The neural basis of language development: Changes in lateralization over age

Olumide A. Olulade[a], Anna Seydell-Greenwald[a], Catherine E. Chambers[a], Peter E. Turkeltaub[a], Alexander W. Dromerick[a], Madison M. Berl[b], William D. Gaillard[b], and Elissa L. Newport[a,1]

[a]Center for Brain Plasticity and Recovery, Georgetown University Medical Center and MedStar National Rehabilitation Hospital, Washington, DC 20057; and [b]Center for Neuroscience and Behavioral Health, Children's National Hospital, Washington, DC 20010

Contributed by Elissa L. Newport, July 17, 2020 (sent for review April 8, 2019; reviewed by Rachel I. Mayberry and Nicholas B. Turk-Browne)

We have long known that language is lateralized to the left hemisphere (LH) in most neurologically healthy adults. In contrast, findings on lateralization of function during development are more complex. As in adults, anatomical, electrophysiological, and neuroimaging studies in infants and children indicate LH lateralization for language. However, in very young children, lesions to either hemisphere are equally likely to result in language deficits, suggesting that language is distributed symmetrically early in life. We address this apparent contradiction by examining patterns of functional MRI (fMRI) language activation in children (ages 4 through 13) and adults (ages 18 through 29). In contrast to previous studies, we focus not on lateralization per se but rather on patterns of left-hemisphere (LH) and right-hemisphere (RH) activation across individual participants over age. Our analyses show significant activation not only in the LH language network but also in their RH homologs in all of the youngest children (ages 4 through 6). The proportion of participants showing significant RH activation decreases over age, with over 60% of adults lacking any significant RH activation. A whole-brain correlation analysis revealed an age-related decrease in language activation only in the RH homolog of Broca's area. This correlation was independent of task difficulty. We conclude that, while language is left-lateralized throughout life, the RH contribution to language processing is also strong early in life and decreases through childhood. Importantly, this early RH language activation may represent a developmental mechanism for recovery following early LH injury.

brain | language | lateralization | development | fMRI

Based on examinations of adults with acquired brain injury, language has long been hypothesized to be lateralized to left-hemisphere (LH) inferior frontal and superior temporal areas in adults (1, 2). This notion of LH language dominance has subsequently received ample support from examinations of brain structure revealing asymmetries in the size of temporal and inferior frontal brain areas (3–7), behavioral studies demonstrating a right ear and right visual field (thus LH) advantage for processing of language stimuli (8, 9), studies showing language impairments induced by experimental disruption of the left hemisphere (10–13), and electrophysiological stimulation and functional neuroimaging studies revealing left-lateralized activation during language tasks (14–20). While handedness modulates language dominance to some degree (21–23), it is clear that language function (especially the production and processing of syntax) is left hemisphere-dominant in the vast majority of adults.

What is less clear is whether this strong left dominance is present at birth or appears gradually during development. Here the existing bodies of evidence, each extensive, suggest apparently inconsistent findings. On one hand, clinical findings suggest that the LH and right hemisphere (RH) are equipotential and equally involved in language early in life, with gradually increasing involvement of the left hemisphere through childhood (24, 25). This generalization was initially based on observations from clinical studies of 102 individuals with early unilateral brain injuries, some of whom underwent subsequent hemispherectomies (24). As summarized by Lenneberg (25), in children whose lesions occurred prior to the onset of speech (18 to 24 mo), left- and right-hemisphere injury equally resulted in delayed or absent language development (47% of children for LH and 51% for RH). Hemispherectomy performed before age 13 resulted in permanent aphasia only for a small percentage of patients regardless of hemisphere (6% for LH and 12% for RH patients), whereas hemispherectomy in adults resulted in permanent aphasia for all LH patients and not a single RH patient. These and subsequent clinical observations (26) suggest that the adult lateralization pattern is not yet established in young children and that both hemispheres participate equally in language during early development.

On the other hand, anatomical and functional evidence in healthy children suggests that left dominance is present even in infants and neonates. For example, the above-mentioned anatomical asymmetries have been observed in infant brains as well (4, 27, 28), and language-evoked brain activation measured using electroencephalography (EEG) evoked potentials, near-infrared spectroscopy (NIRS), and functional MRI (fMRI) is left-lateralized early in life (15, 29–38). While some studies have found that language is somewhat less lateralized in children and that left lateralization increases from childhood to adulthood (31–33, 35), others have found no differences in lateralization between children and adults (34, 39) and, among those who did, there is disagreement about whether the increase in left lateralization is driven by a decrease of RH activation or an increase of LH activation. Either way, the common finding of left lateralization of language early in life is

Significance

Two types of evidence suggest different pictures of how language is represented in the brain during development. Studies of the anatomy, physiology, and fMRI activation of the two hemispheres show that language is lateralized to the left hemisphere from birth. In contrast, damage to the left versus right hemisphere in young children is equally likely to result in language impairment, suggesting that language is bilaterally represented in early development. The present study resolves this paradox by examining fMRI language activation in different ways. While group averages show LH lateralization throughout development, young children show RH language activation that declines systematically with age. Most important, this RH activation in children represents a possible mechanism for explaining language recovery following early stroke.

Author contributions: O.A.O., W.D.G., and E.L.N. designed research; O.A.O., A.S.-G., M.M.B., and W.D.G. performed research; O.A.O., A.S.-G., C.E.C., P.E.T., and M.M.B. analyzed data; A.W.D. discussed analysis and interpreted significance of results; E.L.N. supervised data analysis; and A.S.-G., C.E.C., P.E.T., A.W.D., M.M.B., W.D.G., and E.L.N. wrote the paper.

Reviewers: R.I.M., University of California San Diego; and N.B.T.-B., Yale University.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: eln10@georgetown.edu.

First published September 8, 2020.

seemingly at odds with the clinical observation that early damage to either hemisphere is equally likely to result in language deficits.

This apparent conflict may at least partly be caused by the way laterality data from functional imaging studies are traditionally analyzed. The vast majority of studies to date have investigated language lateralization using a laterality index (LI) comparing LH and RH activation. While the precise method for quantifying LH and RH activation and computing the LI differs across studies (40, 41), the LI in general compares the difference between LH and RH activation with the total activation. LIs near 0 indicate bilateral and equal activation in the two hemispheres, whereas positive LIs indicate left lateralization. This measure allows quantification of lateralization regardless of absolute activation levels, which may be influenced by age or other factors that may differ over age, such as stimulus complexity, effort, and task difficulty (17, 42–45). However, because the LI is a difference score, potentially important information may be lost. An increase in LI with age could reflect a decrease of RH activation, an increase of LH activation, or both; the extent to which the LH and the RH are each involved in language processing is not directly reflected in LI scores. Activation maps presented along with LI data usually show group data, averaged across participants. Lower RH activation in these group maps does not necessarily reflect lower RH activation in individual participants; it could simply indicate that RH activation is not as consistent across individuals as LH activation.

For these reasons, the present fMRI study focuses on activation patterns in each hemisphere rather than on lateralization per se, and on analyses of individual activation maps rather than group averages, through childhood to young adults. Doing this requires a task that activates the brain strongly and reliably enough to compute activation maps for individuals and also requires that task difficulty be kept fairly stable over age.

We examined brain activation in 39 children (ages in years;months from 4;6 through 13;0) and 14 adults (ages 18;5 through 29;2) during a well-studied and highly reliable language-comprehension task called the Auditory Description Decision Task (ADDT) (37, 46). All participants were right-handed, neurologically healthy native speakers of English and had IQs in the normal or above-normal range. During forward-speech blocks, participants heard sentences defining a noun and pushed a button if the sentence was correct (e.g., "a big gray animal is an elephant"). During the reverse-speech control condition, participants heard the same audio files played backward and pushed a button if they heard a beep at the end of the file. To equate task difficulty across ages, all target words were selected from the 5,000 most common words in print; within this range, the word frequency of the target nouns was varied across the three child age groups, based on norms from age-appropriate reading materials [targets were chosen from the 2,500 most frequent words for the youngest age group, from the 3,500 most frequent words for the middle age group, and from the 5,000 most frequent words for the oldest group (47)], while leaving the syntactic frame of the sentences the same. Response accuracy was high and did not differ significantly between the child age groups (Table 1), suggesting that this manipulation matched task difficulty fairly well. Adults received the same stimuli as the oldest children.

Several analyses contrasting the fMRI response to forward vs. reverse speech were performed. For comparability with other studies, we first did a random-effects group analysis to generate and compare language activation maps for our four age groups: young (4;6 through 6;8, $n = 10$), middle (7;4 through 9;10, $n = 14$), and oldest (10;0 through 13;0, $n = 15$) children and adults (18;5 through 29;2, $n = 14$). We then turned to individual activation maps in order to determine significant activation for each individual participant in LH language areas and their RH homologs. To examine the consistency of these activation patterns across individuals in each group, we generated penetrance maps showing the percentage of subjects in each group who showed significant activation in each brain voxel and determined the percentage of subjects with significant activation in anatomically defined regions of interest (ROIs) in left and right inferior frontal and superior temporal cortex. Finally, we performed a whole-brain analysis across all participants to identify brain areas in which language activation was correlated with age. While the first of these analyses is frequently performed on imaging data, the other analyses are less common and are designed to illuminate the patterns of activation that appear in substantial numbers of children and that may change systematically over age but may not appear in a traditional mean activation analysis.

Results

Random-Effects Group Analyses. In order to establish the consistency of our findings with others' in the literature, we began with an analysis of fMRI activation for forward versus reverse speech over the four age groups. In agreement with findings of other studies, in this group analysis all age groups show strongly left-lateralized language activation, predominantly in LH inferior frontal and superior temporal cortex (Fig. 1). Right cerebellar activations are also present in all age groups. A group x condition ANOVA on the three child age groups showed no significant effect of age anywhere in LH language areas or their RH homologs at our usual single-voxel and cluster-size thresholds of $P < 0.001$ and $k < 0.05$. Including the adult group produced a significant age effect in both LH and RH superior temporal gyrus; adults show less activation in these regions than any of the child age groups. This is likely due to our use of the same word frequency level for adult participants as for the oldest children, resulting in a slightly easier task and higher level of accuracy for this group. Importantly, however, there were no significant changes in lateralized activation over age shown in these analyses.

Individual Activation Maps. The ADDT produces such strong and reliable activation for forward vs. reverse speech that we do not have to focus only on group analyses but in addition can examine individual activation maps (smoothed with an 8-mm full width at half-maximum [FWHM] Gaussian kernel and thresholded using an uncorrected single-voxel threshold of $P < 0.001$ and a cluster-extent threshold given which family-wise error rate [FWE] was <0.05). Fig. 2 shows example individual activation maps for subjects in each of the age groups. In contrast to the group analyses, the individual activation maps show striking differences over age, particularly in RH activation. While the majority of adults had little to no RH activation, RH activation was strong and extensive in many of the youngest children, in some cases comparable to that of the LH. When RH activation occurred, it was roughly the mirror image of that observed in the LH, with peaks in inferior frontal and superior temporal cortex. We therefore used several

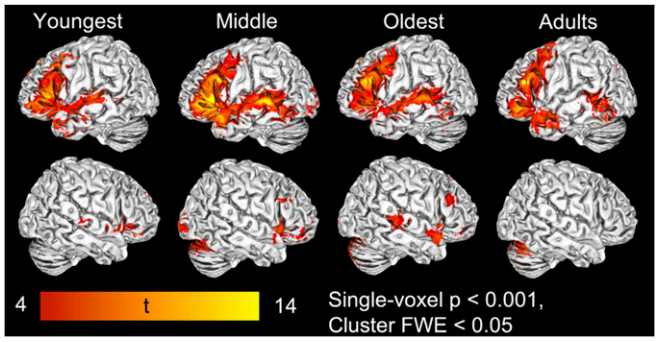

Fig. 1. Random-effects group activation maps contrasting activation evoked by listening to forward (comprehensible) and backward (incomprehensible) speech. All age groups (youngest, $n = 10$; middle, $n = 14$; oldest, $n = 15$; adults, $n = 14$) show strong left lateralization.

Table 1. Participant age groups and performance measures

	Youngest (4;6 to 6;8)	Middle (7;4 to 9;10)	Oldest (10;0 to 13;0)	Adults (18;5 to 29;2)	Group differences (children only)	Group differences (all groups)
N (females)	10 (6)	14 (8)	15 (7)	14 (7)		
Mean FSIQ (SD)	113.80 (14.81)	117.14 (14.73)	112.93 (10.35)	120.71 (6.40)	$F(2,36) = 0.4$, $P = 0.68$ NS	$F(3,49) = 1.23$, $P = 0.31$ NS
Mean VIQ (SD)	106.75 (7.80)	118.57 (17.08)	114.60 (11.99)	121.93 (7.55)	$F(2,36) = 0.58$, $P = 0.57$ NS	$F(3,49) = 1.33$, $P = 0.28$ NS
% correct (SD)	86.30 (11.04)	89.05 (9.99)	93.33 (7.35)	99.64 (0.98)	$F(2,36) = 1.81$, $P = 0.18$ NS	$F(3,49) = 6.58$, $P < 0.001$
RT in msec (SD)	2823 (187)	2945 (466)	3009 (99.59)	3053 (122)	$F(2,36) = 1.15$, $P = 0.33$ NS	$F(3,49) = 1.63$, $P = 0.19$ NS

FSIQ = Full scale IQ; VIQ = Verbal IQ; RT = reaction time.

additional methods of analysis designed to examine whether these differences were reliable within and across the age groups. While significance in group activation analyses requires consistent activation across most or all participants in the group, other types of analyses can better reveal activations that are consistent across a sizeable portion of the group or that relate closely to age.

Penetrance Maps. Because our task permits us to construct an activation map for each individual, we can overlap these maps to examine how many individuals in each age group show significant activation at each voxel throughout the brain. As already revealed by the group activation maps, the proportion of participants with overlapping activations was highest in the usual LH language areas: left superior temporal cortex (STC) and left inferior frontal cortex (IFC) for all age groups (Fig. 3). Importantly, however, and not revealed by the group activation maps, 50% of the youngest children also showed significant overlapping activation in RH IFC (Brodmann area [BA] 45/47) and STC (BA 21/22). This proportion decreased across age groups, reaching 0% in the adults. Table 2 presents the peak overlap percentages and locations for all age groups.

Proportion of Subjects with Significant RH Activation. Because a reduction in map overlap could simply reflect sparser activations or larger variability in where activation occurs, we also calculated the proportion of subjects in each group who showed significant activation anywhere in anatomically defined ROIs in left and right inferior frontal cortex and superior temporal cortex. All participants showed significant activations in left IFC and left STC. Fig. 4 shows the proportion of participants in each group whose individual activation maps also showed significant activation in right IFC and in right STC. This proportion is at 90% in the youngest group but drops dramatically and systematically across the age groups.

Correlation of fMRI Activation with Age: Whole-Brain Analysis. We also conducted a whole-brain analysis of covariance (ANCOVA) searching for voxels whose activation correlated negatively with age across all subjects. This analysis revealed large clusters along left and right STC (BA 21/22) and an additional cluster including portions of the right insula and IFC (BA 13/44/45/47). When the analysis was confined to the children only, only the right IFC correlation remained (also see ref. 37). The bilateral STC correlations were driven predominantly by a group difference between children and adults (with lower STC activation bilaterally in adults; see above) rather than a consistent activation decrease over age. In contrast, the age correlation with activation in the right insula and right IFC appears to be a genuine and consistent decline in activation across age. These results were virtually identical when task performance (proportion of correct responses) and/or reaction times were included as covariates of no interest in the analysis and therefore are unlikely to be driven by differences in task error rates, difficulty, or effort across age. Fig. 5 shows the right IFC cluster displaying a significant decrease of language activation over age, along with scatterplots illustrating the correlation.

Correlation of fMRI Activation with Age: ROI Analysis. In addition to searching the whole brain for correlations between age and activation, we also tested for correlations between age and activation in RH "language areas" (quantified as the number of voxels with significant individual map activation in RH areas mirroring the LH language network, but similar results are obtained for any of several different ways of quantifying RH activation). A significant negative correlation between RH activation and age was observed regardless of whether performance was partialled out ($r = -0.33$ with and $r = -0.40$ without correction for performance). No such correlation was observed between age and LH activation ($r = -0.09$ with and $r = -0.12$ without correction for performance).

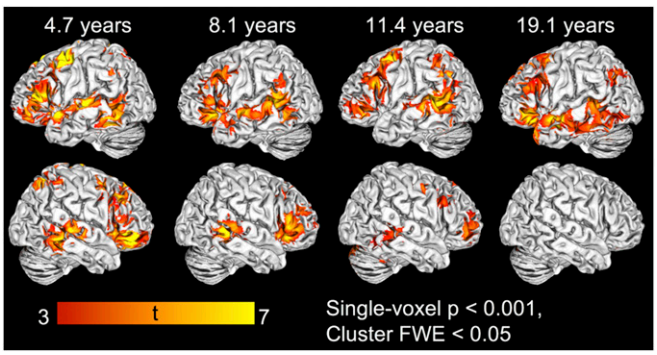

Fig. 2. Examples of individual activation maps in each of the age groups. Strong activation in right-hemisphere homologs of the left-hemisphere language areas is evident in the youngest children, declines over age, and is entirely absent in most adults. The penetrance maps in Fig. 3 summarize this finding across all subjects.

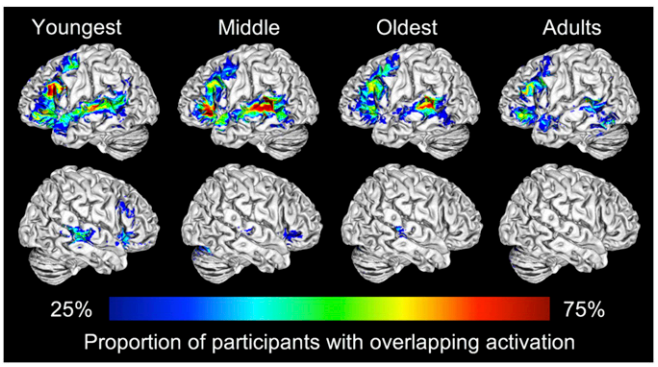

Fig. 3. Penetrance maps, illustrating the proportion of participants in each age group with significant activation in each voxel. Among the youngest participants ($n = 10$), the overlap in some parts of the right hemisphere reaches 50%, but there is no RH overlap among the adults ($n = 14$). The middle ($n = 14$) and oldest ($n = 15$) groups of children show intermediate amounts of RH overlap. (See also Table 2.)

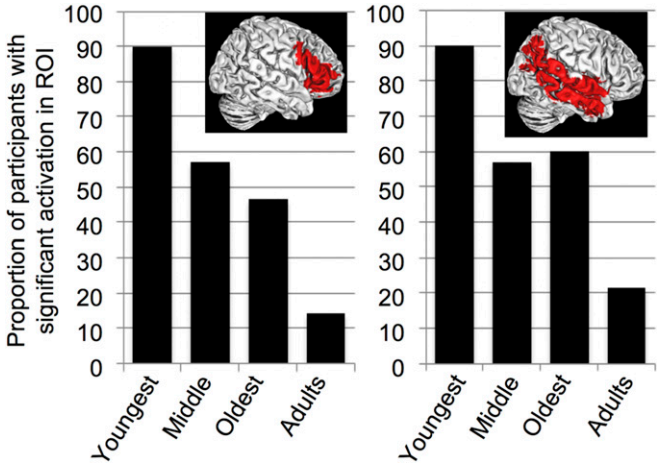

Fig. 4. Proportion of participants in each group (youngest, $n = 10$; middle, $n = 14$; oldest, $n = 15$; adults, $n = 14$) whose individual activation maps showed significant activation in anatomically defined ROIs in right IFC (*Left*) and in right STC (*Right*).

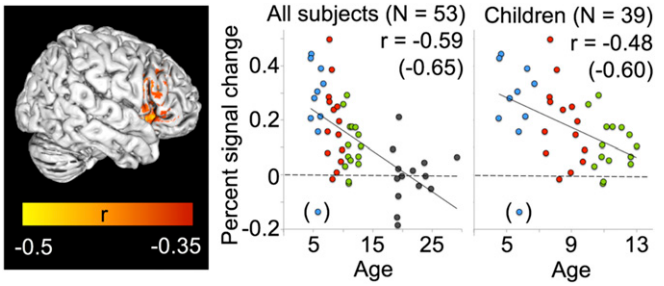

Fig. 5. In a whole-brain analysis, voxels showing a negative correlation between age and language activation appear in right IFC. The cluster shown was identified by looking for correlations across all participants ($n = 53$), but the correlation still holds when the analysis is conducted only on the children ($n = 39$) and adults are excluded from the analysis. It also holds when task performance and reaction times are partialled out. The scatterplots show signal change for forward–backward speech over age, with the age groups indicated in different colors (youngest children in blue, $n = 10$; middle children in red, $n = 14$; oldest children in green, $n = 15$; adults in black, $n = 14$). Correlations are indicated with and without an activation outlier (in parentheses) in the youngest group.

No Correlation between RH Activation and Performance. In healthy participants and stroke patients, increasing sentence complexity, task difficulty, or errorful performance can produce increasing RH activation (17, 43–45). As described above, the design of the ADDT attempted to equate sentence and task difficulty and response accuracy by keeping the sentence frame the same but altering target word frequency in accord with reading material norms across age groups; indeed, response accuracy was high (above 80% correct) and roughly equal across age groups. Nonetheless, to be certain that the higher level of RH activation in younger children was not due to these factors, we examined the correlation between RH activation and task performance. This correlation ($r = -0.25$) did not reach significance. The results were virtually identical when the correlation analyses were repeated without the adult participants.

Discussion

Our analyses demonstrate that, while at the group level even young children show left-lateralized language activation, a large proportion of the youngest children also show significant activation in right-hemisphere homologs of the left-hemisphere language areas (most notably, right-hemisphere inferior frontal and superior temporal cortex). This proportion decreases with age, as does the RH activation itself. No similar activation decrease with age was observed for the LH. These findings suggest that the neural network for language processing in young children includes these RH homologs as well as the typical LH language areas, and that this distribution becomes systematically more localized to the LH over age. While some previous studies have shown that the laterality index increases over age (31–33), their emphasis has been on the fact that language is left-lateralized throughout development. Moreover, the laterality index (measuring the ratio of LH to RH activation) does not reveal whether this change is due to changes in the activation of the LH, the RH, or both. In the present study, we conducted analyses to reveal the details of these activations separately, showing that the developmental changes for our sentence-processing task involve changes only in RH activation.

In healthy adults and adult stroke patients, RH activation increases with sentence complexity, task difficulty, and errorful performance (17, 43–45). However, task complexity and errorful performance in our task do not appear to be the cause of the greater RH activity found here in young children. The ADDT was designed to maintain task difficulty at a comparable level across age groups by maintaining the same sentence frame but changing the target word frequencies used over age groups. As a result, accuracy was high and fairly constant across age. In addition, we found no significant correlation between task performance and RH activation, and the negative correlation between RH activation and age remained when task performance and reaction time were partialled out. This suggests that the RH activations observed here reflect genuine functional involvement of the RH in language processing in young children and are not artifacts of our specific task.

One more possibility is that RH activation in young children is due to developmental differences in the way they process the sentences we have presented in the ADDT and not to changes in neural organization. That is, perhaps RH activity in young children arises not from early RH involvement in the aspects of language processing (lexical activation, syntactic parsing, and semantic interpretation) that will later become localized to the LH, but rather to young children's greater reliance on typical RH processes such as attention to emotional prosody or intonation in our task. For several reasons we do not believe this underlies our results. First, target sentences in the ADDT were spoken in a flat, neutral intonation, and "true" sentences had the same intonation contours as "false" sentences; reliance on prosodic information would therefore not explain the high levels of accuracy that even the youngest children achieved. In addition, children in the broad age range that

Table 2. Peak activation overlap in left and right IFC and STC for each age group

	Youngest	Middle	Oldest	Adults
Left IFC (peak MNI coordinates)	80% (−54,18,17)	86% (−48,25,−1)	60% (−55,19,22)	71% (−52,27,18)
Right IFC (peak MNI coordinates)	50% (51,21,−11)	36% (52,25,−6)	20% (54,34,14)	No overlap
Left STC (peak MNI coordinates)	80% (−58,−25,0)	100% (−59,−34,1)	73% (−63,−31,0)	71% (−57,−41,1)
Right STC (peak MNI coordinates)	50% (53,-37,3)	29% (62,−22,0)	40% (65,−30,−1)	No overlap

we tested (ages 5 through 13) are beyond the early stages of acquisition in which greater reliance on prosody than syntax and semantics might be expected. Studies of sentence processing in children in the age range of our younger subjects show that they do use lexical and syntactic information for sentence comprehension, sometimes showing less ability to use prosodic information than older children and adults (48, 49). Finally, the reverse-speech condition included the same variations in pitch, duration, and loudness that characterize prosody in the forward-speech condition, so neural activation to these properties would be subtracted in our analyses.*

We therefore believe that the higher levels of RH activation in a sentence-processing task, and the slow decline in this activation over development, are reflections of actual changes in the neural distribution of language functions and not merely developmental changes in sentence-comprehension strategies used in the ADDT. We suggest two accounts of what might underlie these changes. One possibility is that children are becoming more efficient language processors long after they have acquired the basic structures of their language, and developmental changes in expertise and the efficiency of language processing may be accompanied by greater localization of these functions in the brain. Another possibility, perhaps related, is that the language system increasingly separates and neurally segregates its subcomponents over this period (see below). Further research on the efficiency of language processing through development, and on the effects of expertise and maturation on the neural organization of skills, is needed.

Conclusions and Future Directions

Our results suggest that, early in life, right-hemisphere homologs of left-hemisphere inferior frontal and superior temporal language areas functionally contribute to sentence processing. This involvement of the right hemisphere in sentence processing decreases throughout childhood and early adolescence and is mostly absent by early adulthood. These findings further suggest that, if we were able to present the same task and conduct the same analyses with even younger children, it is likely that we would see even greater functional involvement of the RH in language processing than we see in our youngest participants. There are some indications of this in studies of infants (29, 50, 51), but most studies of infant language activation have focused on whether there is any LH lateralization at young ages; further research is needed to examine the relative degree and extent of RH activation in infants and very young children.

It is important to note that, when we refer to the activation from our task reflecting the neural distribution for "language processing," we use this term in the conventional sense of focusing specifically on sentence processing, including sentence structure and meaning. These are the particular aspects of language that are conventionally thought of as left-lateralized in almost all healthy adults and, in previous research, in young children as well. There are, however, other language skills, including the processing of prosody (for example, vocal emotion and intonational contrasts between statements and questions) and the processing of discourse devices (for example, story coherence, metaphor, and the like) that are hypothesized to be right-lateralized in healthy adults. These abilities have seldom been studied in fMRI tasks with young children, but the available evidence suggests that infants and children also show RH lateralization to prosody when analyses are done via the usual group-averaging methods (52–54). In ongoing work, we are using our own analyses to ask whether young children, while right-lateralized for prosody, also show more LH activity than adults on the same task, in complement to our present results on sentence processing. Together with the present findings, such results suggest that many aspects of language are more broadly represented in young children and then gradually become segregated into their LH and RH components with increasing age and expertise in language.

We are also using the same sentence-processing task employed with healthy children and adults in the present work (the ADDT) to examine language activation in teenagers and young adults who have had a major LH stroke at birth (55). In accord with previous studies of language outcomes after LH perinatal stroke or early chronic LH seizure disorders (11, 24–26, 38, 46, 56, 57), we are finding that these RH homologs are surprisingly capable of supporting basic and even complex language skills when the LH language network is damaged early in life (55). As the account of such findings, it is often suggested that the young brain is highly plastic and that, for this reason, language can "reorganize" to areas that are not ordinarily the substrates for adult language. In contrast with this view—in light of the present results and in accord with earlier suggestions by Lenneberg (25) and Berl et al. (37)—we hypothesize that the normal involvement of the RH homologs in language processing during very early childhood may permit the maintenance and enhancement of RH language development when the LH is injured. In this hypothesis, the declining involvement of the RH in sentence processing over development—and the increasing dedication of the RH homolog areas to processing other aspects such as prosody—may explain why language recovery after LH stroke is not as good in adults as it is in children. In ongoing work, we are further investigating this "Developmental Origins hypothesis" (58) as an explanation of the successful use of the RH for language processing after early LH damage.

Materials and Methods

Participants. Our participants were neurologically healthy children ranging in age from 4 to 13 who served as control participants in a previously reported study of language reorganization in epilepsy and as participants in a study of regional developmental change (37, 38, 46), and neurologically healthy young adults run on the same scanner using the same paradigm. The study was approved by the Institutional Review Boards at the Children's National Medical Center and at Georgetown University Medical Center; all participants provided consent (adults) or parental consent and child assent (children). In addition to the fMRI task, all participants completed an IQ test [Differential Abilities Scale for children ages 4 to 5 (59); Wechsler Abbreviated Scale of Intelligence (WASI) for children ages 6 to 12 (60); WASI II for adults (61)] and provided information regarding language background and personal and familial handedness. The final sample included only right-handed native speakers of English with IQs in the average or above-average range and who had no significant exposure to another language before the age of 4. As shown in Table 1, the groups did not differ significantly in IQ or in reaction time during the in-scanner task. The three child groups also did not differ in in-scanner task accuracy, but the adult group performed at somewhat higher accuracy (likely due to our using the same target words for them as for the oldest children). All participant groups were above 85% correct on the task. Note that LI increases with age for this dataset have already been reported elsewhere (37) (Fig. 3); here, we focus more specifically on how RH activation changes across age.

Magnetic Resonance Imaging. MRI was performed at Georgetown University's Center for Functional and Molecular Imaging using a 3-T Siemens MAGNETOM Trio scanner and a 12-channel head coil. Prior to scanning, participants were trained on lying still and performing the task in a mock scanner setup that provides immediate feedback when the participant moves too much. During the actual scan, participants' heads were stabilized using foam padding, and motion was monitored throughout so that scans could be repeated if too much motion occurred.

Auditory stimuli were presented via Sensimetrics model S14 insert earphones, over which the participants wore Bilsom ear defenders to attenuate scanner noise. Sound intensity was set prior to the scan to the maximum level participants were comfortable with, to ensure that stimuli were clearly audible over the scanner noise. Following a brief localizer scan to allow volume

*In listening to stimuli constructed to test prosody in another experiment, we find that the prosody of words and sentences is identifiable in reverse as well as forward speech, supporting the argument that activation to prosodic processing would be subtracted in the present analyses.

placement with slices parallel to the anterior commissure–posterior commissure (ACPC) line, each participant was scanned using the following sequences.

High-resolution anatomical image. Siemens MPRAGE, 176 sagittal slices; repetition time (TR), 1.9 s; echo time (TE), 2.52 ms; flip angle, 9°; $1 \times 1 \times 1$-mm^3 voxels covering the entire brain. Participants watched an age-appropriate movie during localizer and MPRAGE acquisition.

Functional sequence. Echo-planar imaging in 50 horizontal slices, acquired in descending sequence; TR, 3 s; TE, 30 ms; flip angle, 90°; $3 \times 3 \times 3$-mm^3 voxels covering the entire brain. Functional runs consisted of 100 whole-brain volumes and took 5 min. Runs were repeated if too much motion occurred; analyses include only one run per subject.

In-scanner task. During functional runs, participants performed an auditory description decision task developed by Gaillard et al. (46) (also see refs. 37 and 38). In the forward-speech condition, participants heard short sentences describing a noun (e.g., "a big gray animal is an elephant") and pushed a button at the end of the sentence if the description was correct (which was the case for 70% of the sentences). Task difficulty was adjusted across age groups by selecting nouns at varying levels of word frequency. This ensured that all participants performed the task with high accuracy, avoiding differences across age in activation arising from errors, confusion, or uncertainty. In the reverse-speech control condition, the same audio files were played backward, and participants responded if they heard a beep at the end of the utterance (again, in 70% of the items). This condition controlled for low-level auditory features and also for motor activity related to button pushes. Sentences in both conditions were presented every 3 s, leaving about 1 s of response time after each sentence. A 5-min functional run comprised five blocks each of the forward and reverse conditions presented in interleaved order, starting with the reverse condition. Each block lasted 30 s and contained 10 utterances.

Preprocessing of Neuroimaging Data. After discarding the first three volumes (to allow for saturation of T1 effects) and the last three volumes of each dataset (to keep the number of volumes constant across conditions), neuroimaging data underwent the following preprocessing using SPM8 (Wellcome Trust Centre for Neuroimaging at University College London) (62): 1) slice-time correction, 2) rigid-body motion correction to the first volume of the run, 3) determination of the rigid-body transformation needed to align functional and anatomical images, 4) spatial normalization into Montreal Neurological Institute (MNI) standard space based on automatic segmentation of anatomical images using VBM8 (63), and 5) spatial smoothing with an 8-mm FWHM Gaussian kernel. Functional time series and motion-estimation files were inspected to identify volumes with obvious artifacts or extreme (>0.75 mm) intervolume motion, and a "bad scan" regressor was created to remove the influence of these volumes from the statistical analyses. Datasets with more than 25% bad scans were excluded from further analysis, resulting in the slightly different group sizes reported in Table 1.

Analysis of Neuroimaging Data.

First-level general linear model. Separately for each participant, we fit the time course of each voxel in the dataset using a general linear model with the following predictors: predictors 1 and 2 modeled the stimulation time courses for experimental and control conditions, convolved with a canonical hemodynamic response function, predictors 3 through 8 captured the motion estimations for translation along and rotation about the *x*, *y*, and *z* directions, predictor 9 captured volumes affected by motion and other artifacts, predictor 10 captured the global signal to remove the influence of signal variations affecting the entire brain (such as those caused by subjects taking a deep breath), and predictors 11 through 14 were cosine basis functions (0.5, 1, 1.5, and 2 cosines) serving as a high-pass filter.

Individual language activation maps. Individual "language activation maps" were created by contrasting the resulting beta images for the experimental and control conditions using a *t* test [while correcting for serial autocorrelations using the AR(1) autoregression model]. To apply consistent threshold criteria to all subjects' maps, we decided on a single-voxel threshold ($P < 0.001$) that was the median (and close to the average of 0.0013) of the *P* values corresponding to a false discovery rate (FDR) of less than 5% across all subjects. This was combined with a cluster-extent threshold given which the family-wise error rate at the cluster level was smaller than 5%. (Mean and median of this threshold across subjects coincided at 269 mm^3, with an SD of 57 mm^3.)

Group activation maps. Group activation maps were created separately for each age group by submitting the unthresholded individual activation maps of all subjects in the group to a random-effects analysis. Like the individual maps, group activation maps were subjected to a single-voxel threshold of $P < 0.001$ (which was stricter than the single-voxel threshold corresponding to an FDR of less than 5% for all groups) and a cluster-level threshold given which the FWE was less than 5%. (The threshold cluster size was 209, 231, 249, and 276 mm^3 for the youngest to oldest group, respectively.) We also ran a random-effects analysis incorporating the activation maps of all subjects, regardless of age. This overall activation map was also thresholded at $P < 0.001$, and the cluster-size threshold corresponding to FWE < 0.05 was 341 mm^3. Finally, to determine whether there was a main effect of age, we performed one ANOVA across the four age groups and another ANOVA across only the three child age groups. As before, maps were thresholded using a single-voxel threshold of $P < 0.001$ and a cluster-size threshold corresponding to FWE < 0.05, which corresponded to a cluster size of 334 mm^3 for the ANOVA across all age groups and 314 mm^3 for the ANOVA across the child age groups.

Penetrance maps. Penetrance maps were computed for each subgroup by determining the percentage of individual activation maps showing significant activation in each voxel. It is important to note the potential differences of what can be revealed by group activation maps as compared with penetrance maps. Both penetrance maps and group activation maps will reveal areas in which there is consistent activation across most or all subjects' individual maps. Thus we expected both map types to highlight LH frontal and temporal cortex. However, they differ in what they reveal about areas of low and/or inconsistent activation: Group activation maps will show activation in areas where there is consistent activation across subjects' individual maps, even if it is slightly subthreshold for all of the individuals. In contrast, only penetrance maps will highlight areas in which activation may not be consistent across all subjects but is significant in a sizeable proportion of them.

Anatomical regions of interest. Anatomical ROIs were generated from the WFU PickAtlas (64), using the IFG (inferior frontal gyrus) label for the inferior frontal ROI and the labels for Brodmann areas 22, 21, and 39 for the temporal ROI, following Berl et al. (38).

Whole-brain correlation analysis. To identify voxels whose activation showed a significant positive or negative correlation with age, we performed a whole-brain ANCOVA across all participants, as well as across the children only. (We also repeated the analyses while including task performance and/or reaction times as covariates of no interest, which did not appreciably change the results.) The resulting maps were thresholded at a single-voxel threshold of $P < 0.005$, combined with a cluster-size threshold of 802 mm^3, corresponding to FWE < 0.05.

Figures. Figures were created using Mango software (65), with statistical maps overlaid on a surface rendering of the Colin27 T1 image in MNI space. To better illustrate activations falling into sulci (e.g., superior temporal sulcus [STS]), activations were projected up to 10 mm outward onto the surface.

Data Availability. The MRI data reported in this paper have been deposited in the Open Science Framework, http://osf.io/zkpv2.

ACKNOWLEDGMENTS. We thank our reviewers for very helpful comments. This work was supported by funds from Georgetown University and MedStar Health and from the Feldstein Veron Innovation Fund to the Center for Brain Plasticity and Recovery; NIH Grants K18DC014558 (to E.L.N.), K23NS065121 (to M.M.B.), R01NS244280 (to W.D.G.), and R01DC016902 (to E.L.N. and W.D.G.); a KL2 grant (to A.S.-G.) from NIH TL1TR001431 Georgetown–Howard Universities Clinical and Translational Science Award; and M01RR020359 and P30HD040677 to the Intellectual and Developmental Disabilities Research Center U54 HD090257 at the Children's National Health System (which also includes Georgetown University).

1. P. M. Broca, Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bull. Soc. Anat.* **6**, 330–357 (1861).
2. C. Wernicke, *Der Aphasische Symptomencomplex—Eine Psychologische Studie auf Anatomischer Basis*, (Cohn & Weigert, Breslau, Germany, 1874).
3. N. Geschwind, W. Levitsky, Human brain: Left-right asymmetries in temporal speech region. *Science* **161**, 186–187 (1968).
4. J. A. Wada, R. Clarke, A. Hamm, Cerebral hemispheric asymmetry in humans. Cortical speech zones in 100 adults and 100 infant brains. *Arch. Neurol.* **32**, 239–246 (1975).
5. G. Falzi, P. Perrone, L. A. Vignolo, Right-left asymmetry in anterior speech region. *Arch. Neurol.* **39**, 239–240 (1982).
6. A. L. Foundas, C. M. Leonard, R. Gilmore, E. Fennell, K. M. Heilman, Planum temporale asymmetry and language dominance. *Neuropsychologia* **32**, 1225–1231 (1994).
7. A. L. Foundas, K. F. Eure, L. F. Luevano, D. R. Weinberger, MRI asymmetries of Broca's area: The pars triangularis and pars opercularis. *Brain Lang.* **64**, 282–296 (1998).
8. W. H. Moore Jr., W. E. Weidner, Bilateral tachistoscopic word perception in aphasic and normal subjects. *Percept. Mot. Skills* **39**, 1003–1011 (1974).

9. W. H. Moore Jr., W. E. Weidner, Dichotic word-perception of aphasic and normal subjects. *Percept. Mot. Skills* **40**, 379–386 (1975).
10. J. Wada, T. Rasmussen, Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance. *J. Neurosurg.* **17**, 266–282 (1960).
11. T. Rasmussen, B. Milner, The role of early left-brain injury in determining lateralization of cerebral speech functions. *Ann. N. Y. Acad. Sci.* **299**, 355–369 (1977).
12. A. Pascual-Leone, J. R. Gates, A. Dhuna, Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. *Neurology* **41**, 697–702 (1991).
13. J. T. Devlin, K. E. Watkins, Stimulating language: Insights from TMS. *Brain* **130**, 610–622 (2007).
14. G. Ojemann, J. Ojemann, E. Lettich, M. Berger, Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J. Neurosurg.* **71**, 316–326 (1989).
15. D. L. Molfese, R. B. Freeman Jr., D. S. Palermo, The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain Lang.* **2**, 356–368 (1975).
16. J. R. Binder *et al*., Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. *Arch. Neurol.* **52**, 593–601 (1995).
17. M. A. Just, P. A. Carpenter, T. A. Keller, W. F. Eddy, K. R. Thulborn, Brain activation modulated by sentence comprehension. *Science* **274**, 114–116 (1996).
18. K. Stromswold, D. Caplan, N. Alpert, S. Rauch, Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* **52**, 452–473 (1996).
19. S. Y. Bookheimer *et al*., A direct comparison of PET activation and electrocortical stimulation mapping for language localization. *Neurology* **48**, 1056–1065 (1997).
20. J. A. Springer *et al*., Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain* **122**, 2033–2046 (1999).
21. S. Knecht *et al*., Handedness and hemispheric language dominance in healthy humans. *Brain* **123**, 2512–2518 (2000).
22. I. E. Sommer, A. Aleman, M. Somers, M. P. Boks, R. S. Kahn, Sex differences in handedness, asymmetry of the planum temporale and functional language lateralization. *Brain Res.* **1206**, 76–88 (2008).
23. J. P. Szaflarski *et al*., Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology* **59**, 238–244 (2002).
24. L. S. Basser, Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. *Brain* **85**, 427–460 (1962).
25. E. Lenneberg, *Biological Foundations of Language*, (Wiley, New York, 1967).
26. E. Bates *et al*., Differential effects of unilateral lesions on language production in children and adults. *Brain Lang.* **79**, 223–265 (2001).
27. S. F. Witelson, W. Pallie, Left hemisphere specialization for language in the newborn. Neuroanatomical evidence of asymmetry. *Brain* **96**, 641–646 (1973).
28. J. G. Chi, E. C. Dooling, F. H. Gilles, Left-right asymmetries of the temporal speech areas of the human fetus. *Arch. Neurol.* **34**, 346–348 (1977).
29. G. Dehaene-Lambertz, S. Dehaene, L. Hertz-Pannier, Functional neuroimaging of speech perception in infants. *Science* **298**, 2013–2015 (2002).
30. M. Peña *et al*., Sounds and silence: An optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11702–11705 (2003).
31. J. P. Szaflarski, S. K. Holland, V. J. Schmithorst, A. W. Byars, fMRI study of language lateralization in children and adults. *Hum. Brain Mapp.* **27**, 202–212 (2006).
32. S. K. Holland *et al*., Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage* **14**, 837–843 (2001).
33. S. K. Holland *et al*., Functional MRI of language lateralization during development in children. *Int. J. Audiol.* **46**, 533–551 (2007).
34. A. G. Wood *et al*., Language cortex activation in normal children. *Neurology* **63**, 1035–1044 (2004).
35. R. Everts *et al*., Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Hum. Brain Mapp.* **30**, 473–483 (2009).
36. K. Lidzba, E. Schwilling, W. Grodd, I. Krägeloh-Mann, M. Wilke, Language comprehension vs. language production: Age effects on fMRI activation. *Brain Lang.* **119**, 6–15 (2011).
37. M. M. Berl *et al*., Regional differences in the developmental trajectory of lateralization of the language network. *Hum. Brain Mapp.* **35**, 270–284 (2014).
38. M. M. Berl *et al*., Characterization of atypical language activation patterns in focal epilepsy. *Ann. Neurol.* **75**, 33–42 (2014).
39. W. D. Gaillard *et al*., Developmental aspects of language processing: fMRI of verbal fluency in children and adults. *Hum. Brain Mapp.* **18**, 176–185 (2003).
40. M. Wilke, V. J. Schmithorst, A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *Neuroimage* **33**, 522–530 (2006).
41. P. Chlebus *et al*., fMRI evaluation of hemispheric language dominance using various methods of laterality index calculation. *Exp. Brain Res.* **179**, 365–374 (2007).
42. K. Murphy, H. Garavan, Artifactual fMRI group and condition differences driven by performance confounds. *Neuroimage* **21**, 219–228 (2004).
43. D. A. Fair *et al*., The functional organization of trial-related activity in lexical processing after early left hemispheric brain lesions: An event-related fMRI study. *Brain Lang.* **114**, 135–146 (2010).
44. H. M. van Ettinger-Veenstra *et al*., Right-hemispheric brain activation correlates to language performance. *Neuroimage* **49**, 3481–3488 (2010).
45. J. D. Yeatman, M. Ben-Shachar, G. H. Glover, H. M. Feldman, Individual differences in auditory sentence comprehension in children: An exploratory event-related functional magnetic resonance imaging investigation. *Brain Lang.* **114**, 72–79 (2010).
46. W. D. Gaillard *et al*., Atypical language in lesional and nonlesional complex partial epilepsy. *Neurology* **69**, 1761–1771 (2007).
47. J. B. Carroll, P. Davies, B. Richman, *The American Heritage Word Frequency Book*, (Houghton Mifflin, Boston, MA, 1971).
48. Y. Choi, R. Mazuka, Young children's use of prosody in sentence parsing. *J. Psycholinguist. Res.* **32**, 197–217 (2003).
49. J. Snedeker, S. Yuan, Effects of prosodic and lexical constraints on parsing in young children (and adults). *J. Mem. Lang.* **58**, 574–608 (2008).
50. D. Perani *et al*., Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16056–16061 (2011).
51. L. A. Petitto *et al*., The "perceptual wedge hypothesis" as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain Lang.* **121**, 130–143 (2012).
52. E. Plante, S. K. Holland, V. J. Schmithorst, Prosodic processing by children: An fMRI study. *Brain Lang.* **97**, 332–342 (2006).
53. I. Wartenburger *et al*., The processing of prosody: Evidence of interhemispheric specialization at the age of four. *Neuroimage* **34**, 416–425 (2007).
54. F. Homae, H. Watanabe, T. Nakano, K. Asakawa, G. Taga, The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* **54**, 276–280 (2006).
55. E. L. Newport *et al*., Revisiting Lenneberg's hypotheses about early developmental plasticity: Language organization after left-hemisphere perinatal stroke. *Biolinguistics (Nicos)* **11**, 407–422 (2017).
56. J. R. Booth *et al*., Developmental and lesion effects in brain activation during sentence comprehension and mental rotation. *Dev. Neuropsychol.* **18**, 139–169 (2000).
57. K. Lidzba, M. Staudt, M. Wilke, W. Grodd, I. Krägeloh-Mann, Lesion-induced right-hemispheric language and organization of nonverbal functions. *Neuroreport* **17**, 929–933 (2006).
58. E. L. Newport, B. Landau, W. D. Gaillard, "*The developmental origins hypothesis: A new account of early developmental plasticity*" (Center for Brain Plasticity, 2020).
59. C. Elliot, *Differential Abilities Scale*, (Psychological Corporation: Harcourt Brace, San Antonio, TX, 1990).
60. D. Wechsler, *Manual for the Wechsler Abbreviated Scale of Intelligence*, (Psychological Corporation, San Antonio, TX, 1999).
61. D. Wechsler, *Wechsler Abbreviated Scale of Intelligence*, (Psychological Corporation, San Antonio, TX, ed. 2, 2011).
62. Wellcome Trust Centre for Neuroimaging, SPM8. https://www.fil.ion.ucl.ac.uk/spm/doc/. Accessed 2 October 2014.
63. C. Gaser, VBM8 Toolbox. dbm.neuro.uni-jena.de/vbm8/. Accessed 6 December 2011.
64. J. A. Maldjian, P. J. Laurienti, R. A. Kraft, J. H. Burdette, An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19**, 1233–1239 (2003).
65. Mango, Version 3.6. http://ric.uthscsa.edu/mango. Accessed 6 November 2015.