

Quantification of Head Motion in Children During Various fMRI Language Tasks

Weihong Yuan,^{1*} Mekibib Altaye,² Jen Ret,³ Vincent Schmithorst,¹
Anna W. Byars,⁴ Elena Plante,⁵ and Scott K. Holland¹

¹Cincinnati Children's Hospital Medical Center, Pediatric Neuroimaging Research Consortium, Cincinnati, Ohio

²Cincinnati Children's Hospital Medical Center, Center for Epidemiology and Biostatistics, Cincinnati, Ohio

³Miami University, School of Psychology, Cincinnati, Ohio

⁴Cincinnati Children's Hospital Medical Center, Division of Neurology, Cincinnati, Ohio

⁵University of Arizona, Department of Speech, Language, & Hearing Sciences, Tucson, Arizona

Abstract: *Purpose:* Head motion during functional MRI scanning can lead to signal artifact, a problem often more severe with children. However, the documentation for the characteristics of head motion in children during various language functional tasks is very limited in the current literature. This report characterizes head motion in children during fMRI as a function of age, sex, and task. *Methods:* Head motion during four different fMRI language tasks was investigated in a group of 323 healthy children between the age of 5 and 18 years. A repeated measures ANOVA analysis was used to study the impact of age, sex, task, and the interaction of these factors on the motion. *Results:* Pediatric subjects demonstrated significantly different amounts of head motion during fMRI when different language tasks were used. Word-Picture Matching, the only task that involved visual engagement, suffered the least amount of motion, which was significantly less than in any of the other three tasks; the latter were not significantly different from each other. Further examination revealed that the main effect of language task on motion was significantly affected by age, sex, and their interaction. *Conclusion:* Our results suggest that age, sex, and task are all associated with the degree of head motion in children during fMRI experiments. Investigators working with pediatric patients may increase their success by using task components associated with less motion (e.g., visual stimuli), or by using this large scale dataset to estimate the effects of sex and age on motion for planning purposes. *Hum Brain Mapp* 30:1481–1489, 2009. © 2008 Wiley-Liss, Inc.

Key words: fMRI; head motion; language task; children

Contract grant sponsor: NIH; Contract grant number: RO1-HD38578.

*Correspondence to: Weihong Yuan, Cincinnati Children's Hospital Medical Center, Pediatric Neuroimaging Research Consortium, 3333 Burnet Ave., Cincinnati, OH 45229.

E-mail: weihong.yuan@cchmc.org

Received for publication 26 April 2007; Revised 28 April 2008; Accepted 29 April 2008

DOI: 10.1002/hbm.20616

Published online 17 July 2008 in Wiley InterScience (www.interscience.wiley.com).

© 2008 Wiley-Liss, Inc.

INTRODUCTION

One of the fundamental limitations in conducting fMRI studies is the sensitivity of the fMRI method to the effects of involuntary head motion during volume acquisition in EPI scanning. The resulting motion artifact can obscure brain activation and invalidate the statistical parameters computed from the data [Grooten et al., 2000; Liao et al. 2006].

Many studies have proposed techniques to proactively restrain the subjects' head position during scanning and/

or retrospectively correct for motion using coregistration algorithms in postprocessing [Biswal et al., 1997; Friston et al., 1995; Woods et al., 1998]. More recent efforts include using motion data as regressor in GLM analysis [Schmithorst and Holland, 2007] and extracting the motion artifact as an independent component [Liao et al., 2006]. However, none of these efforts are able to completely eliminate or “undo” the motion and the motion artifact resulting from long scanning time, discomfort, distractibility, or other factors. On the other hand, reports regarding the characteristics of head motion (for adults or children) and the related confounding factors are rather limited. Hoeller et al. [2002] investigated the impact of motion artifact on the fMRI results of sensorimotor cortical activations. They found that differences between motor tasks led to significantly different motion artifact, while no motion artifact was observed with sensory tasks. Most relevant to our interest is a study conducted by Seto et al. [2001] regarding the characteristics of head motion measured in a battery of motor tasks commonly used in fMRI paradigms. It was found that, for the adult subjects under study, the degree of motion was strongly dependent on the subject group (stroke patient vs. healthy adult) and task conditions (hand gripping, ankle flexion, etc). These studies, although different from the language tasks we are interested in, provide an important indication about the factors (age, task) that are likely to be associated with head motion.

Motion is more problematic in children than in adults and is related to age, sex, attention span, and various neuropsychological and cognitive characteristics that might be associated with development [Byars et al., 2002; Kotsoni et al., 2006]. The motivation for this study comes from the belief that an in-depth understanding of the characteristics of motion during various language functional tasks will lead to better experimental designs and improved strategies for analysis of fMRI data from children. This goal is important in light of the growing applications of fMRI in pediatric studies and its increasing translation into clinical applications for children with various neurological disorders.

In this study, we retrospectively analyzed head motion data obtained from a large number of healthy children while performing four different language tasks. Our aim is to investigate how the nature of different language tasks, as well as the subject’s age and sex, affect the amount of head motion during fMRI scanning. We also examine the interaction of these factors, i.e., whether or not the influence of one factor depends on the other factors.

METHODS

A total number of 422 children between the age of 5 and 18 years (192 girls, 230 boys) participated in a large-scale fMRI study of language development (NIH-R01-HD38578) between 2000 and 2005. Of these subjects, 323 (155 girls, 168 boys) finished all the four language tasks in one

fMRI session, and their data were used in our statistical analysis.

MRI/fMRI scans were all performed on a 3 Tesla Bruker Biospec 30/60 MRI scanner (Bruker Medizintechnik, Karlsruhe, Germany). fMRI scan parameters were: TR/TE = 3,000/38 ms, BW = 125 kHz, FOV = 25.6 × 25.6 cm, matrix = 64 × 64; slice thickness = 5 mm. Details for the procedure can be found in Holland et al., 2007. Image reconstruction and postprocessing were conducted using Cincinnati Children’s Hospital Image Processing Software (CCHIPS[®]), an IDL[®]-based program (ITT Visual Information Solutions, Boulder, CO) [Schmithorst et al., 2000].

This study was approved by the Institutional Review Board at the Cincinnati Children’s Hospital Medical Center (CCHMC). Written consent was obtained from the parents, and the subjects gave either verbal or written assent.

Prior to the fMRI session, children were given an orientation to the MR scanning process that consisted of watching an 8-min video taped introduction to the procedure followed by specific training for each of the language tasks. The training was done at a computer workstation running the same paradigms presented during the fMRI scanning. Finally, young children were given special preparation and desensitization to the MRI scanner itself that consisted of increasing duration of time inside the scanner room performing experiments with the magnetic field and non-ferromagnetic metallic objects in addition to multiple entries into the magnet bore both on and off of the bed [Byars et al., 2002].

The Language Tasks

The fMRI tasks used in the study were four child-friendly language tasks, namely, Syntactic Prosody, Story Processing, Word-Picture Matching, and Verb Generation. These tasks are described in detail elsewhere [Holland et al., 2007] and were designed to tap a range of language skills. For the purposes of this article, it is important to note that the Syntactic Prosody and Word-Picture Matching tasks were designed to tap language skills that develop early in childhood whereas Story Processing and Verb Generation tapped skills thought to have a protracted period of development. Furthermore, all tasks except Word-Picture Matching used auditory stimuli only. Word-Picture Matching employed both auditory and visual stimuli. All tasks except Story Processing required children to make responses during the scan. No response during the scan was required for the Story Processing task. Verb Generation required covert generation of verbs associated with nouns presented aurally. Syntactic Prosody and Word-Picture Matching recorded responses via pressing buttons in the right or left hand.

The fMRI Paradigm

The four language tasks were delivered in a random order during a single fMRI scanning session. The total

time needed to complete the session was between 40 to 60 min, including scanning time for all four functional tasks and anatomical imaging, as well as other necessary preparation time. Those subjects who were not able to complete all four tasks in one session were excluded from the analysis in this report because of the nature of the statistical methodology (repeated measures ANOVA) used in our study.

All four tasks used a similar block-design format. For the Story Processing, Word-Picture Matching, and Verb Generation tasks, each block was 30-sec long and the five cycles of task-control blocks led to total scanning time of 330 sec for each paradigm. For Syntactic Prosody, the task period was 45-sec long, which led to the total scanning time of 405 sec.

The control tasks were designed to account for the sensory/motor and other nonlanguage specific components during language tasks. A tone task was used as the control task for Story Processing and Syntactic Prosody tasks. An image discrimination task and bilateral finger tapping were used in the paradigms for Word-Picture Matching task and Verb Generation task, respectively.

Head Motion Monitoring and Calculation

As an initial compliance screening, operators monitored the subject closely during scanning via closed circuit TV. If subject motion was detected at this stage, the scan was halted. Instructions about the importance of remaining still were repeated via the intercom and headphones, and the scan was repeated. As a second level of screening for head motion, the EPI data were reconstructed online and reviewed as a cine loop on the operator's console immediately following completion of each scan, while the next task was administered. Datasets that were clearly contaminated by gross head motion based on visual inspection could be repeated at the end of the series if the subject remained compliant. These methods were designed to reduce the failure rate and improve the overall quality of the data set but did not provide a consistent or quantitative assessment of head motion. Analysis of the failure rates of children performing these tasks is presented elsewhere [Byars et al., 2002].

A quantitative "motion parameter" was developed using the pyramid method of coregistration developed by Thevenaz and Unser [Thevenaz et al., 1998]. Among the series of images in each scan, the first image was used as the reference frame and each of the subsequent frames was tested and evaluated for motion relative to the reference using the transformation matrix calculated by the realignment algorithm. The norm of the 3D displacement was computed for each frame based on the realignment transformation matrix. It should be noted that, if there is rotation, the displacement is not the same for all voxels in the volume. In that case, the maximum voxel displacement is used. The median value of the displacement for the entire time series of frames was calculated and used as the motion

parameter in the present study. This value is calculated in units of pixels with the dimensions of each pixel being $4 \times 4 \times 5$ mm. As described in the Statistical Analysis, predicted marginal means based on the GLM model were used to compare and demonstrate the effect and interaction of the various factors.

Ardekani et al. [2001] tested the reliability of this algorithm in a dataset with known motion and found that it was very accurate. In the present study, the results of the quantitative analysis were compared to the judgment of experienced neuroimaging experts. Four experts in the research group (SKH, JR, VJS, AWB) agreed on a guideline (Appendix) for motion in fMRI time series data. This group then used the guideline to rate a randomly-selected subset of 20 scans by visually examining the cine loop composed from the sequence of 110 EPI volumes comprising each functional image series. Scans were rated on a scale from 1 to 4, where a 1 corresponded to the absence of visually detectable motion and a 4 to motion sufficient to render the data unusable.

In addition to the aforementioned measures, we also examined the reliability of the motion calculation algorithm with a recently developed statistical image analysis method specifically designed for the suboptimal fMRI datasets [Szafarski et al., 2006]. A cost function, the normalized root mean square deviation from reference frame, was calculated on a frame by frame basis as the criterion for the presence of motion. Comparing this cost-function based method with the three-dimensional retrospective analysis as used in the present study, we found a very good correlation (0.87–0.9, not presented in the present report).

Statistical Analysis

Statistical analysis was done using the Statistical Package for Social Sciences, v.12 (SPSS, Chicago, IL). Correlation analysis was used to examine the inter-rater reliability of the expert ratings and their accordance with the motion parameter. Correlation analysis was also used to study the association between head motion and the time course of task/control cycles. A standard ANOVA was used to test the impact of task order as a possible factor that affects head motion.

A General Linear Model, repeated measures, three-way ANOVA was performed with fMRI task as a within-subjects factor, and age and sex as between-subjects factors. The four language tasks represented the four levels for the within-subjects factor. When a factor was significantly associated with the observed motion, a multiple comparison was used to study the influence of the specific level within that factor.

Marginal means, i.e., the weighted average of the conditional means, with weights being the frequency of occurrence in subgroups, were estimated from the model after adjusting for specified between- and within-subject factors. The estimated marginal means are predicted, not observed, and are based on the specified linear model. In our study,

we used the estimated marginal means of head motion to produce line plots (also called interaction plot or profile plot) across different levels of the between subject factors to study the pattern and interaction (which is the reflection of effect) of different factors.

RESULTS

As mentioned earlier, 323 children were able to finish all four language tasks in one session and thus were included

TABLE I. Descriptive statistics of head motion (in units of pixles, each pixel has dimensions of $4 \times 4 \times 5$ mm)

Tasks	Age group	Sex	N	Mean	STD	
Prosody processing	Younger	Girl	52	0.496	0.343	
		Boy	52	0.590	0.613	
		Total	104	0.543	0.497	
	Middle	Girl	60	0.353	0.246	
		Boy	79	0.384	0.312	
		Total	139	0.370	0.285	
	Older	Girl	43	0.274	0.178	
		Boy	37	0.371	0.219	
		Total	80	0.319	0.202	
	Total	Girl	155	0.379	0.281	
		Boy	168	0.445	0.424	
		Total	323	0.413	0.364	
	Story processing	Younger	Girl	52	0.390	0.259
			Boy	52	0.813	1.076
			Total	104	0.602	0.807
Middle		Girl	60	0.341	0.274	
		Boy	79	0.396	0.277	
		Total	139	0.372	0.276	
Older		Girl	43	0.236	0.204	
		Boy	37	0.354	0.262	
		Total	80	0.290	0.239	
Total		Girl	155	0.328	0.257	
		Boy	168	0.516	0.667	
		Total	323	0.426	0.520	
Word-picture matching		Younger	Girl	52	0.330	0.187
			Boy	52	0.325	0.205
			Total	104	0.328	0.195
	Middle	Girl	60	0.308	0.197	
		Boy	79	0.323	0.200	
		Total	139	0.317	0.198	
	Older	Girl	43	0.264	0.402	
		Boy	37	0.251	0.123	
		Total	80	0.258	0.304	
	Total	Girl	155	0.303	0.267	
		Boy	168	0.308	0.189	
		Total	323	0.306	0.229	
	Verb generation	Younger	Girl	52	0.604	0.926
			Boy	52	0.413	0.202
			Total	104	0.509	0.673
Middle		Girl	60	0.310	0.147	
		Boy	79	0.451	0.550	
		Total	139	0.390	0.430	
Older		Girl	43	0.226	0.155	
		Boy	37	0.365	0.358	
		Total	80	0.290	0.276	
Total		Girl	155	0.385	0.569	
		Boy	168	0.420	0.427	
		Total	323	0.403	0.500	

TABLE II. Cross correlation between head motion based on calculation and expert rating

	Expert rating				Head motion calculation
	E1	E2	E3	E4	
E1	1	—	—	—	—
E2	0.91	1.00	—	—	—
E3	0.95	0.94	1.00	—	—
E4	0.95	0.95	0.96	1.00	—
Head motion calculation	0.77	0.79	0.77	0.80	1.00

E1~4: Expert 1~4.

in the statistical analysis of motion. The descriptive statistics of the head motion data for these subjects is listed in Table I.

Concordance of Expert Ratings and the Motion Parameter

From the entire 1,292 (= 323 subjects × 4 tasks) fMRI motion data samples, 20 cases were randomly selected without regard to the age and sex of the subjects or the task during fMRI scanning. Correlations were calculated between the ratings by the four experts and also between the ratings and the motion parameter calculated from the coregistration algorithms. As shown in Table II, the ratings of head motion by the four experts are highly correlated with each other at the level of $R > 0.91$ and they are also consistent with the quantitative motion parameter at the level of $R > 0.77$.

Influence of Task Order on Head Motion

The four language tasks were presented to each subject in a counterbalanced order (Table III). An ANOVA was performed for the head motion in each task using task order as the factor that categorized the samples into eight sub-groups for each task. Table IV shows that, for any task, the task order as used in our experiment did not have a significant impact (at a P -level of 0.05) on head motion. In other words, when a subject performed a task, Syntactic Prosody for example, the motion was not significantly associated with what other tasks had been performed before the current one. It should be noted that the impact of task order approaches significance for verb generation task ($P = 0.06$). We conducted multiple comparisons with Post Hoc tests to examine motion in the verb generation task performed in different task orders. The motion for the verb generation task performed in task order “E” (as in Table III) was found significantly different when compared with other task orders ($P < 0.05$, with the order “F” being the only exception). Further examination showed that two subjects (out of 47 subjects scanned in task order “E”) had very large head motion (motion = 5.89 and 3.22). Without these two subjects, the motion for

TABLE III. Eight different combinations of task orders used in the experiment

	Task order			
	1	2	3	4
A	Story processing	Prosody processing	Word-picture	Verb generation
B	Story processing	Prosody processing	Verb generation	Word-picture
C	Verb generation	Word-picture	Prosody processing	Story processing
D	Verb generation	Word-picture	Story processing	Prosody processing
E	Prosody processing	Story processing	Verb generation	Word-picture
F	Prosody processing	Story processing	Word-picture	Verb generation
G	Word-picture	Verb generation	Story processing	Prosody processing
H	Word-picture	Verb generation	Prosody processing	Story processing

this task order (order “E” in Table III) is not statistically different from motion in any other order and the impact of task order on motion in verb generation (ANOVA) no longer approaches significance ($P = 0.28$). This suggests that the overall order effect at $P < 0.06$ for verb generation is the results of a very small number of subjects who happened to have large motion during the experiment.

Task-Correlated Motion

The correlation between head motion and the corresponding time course of task/control cycles was tested for each subject during each individual task to see if performing a certain language task induced a different magnitude of motion compared to control phases. Table V and Figure 1 show that the correlation analysis failed to establish a strong association between head motion and the task/control cycle. Using the same methodology we also calculated the head motion during the target behavior phase only, for all of the four tasks, to test whether performing target behavior alone would result in greater head motion. The observed mean head motion calculated this way was within the range of 0.01 to 0.02 (unit in pixel) from the value calculated for the entire time series for each task. None of these differences was significant. In summary, our analysis indicates that motion during each fMRI task did not depend on whether the subject was performing the

target behavior or the control condition for the target language behavior.

Analysis of Task, Age, and Sex Effects

A repeated measures three-way ANOVA using the general linear model (GLM) in SPSS was applied to test the impact of task (within-subject factor), age and sex (between-subject factors), and the interaction of these factors on motion observed in fMRI scanning. Based on the multivariate test (Wilks’ test), the subjects demonstrated significantly different head movement during different language tasks ($F(3,315) = 14.18, P < 0.0005$). The pattern of differences observed among tasks is significantly different among different age groups ($F(6,630) = 3.94, P = 0.001$). The sex of the subject also exerts significant impact on the motion in different tasks ($F(3,315) = 4.41, P = 0.005$). Finally, the combination of these two between-subject factors (age and sex) interacts with the task factor and significantly affects the motion outcome ($F(6,630) = 3.56, P = 0.002$). It should be noted that the result reported for age is based on a categorical scale. During the initial analysis, age was used as a continuous variable and similar results were observed. We retrospectively group the data into three age levels: Younger (5 to 9 years), Middle (10 to 14 years), and Older (15 to 18 years). The comparison of head motion using the two different age categorizations is

TABLE IV. ANOVA results for task order (N = 323)

		Sum of squares	df	Mean square	F	P
Prosody processing	Between groups	.824	7	0.118	0.888	0.52
	Within groups	41.749	315	0.133		
	Total	42.573	322			
Story processing	Between groups	3.145	7	0.449	1.683	0.11
	Within groups	84.078	315	0.267		
	Total	87.223	322			
Word-picture matching	Between groups	0.323	7	0.046	0.876	0.53
	Within groups	16.598	315	0.053		
	Total	16.921	322			
Verb generation	Between groups	3.410	7	0.487	1.991	0.06
	Within groups	77.066	315	0.245		
	Total	80.476	322			

TABLE V. Descriptive statistics of the cross-correlation between task/control cycle and head motion for different tasks

Task	N	Mean	Std dev	Max	Min	Median	25% percentile	75% percentile
Prosody processing	323	0.039	0.086	0.349	-0.306	0.037	-0.008	0.092
Story processing	323	-0.127	0.125	0.296	-0.482	-0.145	-0.210	-0.068
Word-picture matching	323	-0.152	0.117	0.387	-0.590	-0.159	-0.224	-0.091
Verb generation	323	-0.0993	0.145	0.403	-0.522	-0.111	-0.188	-0.026

demonstrated in Figure 2. It shows that head motion follows the same trend, i.e., decrease with the increase of age, using either categorization schemes for age factor. Since the main effects and interactions with other factors were similar, the age factor will only be reported in this study as categorical factor for the sake of convenience and clarity in the graphical presentation of age interaction with sex and task.

Figure 3 shows the profile plot of estimated marginal means of head motion as described in the Methods section. This figure illustrates a pattern in which head motion is affected by the main effect of task, as well as by the factor of age group. Conforming to the statistical results reported in the previous paragraph, the motion amplitude appears to vary for each of the different tasks. Post-hoc paired comparisons reveal that the Word-Picture Matching task induces the least motion and differs significantly from the three other tasks (Mean Difference, MD = -0.11, -0.09, -0.12, $P < 0.0001, 0.0001, 0.0001$ when compared with Syntactic Prosody, Verb Generation, and Story Processing task, respectively). No significant difference was observed between any two of the latter three tasks. For all four

tasks, there is a trend for motion to decrease with increasing age, with Word-Picture Matching task being a possible exception. The estimated marginal mean of motion was different for all three age-groups during different tasks, with the younger age group showing the greatest motion (MD = 0.14, $P < 0.0005$ when compared with middle age group; MD = 0.20, $P < 0.0005$ when compared with the older age group). No significant difference exists between the two older age groups averaged across tasks.

Figure 4 demonstrates the influence of task on motion from different sex. Pair-wise comparisons show that girls move significantly less than boys performing the same task (MD = -0.08, $P < 0.02$). The impact of sex on the task factor is also evident in the non-parallel lines for boys and girls in Figure 4.

Figure 5A and B demonstrate the impact of task, age, and sex on motion. The most important observation is that the difference of motion among the four language tasks

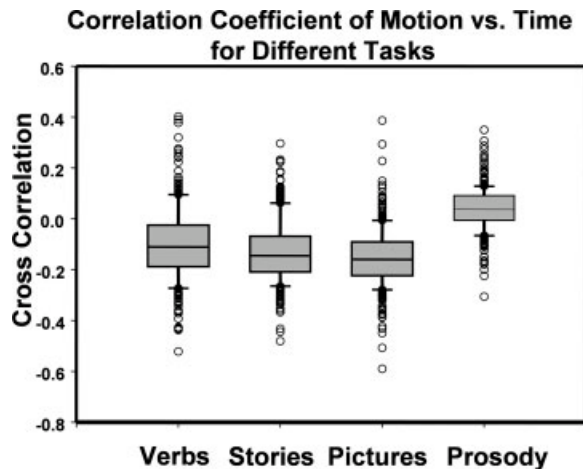


Figure 1.

Box plot of the correlation coefficient between head motion and task/control cycle in different fMRI tasks. The correlation coefficient was calculated between head motion versus the time course (the interleaved 30-sec on-period and 30-sec off-period). For all four tasks, this correlation is found to be very small for the majority of the subjects.

Motion vs. Age with Different Age Categorization

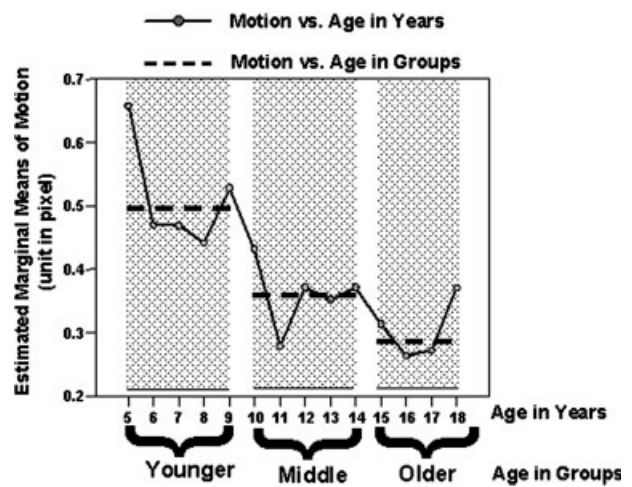


Figure 2.

Head motion data with different age categorization methods. After calculating head motion at each year of age, the inverse correlation between motion and age still follows the same trend as it is when age is categorized into three groups (Younger, Middle, and Older). In general, both categorization methods show that younger children have more motion than older children. The data displayed here is based on the head motion data collapsed across sex for ease of presentation.

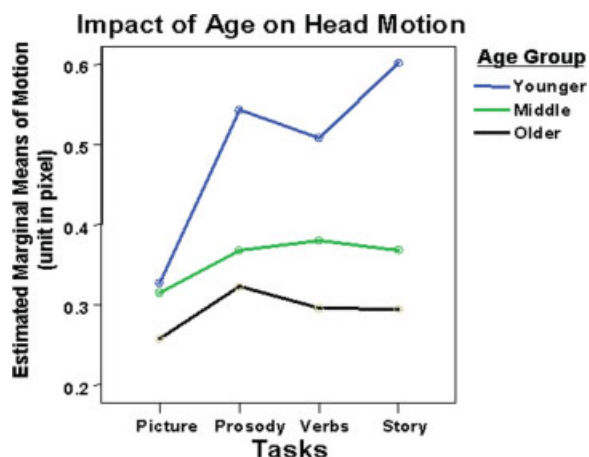


Figure 3.

Head motion in children in three age groups for different language tasks. Different tasks demonstrate different patterns of age difference in motion.

depends on age in conjunction with sex of the subject. For example, as mentioned previously, the Word-Picture Matching task induced the smallest motion. From the observation in Figure 5, it appears that the Word-Picture Matching task is the only task in which subject motion is minimal in all age groups and for both boys and girls. In the Syntactic Prosody task, the motion for girls decreased with increasing age, whereas boys had approximately the same degree of motion for the two older age groups. Girls in the Younger age group, performing the Verb Generation task, provide the only case for which the girls have significantly greater motion than boys. For this task, boys in all three age groups appear to have similar motion amplitude. For the Story Processing task, girls have less motion than the boys in all three age groups. The difference of motion between boys and girls is the largest in the Story Processing task: 0.322 ± 0.040 for girls vs. 0.521 ± 0.040 for boys (Mean \pm SE), with the most extreme difference occurring between the youngest boys and girls as shown in Figure 5. Averaged across tasks, the boys in the Younger age group have the greatest degree of head motion among all categories of children.

DISCUSSION

In this report we attempt to quantify the characteristics of head motion in various fMRI language tasks in a pediatric population. We employed the Thevenaz and Unser coregistration algorithm [Thevenaz et al., 1998] to quantify the degree of motion in boys and girls ranging in age from 5 to 18 years. This algorithm has been tested using datasets with simulated known motion data and it was found to be a simple yet accurate and robust routine method for evaluating head motion during fMRI scanning [Ardekani et al., 2001]. In this study, a high correlation existed

between the algorithm and expert rating as demonstrated in Table II.

An overview of the results of our motion analysis of fMRI data from 323 children performing four language paradigms leads to the following general observations: First, as one might expect, younger children tended to move more than older children. Second, boys in general moved more than girls for tasks that do not involve visual system. Third, and perhaps most important, tasks that involve active responses and multi-sensory stimulation (i.e., auditory and visual) appear less susceptible to head motion than tasks that do not include a visual component or active responses. The Word-Picture Matching task used in this study suffered the least amount of motion overall (0.306 pixels), which was significantly lower than all other tasks across both sex and age groups. Surprisingly, for this task motion did not differ significantly between boys and girls or between younger and middle age groups. It appears that a visual component may equalize motion in fMRI paradigms across groups. This observation clearly has important implications for designing appropriate fMRI tasks for use in children and perhaps in adults as well.

Looking quantitatively at the motion data in Table I and Figures 3, 4, and 5, some important features can be discerned. The use of different language tasks clearly resulted in different magnitudes of head motion. For example, the Word-Picture Matching task had the lowest movement among four tasks (mean displacement = 0.306 pixel) for both boys and girls, which can be attributed to the differential workload and attention required in this task compared to the others. It would seem that a purely passive auditory task (story processing) results in the greatest

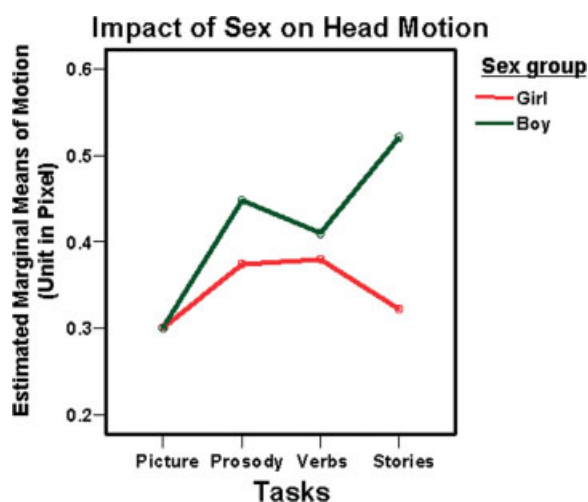


Figure 4.

Impact of sex factor on head motion during different language tasks. The difference of head motion in different tasks is also affected by sex factor. Boys are found to have more motion than girls in three out of four tasks tested.

Impact of Task, Age, & Sex on Head Motion

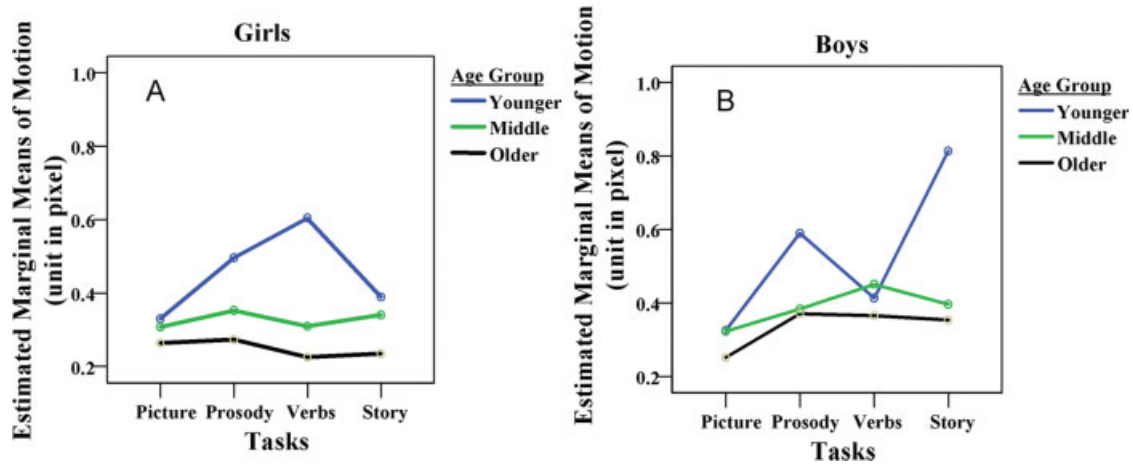


Figure 5.

The influence of task, age, sex, and the interaction of these factors on head motion. The three factors all affect the characteristics of head motion in children. For Syntactic Prosody, Verb generation, and also for Story listening task, the most significant difference is between boys and girls in the younger group.

motion during fMRI in children whereas the active response, multi-sensory task (Word-picturing matching) results in the least motion. The tasks with no visual stimulation, but which require a covert or explicit response (Verb Generation and Syntactic Prosody, respectively), resulted in motion that falls between that measured for the other two tasks. Our interpretation of this pattern is that inclusion of audio and visual stimulation together with button pushes or other explicit responses to the stimuli may reduce motion in fMRI data in children. This point is evidenced by the bottom line in Figure 3, which demonstrates that motion is lowest in the Word-Picture Matching task even for the youngest group of children.

Our results also show that the main effect of task on head motion can be affected by differences in subject age ranges and sex evidenced by the fact that the difference of head motion in the four tasks has a different contrast pattern at different age and sex (Figures 3, 4, and 5). The significant influence of age is demonstrated in Figures 2 and 3. While the general trend is that head motion is greater in the younger children, this difference is less dramatic in the Word-Picture Matching task than in the other three tasks.

The influence of sex is also interesting. For the Syntactic Prosody, Verb Generation, and Story Listening tasks, head motion is more pronounced in boys than in girls. This finding is consistent with the idea that girls usually mature at an earlier age than boys and therefore should be more compliant at earlier ages than the boys. However, it is clear that the nature of the task (i.e., Word-Picture Matching) may mitigate this difference, which may be a reflection of the fact that boys are more sensitive to visually-

engaging stimuli than girls. It should be noted that the Word-Picture Matching task includes visual as well as other elements, and it could be the combined visual, auditory, and manual component that leads to the less motion in children.

The effect of task on motion was not affected by the order in which the task was administered. Although one might suppose that children would be generally less cooperative as the scan session progressed, this was not the case. It may be argued that the relatively short duration of each task, and the use of multiple tasks actually facilitated cooperation by keeping the child engaged in different ways across the entire session. As analyzed in the Results section, the significance level at $P = 0.06$ for the order effect of motion in verb generation task came mainly from the motion of two subjects performing tasks in order “E,” suggesting that task order is not an important effect in a large number of subjects. However, given the popularity of this type of task in fMRI studies of children, close attention should be paid to this factor in future research that combines this particular task with other tasks in the same session.

The large sample size in this study ($N = 323$) provides sufficient power to permit thorough examination of multiple factors influencing head motion in pediatric fMRI data. Studies with less power are unlikely to allow for adequate degrees of freedom to explore three-way interactions of multiple tasks, age groups, sexes, and other variables. The results of this large scale study can permit investigators to understand the trade-offs between the nature of the task they intend to employ, the age group and sex of their

subjects to anticipate the degree to which motion may be problematic in their study. For example, if a study targeting passive processing by young children is planned, the investigators may wish to oversample relative to the cohort dimension estimated by a power analysis to tolerate the greater loss of data due to motion compared with a study that employs a more active paradigm or older children. In this context, this dataset can serve as a guide concerning which factors must be controlled during study design to ensure that head motion does not confound other effects of primary interest. Investigators are advised to carefully consider age and sex contributions in study designs involving children.

CONCLUSIONS

In summary, our study quantified for the first time head motion for normal children during various fMRI language tasks. The language task, age, and sex, as well as their interactions all exerted significant influences on head motion during fMRI scanning. Our results and conclusions should serve as a frame of reference regarding the influence of these factors for future pediatric patient studies in which motion artifact is expected to be more severe than in normal, healthy subjects [Yuan et al. 2006]. More rigorous means of minimizing motion might be necessary in the design of these experiments.

REFERENCES

- Ardekani BA, Bachman AH, Helpert JA (2001): A quantitative comparison of motion detection algorithms in fMRI. *Magn Reson Imaging* 19:959–963.
- Biswal BB, Hyde JS (1997): Contour-based registration technique to differentiate between task-activated and head motion-induced signal variations in fMRI. *Magn Reson Med* 38:470–476.
- Byars AW, Holland SK, Strawsburg RH, Bommer W, Dunn RS, Schmithorst VJ, Plante E (2002): Practical aspects of conducting large-scale functional magnetic resonance imaging studies in children. *J Child Neurol* 17:885–890.
- Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ (1995): Spatial registration and normalization of images. *Hum Brain Mapp* 3:165–189.
- Grootenck S, Hutton C, Ashburner J, Howseman AM, Josephs O, Rees G, Friston KJ, Turner R (2000): Characterization and correction of interpolation effects in the realignment of fMRI time series. *Neuroimage* 11:49–57.
- Holland SK, Vannest J, Mecoli M, Jacola LM, Tillema JM, Karunayaka PR, VJS, Yuan WH, Plante E, Byars AW (2007): Functional MRI of language lateralization during development in children. *Int J Audiol* 46:533–551.
- Kotsoni E, Byrd D, Casey BJ (2006): Special considerations for functional magnetic resonance imaging of pediatric populations. *J Magn Reson Imaging* 23:877–886.
- Liao R, McKeown MJ, Krolak JL (2006): Isolation and minimization of head motion-induced signal variations in fMRI data using

- independent component analysis. *Magn Reson Med* 55:1396–1413.
- Schmithorst VJ, Dardzinski BJ (2000). CCHIPS: Cincinnati Children's Hospital imaging processing software. Available at <http://www.irc.chmcc.org/chips.htm>
- Schmithorst VJ, Holland SK (2007): Sex differences in the development of neuroanatomical functional connectivity underlying intelligence found using Bayesian connectivity analysis. *Neuroimage*. 35:406–419.
- Seto E, Sela G, McIlroy WE, Black SE, Staines WR, Bronskill MJ, McIntosh AR, Graham SJ (2001): Quantifying head motion associated with motor tasks used in fMRI. *Neuroimage* 14:284–297.
- Szaflarski JP, Schmithorst VJ, Altaye M, Byars AW, Ret J, Plante E, Holland SK (2006): A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. *Ann Neurol* 59:796–807.
- Thevenaz P, Unser M (1998): A pyramid approach to subpixel Registration based on intensity. *IEEE Trans Image Process* 7:27–41.
- Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC (1998): Automated image registration: I. General methods and intrasubject, intramodality validation. *J Comput Assist Tomogr* 22:139–152.
- Yuan W, Szaflarski JP, Schmithorst VJ, Schapiro M, Byars AW, Strawsburg RH, Holland SK (2006): fMRI shows atypical language lateralization in pediatric epilepsy patients. *Epilepsia* 47:593–600.

APPENDIX: MOTION RATING SCALE

Instructions

Each dataset should be rated using the 1–4 scale below. If it contains one or more jump(s) as defined below, add 0.5 to the rating.

1. Perfect data, no discernable motion
2. Some minimal motion where all slices contain a change in intensity and/or position. Degree of motion does not pose a threat to the integrity of the data.
3. Continuous, moderate motion where all slices contain a change in intensity and position. Degree of motion poses a threat to the integrity of the study's data
4. Continuous, excessive motion where all slices contain a change in intensity and position; data should clearly be discarded.

Jump

Discrete motion that does not affect all of the slices in a frame. Because slices are acquired sequentially at 125 msec intervals, transient motion of less than 3 sec in duration will not affect all slices. A jump is therefore defined as a change in position or intensity that does not affect ALL slices.