Prosodic Processing by Children: an fMRI Study

Elena Plante[1], **Scott K. Holland**[2], and **Vince J. Schmithorst**[2]

1 *The University of Arizona, Tucson, AZ.*

2 *Children's Hospital Research Foundation, The University of Cincinnati, Cincinnati OH.*

Abstract

Prosodic information in the speech signal carries information about linguistic structure as well as emotional content. Although children are known to use prosodic information from infancy onward to assist linguistic decoding, the brain correlates of this skill in childhood have not yet been the subject of study. Brain activation associated with processing of linguistic prosody was examined in a study of 284 normally-developing children between the ages of five and eighteen years. Children listened to low-pass filtered sentences and were asked to detect those that matched a target sentence. fMRI scanning revealed multiple regions of activation that predicted behavioral performance, independent of age-related changes in activation. Likewise, age-related changes in task activation were found that were independent of differences in task accuracy. The overall pattern of activation is interpreted in light of task demands and factors that may underlie age-related changes in task performance.

Keywords

fMRI; language; prosody; children; brain

Speech processing in natural contexts involves the perception and integration of both segmental (i.e., phonological) and suprasegmental information. For adults, the suprasegmenatal information, or prosody, serves to cue emotional states (emotional prosody) and language structure (linguistic prosody). The latter includes information on sentence type (questions vs. statements), the occurrence of phrasal units within sentences and the likely boundaries of words within phrases. As such, prosodic information shapes how the linguistic information carried by sentences is parsed and interpreted. Furthermore, it can serve as the basis for processing unfamiliar units. For example, the stress patterns of the words *habeus corpus*, when spoken, suggest two words are present rather than some other number. This provides a rough idea of the units of meaning that are likely to be attached to these elements.

For adult listeners, prosody is a robust linguistic phenomenon. Prosodic information is redundantly distributed over a broad range of the acoustic speech spectrum (Grant & Walden, 1996) and adults are able to identify information carried through virtually any segment of the speech spectrum (Grant & Walden, 1996, Lakshminarayanan, Ben Shalom, van Wassenhove, Orbelo, Houde, & Poeppel, 2003). Furthermore, physiologic data suggests that speech prosody is used to disambiguate meaning when the words of a sentence alone do not signal how the phrases should be parsed (Steinhauer, Alter, & Friederici, 1999).

When learners are in the process of acquiring language, prosodic cues are even more important. Infants show evidence of speech preferences that appear to reflect prosodic information shortly after birth (DeCasper & Spence, 1986; Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, &

Contact author: E. Plante, Ph.D., Department of Speech and Hearing Sciences, PO Box 210071, The University of Arizona, Tucson, AZ 85721-0071 USA. Eplante@u.arizona.edu Phone: 520/621-5080. FAX: 520/621-9901..

Amiel-Tison, 1988). Infants are capable of using this information to segment individual words from running speech at seven and a half months of age (Jusczyk, Huston, & Newsome, 1999) although they do not appear to use prosodic cues preferentially until somewhat later (Theissen & Saffran, 2003). By nine months, American infants show a clear bias towards the trochaic stress pattern that characterizes spoken English (Gerken, 2004; Jusczyk, Cutler, & Redanz 1993) and this information takes precedence over segmental information as a cue to word boundaries (Johnson & Jusczyk, 2001; Mattys, Jusczyk, Luce, & Morgan,1999; Theissen & Saffran, 2003). Although prosody alone is an imperfect cue to syntactic structure, there is evidence that infants at nine months are sensitive to the prosodic structure of sentences that they have heard (Gerken, Jusczyk, & Mandel, 1994). Moreover, infants at this age can are sensitive to the specific prosodic structure of phrases they hear and show evidence of abstracting the underlying structure, allowing for generalization (Gerken, 2004). Thus there is evidence that very young children are not only sensitive to prosodic information, but are capable of using this information to evaluate new input.

Neuroimaging studies identify broadly distributed networks associated with prosodic processing tasks. Activation in the superior temporal region appears to be critical for the analysis of the acoustic signal that carries prosodic information (Plante, Creusere, & Sabin, 2002). This area activates regardless of whether the prosodic cues signal emotional or linguistic information and tends towards a right hemisphere lateralization (Dongil, Ackerman, Grodd, Haider, Kamp, Mayer, Riecker, & Wildergruber, 2002; Meyer, Alter, Friederici, Lohmann, & von Cramon 2002; Meyer, Steinhauer, Alter, Friederici, & von Cramon, 2004; Mitchell, Elliott, Barry, Cruttenden, & Wordruff, 2003; Plante et al., 2002). Some researchers have made the distinction that the contributions of the superior temporal region reflects the acoustic properties of the signal, rather than a verbal-nonverbal dichotomy, with rapidly changing acoustic information likely to show left lateralization and slow changing information showing right lateralization (Dongil et al., 2002; Pöepell, 2003). This type of material-specific lateralization is measurable in infancy and is not restricted to cortical auditory areas, but is also reflected early in life at lower levels in the auditory system (Sininger & Cone-Wesson, 2004). However, it is apparent that how adults treat the acoustic signals of prosodic information is shaped by their linguistic experience (Gandour, Dzemidzic, Wong, et al., 2003; Gandour, Wong, Dzemidzic, Lowe, Tong, & Li, 2003). This suggests that there may be some possibility for change in the neural support for prosodic processing with experience.

Additional regions of activation appear to reflect task-specific demands. Typically, subjects are asked to hold information in memory (Meyer, Steinhauer, Alter, Friederici, & von Cramon, 2004; Plante et al., 2002), explicitly rehearse (Meyer et al., 2004) or make judgements about the stimuli (Meyer et al., 2002; Meyer et al., 2004; Plante et al., 2002). These activities appear to recruit frontal and parietal lobe regions. Plante et al. (2002) reported that passive listening to prosodic stimuli resulted in activation in the temporal lobe whereas frontal and parietal activation was seen when subjects were required to compare low-pass filtered sentences to an unfiltered target sentence held in memory. Likewise, Meyer et al (2002) reported left inferior frontal, anterior insula and motor-related cortex activation in association subvocal rehearsal of the prosodic signal. Parietal activations appear lateralized to the left hemisphere (Meyer et al., 2004; Plante et al., 2002) whereas right lateralization in dorsolateral prefrontal regions are seen for prosodic processing tasks. Plante et al reported that a prosodic processing task involving memory and judgment components was right lateralized in this region, whereas sentence processing involved peak activation in left dorsolateral prefrontal areas, suggesting an interaction between the stimulus and memory components of these tasks (2002).

The imaging studies of speech prosody to date have involved adult participants exclusively. Whether or not these same patterns would be seen in the developing brain remains an open question. The study of children presents an opportunity to examine two factors that co-occur

during childhood and may each affect how the brain supports behavioral skills. The first is that of brain maturation. We know, for example, that frontal lobe regions, important when we ask adults to remember or use prosodic information in some way, have a long developmental trajectory relative to other brain regions involved in prosodic processing tasks. Therefore, the extent to which brain regions contribute to a prosody processing task may vary with maturation. It is also possible to have age-related changes in activation because children adopt different strategies (e.g., subvocal rehearsal, sustained attention) that also result in better performance with age. This too would result in some areas becoming increasingly active with age.

The second factor involves level of expertise. It is clear that infants are able to process and use prosodic information. However, the propensity to rely on prosodic cues vs. other types of segmental cues can shift with age and, correspondingly, with the stage of language acquisition. It is also probable that different individuals recognize these cues with greater or lesser success. It may be the case that activation in certain regions of the brain are more predictive of successful prosodic processing than others and that sensitivity to these cues may shift over the course of childhood.

Here we examine the relative contribution of maturational and behavioral differences in a large sample of children between the ages of five and 18 years of age. These participants completed a prosodic processing task similar to one used by Plante et al in adults (2002). This task requires a subject to hold a target sentence in memory and to evaluate a series of low-pass filtered sentences to determine when the latter match the target sentence. As such, the task requires both processing of and memory for prosodic information. Furthermore, it allows for the collection of two types of behavioral information: correct identification of target sentences (true positive responses) and mistaken identification or false positive responses. We hypothesize that because processing of prosodic information appears to require activation of the superior temporal cortex, higher activation in this area will be associated with rates of correct accept responses and lower activation with false positives. Frontal regions that appear related to the interaction of memory and prosodic processing (cf. Plante et al., 2002) should also show a similar pattern in association with rates of correct accepts and false positive responses. Finally, we are prepared that children of different ages may require greater attentional resources to complete the task. This may result in activation in regions associated with attentional networks (e.g., parietal, cingulate), that show age-related changes.

Method

Participants

The participants of this study are a subset of participants of a larger neuroimaging program that focuses on language development in children. As part of this program, 284 children completed a functional neuroimaging study of prosody. and these children serve as the subjects of this report. All children spoke American English as their native and primary language. Their demographic information is provided in Table 1. All children were normally developing by parent report. In addition, IQ using the *Wechsler* series (Wechsler, 1997; Wechsler, 1987). status. The *Oral and Written Language Scales* (Carrow-Woolfolk, 1996) was used to assess broad language skills and confirmed that receptive and expressive abilities were within normal limits. A neurological exam established normal neurological status. Informed consent was obtained for all participants of this study prior to enrollment in the study.

Behavioral Materials and Procedures

The prosody task used a block periodic design consisting of a cue period (not analysed), a 30 second "on" block, during which the syntactic prosody task was performed, and a 30 second "off" block during which a control task was performed. This cycle repeated five times. Children

heard a target sentence during the cue period. Target sentences were constructed using a core content words (i.e., nouns and verbs) that should be readily understood by age 5 years. These were taken from standardized measures of vocabulary designed for use with children ages 2 and above. Following the cue sentence, children were asked to indicate whether low-pass filtered sentences they subsequently heard corresponded to the target sentence. A different target sentence was presented for each cue period so that the children held a new sentence in memory for comparison during each "on" block. The correct target sentences occurred fifteen times in filtered form over the course of the experiment.

Sentences heard during the "on" block were low-pass filtered at 450 Hz so that the prosodic information was preserved but the lexical information was indistinct. Sentences consisted of a variety of syntactic types by varying the sentence type (e.g., statements, questions) and the presence, absence, and position of embedded clauses and prepositional phrases. This resulted in a set of foil, filtered sentences that differed in terms of the syntactic form, the number of syllables (and syllable stress), or both from the target sentences. Therefore, the foil sentences varied in terms of degree of similarity to the target sentence. Children indicated when they thought the target sentence had occurred by button press, which transmitted responses via an infra-red system. Both correct accepts of target sentences and false positive responses were recorded by the MacStim program (White Ant Software, Melbourne, AU) running the paradigm on the Macintosh computer

The "off" task was designed to preserve the elements of working memory (holding a target in memory), sustained attention (listening to sequentially-presented information), and decision-making (target vs. nontarget judgements) while eliminating the element of prosodic processing. Children were given a target tone (a warble tone) and asked to identify when that tone occurred during the "off" block via button press. The same warble tone was used for all off task blocks to reduce task confusion. This tone was embedded among a series of pure tones in each off block to compose a control interval task that would be easy for a 5 year old to perform accurately.

All children were trained to perform the on and off tasks prior to the scan. Training continued until the child's performance indicated that they understood and could complete the task.

Imaging Procedures

All children were scanned using a 3T Bruker Biospec 30/60 MRI scanner (Bruker Biospin, Karlsruhe, Germany). The functional scan consisted of T2*-weighted, gradient-echo EPI images with 145 time points. The first ten of these time points were discarded so that the signal reached T1 relaxation equalibrium prior to when stimulus presentation began. Five millimeter contiguous slices were obtained in the axial plane (TR=3000, TE=38ms, FOV=256x256mm, 64x64 matrix, 24 slices). In order to correct geometric distortion and Nyquist artifacts which can result in poor registration of the functional images to the anatomical images, a 3D phase reference image was also acquired. Functional image correction of the EPI images followed the procedures described by Schmithorst, Dardzinski, and Holland, (2001). Finally, a whole brain, high resolution structural image was acquired using a 3D MDEFT (Modified Driven Equalibrium Fourier Transform) (TR/TE/τ=15.7/4.3/550msec, FOV=194x256mm, 256x122x128 matrix).

Image processing was completed using Cincinnati Children's Hospital Image Processing Software (CCHIPPS, Schmithorst, 2000). The raw EPI data was initially processed with a hamming filter to reduce truncation artifacts and high frequency noise (Lowe & Sorenson, 1997). We evaluated the level of motion in the data using quantitative procedures based on tracking the root means square (rms) displacement of each image volume during realignment. Briefly, this consisted of implementing a pyramid co-registration and normalization approach

using the first image volume as the co-registration template (Thévenaz & Unser, 1998). A measure of motion relative to the first image volume was generated by obtaining the linear transformation parameters produced by this co-registration process. The criterion for unacceptable levels of motion was based on rejecting data that exhibited rms displacement greater than 1.0 mm, over the data run. Only children whose functional images produced motion parameters that met this criterion were included in this report.

Regions that were active during the time children were engaged in the prosody task were identified through an initial regression procedure followed by eliminating active voxels that were not spatially clustered. The relation between the signal change associated with task performance was computed using a General Linear Model with a set of covariates (cosine functions) used to account for respiratory and cardiac effects. Regression coefficients were transformed to t-scores so that values were normalized across subjects and ages. Images from individual subjects were converted to Talairach space. We have previously demonstrated that the Talairach reference frame is adequate for use in children as young as 5 years and have validated its use in the subject cohort.

Regions of activation were determined by combining data for all children and then applying a threshold to establish the extent of group activation. Regions of active contiguous voxels for the full group and regions that showed effects that appeared to be age-related were identified as regions of interest (ROIs) for statistical analysis. The large number of subjects permitted a low uncorrected threshold ($p < 1e-8$ for group activation; $p < 1e-4$ for age effects) in a random-effects analysis. These can be seen in Figures 1 and 2. Masks of these regions were applied to the functional images of the individual subjects in order to extract the mean t values for each ROI in subsequent hypothesis-driven analyses.

Results

Behavioral Results

Table 2 displays the result of the prosody task across subject ages. Across the age ranges subjects had more correct accept responses than false positive responses. The d' statistic was used to describe the differential responses to target and foil sentences. This statistic reflects the difference between the means of the correct accept and false positive response distributions in units of standard deviation. The d' showed an increase with age indicating better differentiation between filtered target and foil sentences. The maximum d' obtained by the oldest subjects indicates the combination of the sentence types and presentation in the scanner environment presented a discrimination challenge even to the most proficient listener group. However, there was also considerable range of ability within each age group. By age 7, all age groups included participants whose correct accept and/or false positive rates were near perfect. The ranges of responses across ages suggest considerable overlap in behavioral performance.

To characterize the nature of performance on the prosody task, we correlated individual subject performance with test scores on both IQ and language measures. There was no significant correlation for either correct accept or false positive responses and IQ scores. This is consistent with a minimal role for general intelligence for this task. In contrast, there is significant correlation between general language test scores and correct accept (r=.22, p=.0002) and false positive responses (r=−.12, p=.0355). Although the magnitude of the relation was small, these results are consistent with the idea that language skills contribute to the performance on the prosody task.

Imaging Results

Figure 1 represent the activation associated with task performance in the full group collapsed across age. These are described in terms of their center and extent in Table 3. In addition to the composite maps that elucidate group activation patterns, we also did a second voxel-wise analysis. The result of this analysis is displayed in Figure 2 and reflects the voxel-wise correlation of GLM z-score with age as a regressor. We refer to the representation of our data in Figure 2 as "age correlation maps." A comparison of this age correlation map with Figure 1 indicates considerable overlap with the active regions in the composite analysis, though with differing extents at the thresholds selected for display corresponding to p<0.01. In addition, three regions active in the full group analysis did not correlate with age. These occurred in the right middle frontal gyrus and the right and left anterior occipital lobes. In addition, the age correlation map indicated two distinct regions of interest with different centroids for the inferior frontal gyrus and anterior insula whereas these appeared as one cluster in Figure 1.

The regions present in either Figure 1 or 2 served as the ROIs for statistical analysis. When the same general region appeared in both figures, but differed in extent of activation, we entered both in the statistical analysis and permitted the one that accounted for the largest proportion of variance to load into the model. The shared variance among regions common to both figures prevented both from loading into any given statistical model.

Correct Accepts (True Positive Responses)—The first analysis addressed the issue of which brain regions contributed to correct identification of target sentences. However, to answer this question, we must first control for the effect of age. Therefore, this analysis statistically controlled for the variance in activation associated with age (in months) by entering this variable into the regression analysis prior to allowing ROIs to enter. We used the procedure for maximizing R^2, adjusted for the number of predictor variables entered, to identify the maximum number of regions that added valid variance to iterative regression solutions. We adopted the adjusted R^2 approach to allow for variables in the regression equation that are not necessarily individually statistically significant, but none-the-less add valid variance to the prediction. After forcing age in months as the initial predictor variable (to remove this source of variance before ROIs were entered), an additional four variables added valid variance to the prediction of correct accepts (F(5,282)=11.50, p<.0001; adjusted R^2 =.1597). Activation in the left precentral sulcus, the left superior temporal and right middle frontal gyri was positively related to correct accept performance levels. In contrast, activation in the cingulate gyrus was negatively related to performance. The results of the full regression model are provided in Table 4.

False Positives—False positives are of interest independently of correct accepts. After age was again statistically controlled, five ROIs predicted this performance metric (F(6,281) =15.43, p<.0001; adjusted R^2 =.2369). As with correct accepts, the left superior temporal region contributed to the regression equation, but was negatively correlated with the proportion of false positive responses. The right anterior insula was also negatively associated with false positives. Predictors that were positively associated with false positive performance included the right precentral sulcus, right middle frontal gyrus, and cingulate gyrus. The results of this regression are presented in Table 4.

Age-related change—To understand the relation between ROIs and age, independent of behavioral performance, we used a regression procedure that statistically controlled the variance associated with both correct accepts and false positives prior to calculating the variance associated with age (in months). After forcing the proportion of correct accepts and false positives into the regression equation, six additional regions predicted age in months (F (7,276)=27.59, p<.0001; adjusted R^2 =.4042). After the behavioral variables (each of which

were statistically significant predictors), activation in the left anterior occipital, right anterior occipital, right middle frontal, left and right precentral and left superior temporal gyrus region predicted age in months. The parameter estimates and probability levels associated with each of these regions are found in Table 4.

Discussion

This large-scale study largely replicated the results of earlier imaging studies of adult subjects who completed prosodic processing tasks. Our pediatric subjects showed a network of activation that included frontal, temporal, and parietal regions as previously described. The superior temporal gyrus showed a right lateralization, which has been seen in a previous study that used a very similar task (Plante et al., 2002). In addition, frontal areas including the inferior and middle frontal gyrus, and precentral sulcus activated to the task. Previous studies (Meyer et al., 2004; Plante et al., 2002) suggested that these areas tend to show a right lateralization for processing prosodic information vs. normal speech. However, the children showed less lateralization than previously described for most areas, with the exception of activation centered on the right middle frontal gyrus which was strongly lateralized to the right hemisphere. An area in the anterior occipital region, bordering on the posterior temporal lobe showed age-related, but not performance-related changes in activation. Activation in this area was reported for conditions that compared sentences with linguistic content to filtered sentences (Meyer et al, 2002; Meyer et al, 2004) and for the phonological content vs. the emotional valence of word pairs (Buchanan et al., 2000). This activation, which increased in the left hemisphere and decreased in the right with age when performance differences were controlled, may have reflected an increasing tendency with age to attempt to fit the linguistic content of the target sentence to the filtered sentences during the task.

This study was able to identify regions for which activation appeared to vary with behavioral success and error rates that were independent of age. These patterns suggest the roles that each of these regions may play in performing the prosody task. Although seven distinct regions were active during the task, only a subset of these regions actually predicted aspects of performance. This subset included areas that showed increases in activation with increasing rates of correct accept responses and decreasing activation with increasing rates of false positives. As expected, the left superior temporal gyrus showed this profile. This pattern suggests that greater engagement of this auditory processing region results in better processing of the filtered sentences, and therefore, better performance. What is interesting in this process is that, although there was a right-greater-than-left asymmetry in activation for this area, it was the left superior temporal gyrus that predicted correct accepts and false positives. This suggests that although the right superior temporal gyrus (STG) may have been registering the longer-duration tonal variations at the sentence or phrase level, activation in the left STG that reflected processing of the shorter, syllable-length prosodic cues that was key in predicting task performance (cf., Gandour et al., 2003; Dongil et al., 2002; or Pöppel, 2003). A syllable-level analysis would be most important when subjects were evaluating filtered sentences of a similar syntactic form as the target. These types of sentences are likely to be the most difficult to discriminate and are most likely to be correctly identified by individuals with the highest rates of correct accept responses.

Frontal activation included several distinct regions that showed independent contributions to both behavioral performance and age. Frontal regions do not generally activate when subjects are simply maintaining attention to prosodic information (Plante et al., 2002). In this study, frontal sites including an area centered around the dorsolateral precentral sulcus in both hemispheres and the right middle frontal gyrus made independent contributions to the prediction of age and behavioral performance. For the precentral sulcus region, increases in the left ROI were associated with increases in correct accepts while increases in the right ROI

were associated with increases in the false positive rate. This suggests that each hemisphere was contributing to processing of different aspects of the task and memory for these aspects might have promoted accurate recall (left) or a mistaken sense of familiarity (right). This would occur if this region in each hemisphere was preferentially contributing to memory for different aspects of the target sentence. Gandour et al. (2003), reported that native speakers of Chinese showed larger left hemisphere activation in this area when asked to make judgements on syllable level stress and right hemisphere activation when asked to judge phrase intonation. In our study, subjects who had a sense of familiarity between a filtered and target sentence may have derived that sense from similarities with either syllable stress patterns within the sentence frame or from the intonational pattern of the sentence as a whole. Reliance on syllable-level cues will differentiate sentences of both similar and distinct syntactic types, leading to high correct accept rates. In contrast, reliance on intonational patterns will differentiate sentences of different syntactic types but not similar ones. Therefore, reliance on these cues will increase false positive rates overall. In contrast to the precentral sulcus ROI, the right middle frontal gyrus showed increases in activation with either correct accept or false positive responses. This suggests that it may have been active when subjects sensed a familiarity to the target sentence for any reason. Alternatively, this region could have contributed more to elements of decision making, which are also common to both responses, rather than comparison of the target and filtered sentences.

Likewise, activity in the right anterior insula was negatively correlated with false positive responses after age-related variance was controlled. Greater activity in this area has been found for tasks requiring attention to tonal patterns (Zatorre, Evans, Meyer, 1994), melodic production tasks (Reiker, Wildgruber, Dogil, & Grodd, 2000; Reiker et al., 2002), and activity is right lateralized during judgement of sentence-level stress (Gandour et al., 2003). The negative relation between activity in the right anterior insula and false positive responses may reflect poor replication of the intonational contour of the target sentence during subvocal rehearsal. Again, this cue has only limited utility for correct accept responses but would be more important for avoiding misidentification of foils.

Both the left anterior insula and the bilateral inferior frontal gyrus were active during task performance, but neither was a significant predictor of either correct accepts or false positive responses. In addition, these areas did not show age-related change. Therefore, for this task, these areas may represent base resources that are necessary but not sufficient for the performance of this particular task. It has been previously suggested that the left anterior insula may reflect a strategy of subvocal rehearsal or the planning of speech (Rieker et al., Ackermann, Wildgruber, Dogil, & Grodd, 2000), which would occur if subjects repeated the target sentence to themselves to hold it in memory. It may be that this did not predict performance because the phonological information rehearsed was not relevant for identifying low-passed sentences. The role of the inferior frontal gyrus in this task is a matter for speculation. The role of this area is suggested by the results of several other studies. First, this region does not activate strongly when subjects are simply asked to attend to low-pass filtered speech (Plante et al., 2002). Therefore, activation is not likely linked to networks involved in either the acoustic processing of prosodic information or the effort in maintaining attention to the stimuli. This area is active in children performing generative tasks (Gaillard, Hertz-Pannier, Mott, Barnett, LeBihan, Theodore, 2000; Gaillard, Sachs, Whitnah, et al., 2003; Holland, Plante, Byars, Strawsburg, Schmithorst, & Ball, 2002) that presumably engage subvocal mechanisms (e.g., verbal fluency tasks) and this region shows age-related shifts during childhood for those tasks (Holland et al, 2001). However the lack of age-related changes for this task or any relation to performance success suggest that it does not represent a use of a subvocal strategy, which should both increase success behaviorally and be used more frequently with age. Therefore, activation in this region may represent a necessary contribution of a portion of the memory network (see

Chein, Ravizza, & Fiez, 2003 or Buckner & Koustaal, 1998 for reviews) that itself is not sufficient for task success.

The overall pattern of results in this study helps to differentiate the role of areas previously described in studies of adults. We found bilateral areas of activation for the prosody task, with both the superior temporal gyrus and middle frontal gyrus showing right lateralization. However, the right hemisphere contributions did not generally predict good prosodic processing. Instead, the differential contribution of temporal and frontal areas to correct accept and false positive responses is consistent with theoretical positions (Dongil et al., 2002; Pöppel, 2003) that the right and left hemispheres contribute to processing of prosodic cues based on the temporal window of the signal. The relative utility of these skills for completing the task appears to explain the to the correlations we observe between brain activation and task performance. The picture that emerges from this study qualifies the relative importance of the historical emphasis on right lateralization in prosodic processing studies. Instead, it appears that this lateralization may reflect processing of a subset of prosodic cues that may or may not be critical for a particular task. Focus on sentence-level intonation may be important for differentiating questions, statements, and exclamations. However, if the task at hand is to identify the stress cues that identify words (Johnson & Jusczyk, 2001; Jusczyk et al., 1999; Mattys et al., 1999; Theissen & Saffran, 2003), left hemisphere contributions may be more important.

References

Buchanan TW, Lutz K, Mirzazade S, Specht K, Shah NJ, Zilles K, Jäncke L. Recognition of emotional prosody and verbal components of spoken language: an fMRI study. Cognitive Brain Research 2000;9:227–238. [PubMed: 10808134]

Buckner, R.L. & Koustaal, W. Functional neuroimaging studies of encoding, priming, and explicit memory retreival. *Proceedings of the National Academy of Science, USA. 95,* 891–898.

Carrow-Woolfolk, E. (1996). *Oral and Written Language Scales* Circle Pines, MN: American Guidance Service.

Chein JM, Ravizza SM, Fiez JA. Using neuroimaging to evaluate models of working memory and their implications for language processing. Journal of Neurolinguistics 2003;16:315–339.

DeCasper AB, Spence MB. Prenatal maternal speech influences newborns' perception of speech sounds. Infant Behavior and Development 1986;9:133–150.

Dongil G, Ackermann H, Grodd W, Haider H, Kamp H, Mayer J, Riecker A, Wildgruber D. The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody, & syntax. The Journal of Neurolinguistics 2002;15:59–90.

Gaillard WD, Hertz-Pannier L, Mott SH, Barnett AS, LeBihan D, Theodore WH. Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. Neurology 2000;54:180. [PubMed: 10636145]

Gaillard WD, Sachs BC, Whitnah JR, Ahmad Z, Balsamo LM, Petrella JR, Braniecki SH, McKinney CM, Hunter K, Xu B, Grandin CB. Developmental Aspects of Language Processing: fMRI of Verbal Fluency in Children and Adults. Human Brain Mapping 2003;18:176 –185. [PubMed: 12599275]

Gandour J, Dzemidzic M, Wong D, Lowe M, Tong Y, Hsieh L, Satthamnuwong N, Lurito J. Temporal integration of speech prosody is shaped by langauge experience: An fMRI study. Brain and Language 2003;84:318–336. [PubMed: 12662974]

Gandour J, Wong D, Dzemidzic M, Lowe M, Tong Y, Li X. A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. Human Brain Mapping 2003;18:149–157. [PubMed: 12599272]

Gerken LA. Nine-month-olds extract structural principles required for natural language. Cognition 2004;93:B89–B96. [PubMed: 15178379]

Gerken LA, Jusczyk PW, Mandel DR. When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. Cognition 1994;51:237–265. [PubMed: 8194302]

Grant KW, Walden BE. Spectral distribution of prosodic *information*. Journal of Speech and Hearing Research 1996;39:228–238. [PubMed: 8729913]

Holland SK, Plante E, Byars AW, Strawsburg RH, Schmithorst VJ, Ball WS. Normal fMRI brain activation patterns in children performing a verb generation task. NeuroImage 2002;14:837–843. [PubMed: 11554802]

Johnson EK, Jusczyk PW. Word segmentation by 8-month-olds: When speech cues count more than statistics. Journal of Memory and Language 2001;44:548–567.

Jusczyk PW, Cutler A, Redanz NJ. Infants' preference for the predominant stress patterns of English words. Child Development 1993;64:675–687. [PubMed: 8339688]

Jusczyk PW, Houston DM, Newsome M. The beginnings of word segmentation in English-learning infants. Cognitive Psychology 1999;39:159–207. [PubMed: 10631011]

Lakshminarayanan K, Ben Shalom D, van Wassenhove V, Orbelo D, Houde J, Poeppel D. The effect of spectral manipulations on the identification of affective linguistic prosody. Brain & Language 2003;84:250–263. [PubMed: 12590914]

Mattys SL, Jusczyk PW, Luce PA, Morgan JL. Phonotactic and prosodic effects on word segmentation in infants. Cognitive Psychology 1999;38:465–494. [PubMed: 10334878]

Mehler J, Jusczyk PW, Lambertz G, Halsted N, Bertoncini J, Amiel-Tison C. A precursor of language acquisition in young infants. Cognition 1988;29:144–178.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY. FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Human Brain Mapping 2002;17:73–88. [PubMed: 12353242]

Meyer M, Steinhauer K, Alter K, Friederici AD, von Cramon DY. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain and Language 2004;89:277–289. [PubMed: 15068910]

Mitchell RLC, Elliott R, Barry M, Cruttenden A, Woodruff PWR. The neural response to emotional prosody as revealed by functional magnetic resonance imaging. Neuropyschologia 2003;41:1410–1421.

Plante E, Creusere M, Sabin C. Dissociating sentential prosody from sentence processing: activation interacts with task demands. NeuroImage 2002;17:401–410. [PubMed: 12482093]

Pöppel D. The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. Speech and Communication 2003;41:425–255.

Reiker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula, and cerebellum. NeuroReport 2000;26:1997–2000.

Reiker A, Wildgruber D, Dogiz G, Grodd W, Akerman H. Hemispheric lateralization effects of rhythm implementation during syllable repetitions: An fMRI study. NeuroImage 2002;16:169–176. [PubMed: 11969327]

Steinhauer K, Alter K, Friederici AD. Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nature Neuroscience 1999;2:191–196.

Steinhauer K, Friederici AD. Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. Journal of Psycholinguistic Research 2001;30:267–295. [PubMed: 11523275]

Sininger YS, Cone-Wesson B. Asymmetric Cochlear Processing Mimics Hemispheric Specialization. Science 2004;305:1581. [PubMed: 15361617]

Theissen ED, Saffran JR. When cues collide: When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. Developmental Psychology 2003;39:706–716. [PubMed: 12859124]

Thévenaz P, Unser M. A pyramid approach to subpixel registration based on intensity. IEEE Trans Image Process 1988;7:27–41.

Wechsler, D. (1997). *Wechsler Adult Intlelligence Scale, Third Edition* San Antonio, Tx: The Psychological Corporation.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence*, Revised. San Antonio, TX: The Psychological Corporation.

Wilke M, Schmithorst VJ, Holland SK. Assessment of spatial normalization of whole-brain magnetic resonance images in children. Human Brain Mapping 2002;17:48–60. [PubMed: 12203688]

Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. Journal of Neuroscience 1994;14:1908–1919. [PubMed: 8158246]
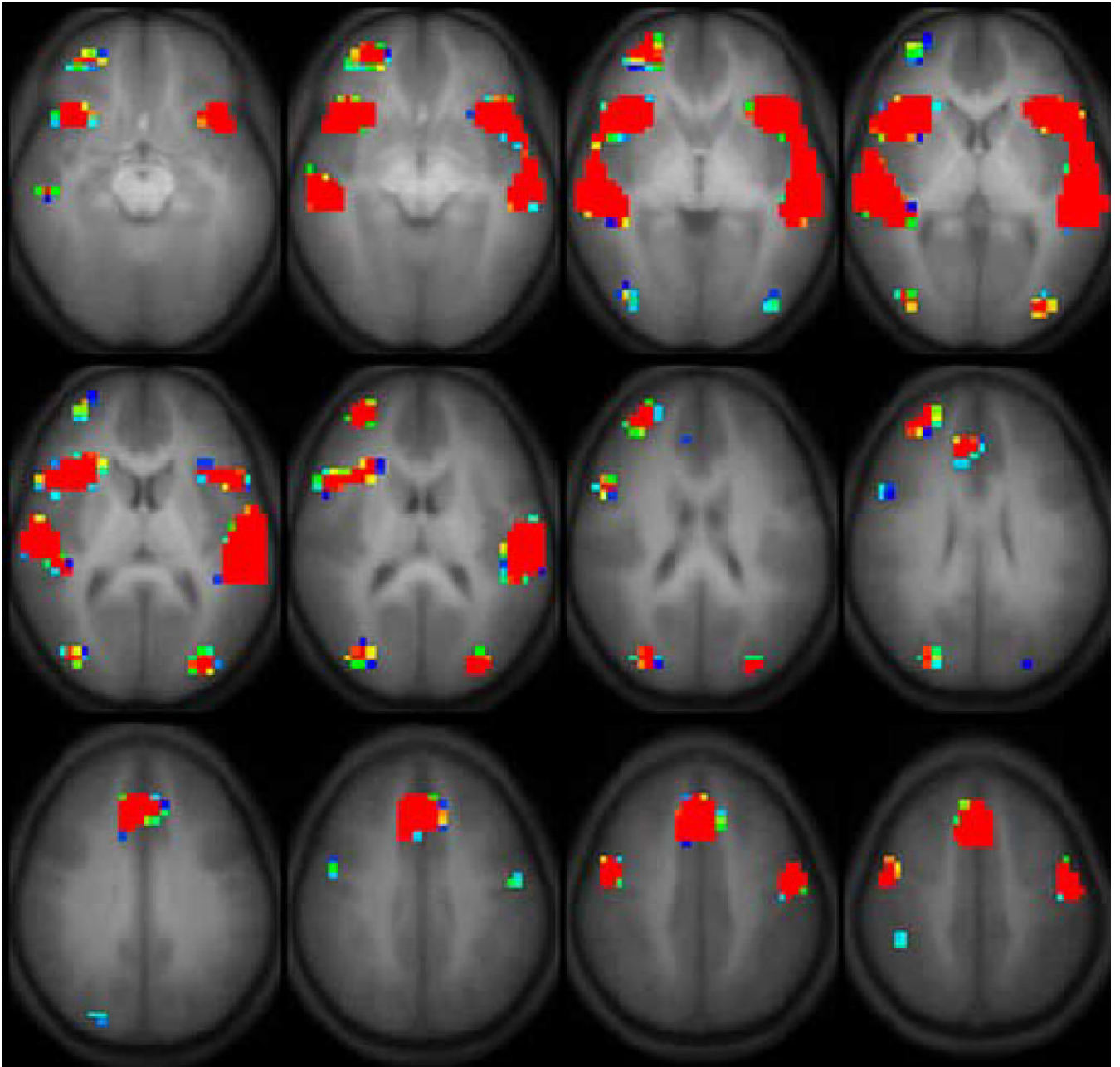
Figure 1.
Activation for regions associated with task performance (collapsed across age).
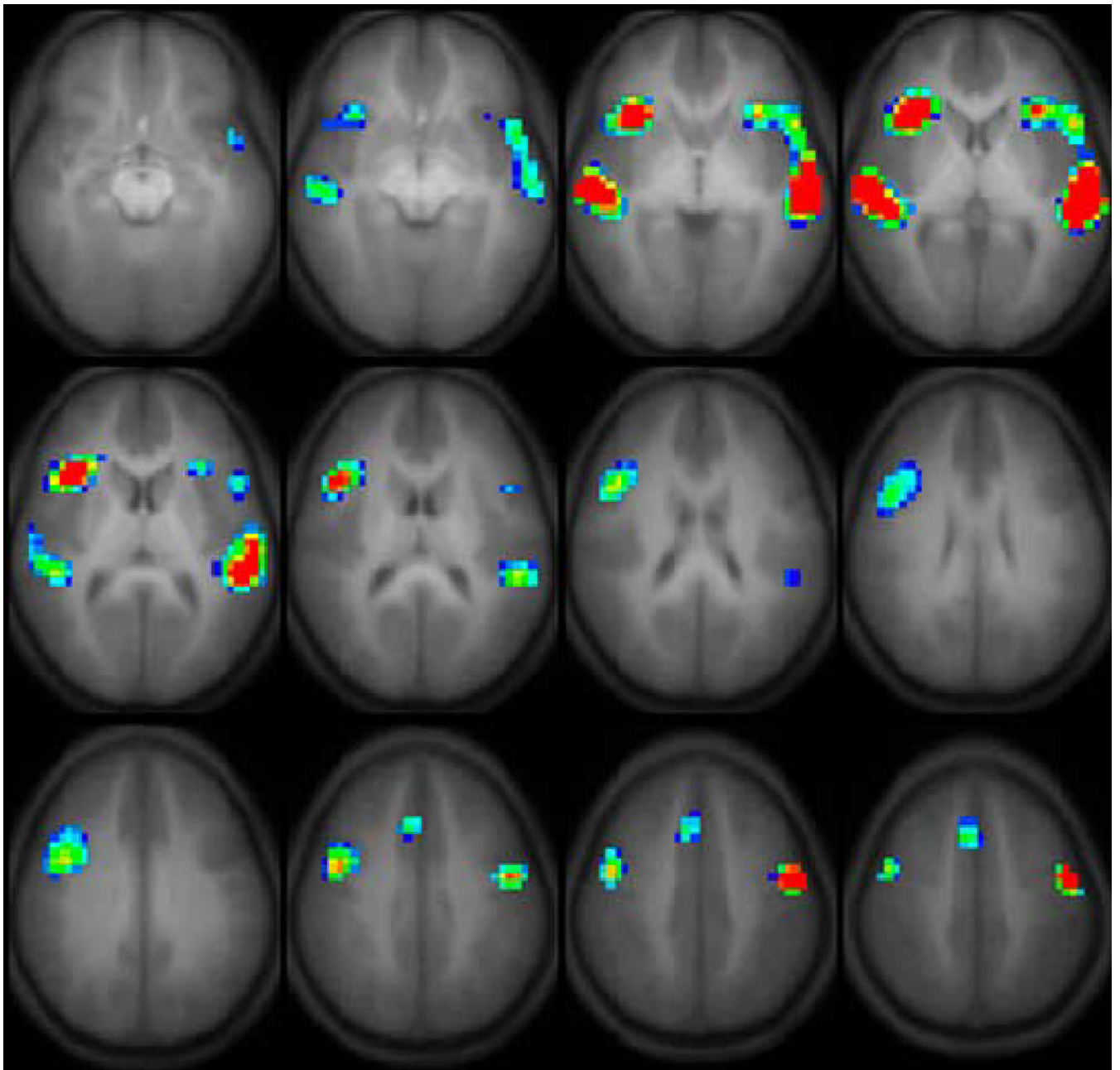
Figure 2.
Activation for regions associated with age.

Table 1

Demographic characteristics of participants

Age in years	Sex	Handedness	*Full Scale IQ*[1] **Mean (SD)**	*Language Quotient*[1] **Mean (SD)**
5	3 male 7 female	10 right	113.4 (4.6)	105.6 (15.1)
6	7 male 6 female	11 right 2 left	113.3 (13.6)	108.6 (13.2)
7	12 male 11 female	23 right	116.5 (13.5)	106.5 (9.4)
8	10 male 13 female	22 right 1 mixed	114.4 (15.4)	102.8 (11.2)
9	12 male 10 female	21 right 1 left	112.5 (14.7)	107.5 (15.0)
10	12 male 13 female	22 right 3 left	112.8 (17.1)	106.4 (16.9)
11	9 male 17 female	23 right 2 left 1 mixed	109.04 (10.2)	103.8 (13.2)
12	13 male 16 female	25 right 4 left	112.5 (11.6)	110.0 (16.8)
13	14 male 13 female	24 right 3 left	106.2 (11.7)	107.6 (10.7)
14	7 male 11 female	16 right 2 left	106.6 (15.8)	109.3 (17.9)
15	11 male 9 female	20 right	105.9 (15.5)	106.9 (12.5)
16	9 male 8 female	17 right	104.4 (14.0)	108.2 (16.1)
17	9 male 11 female	20 right	111.7 (12.7)	112.3 (16.7)
18	12 male 3 female	13 right 2 left	114.9 (13.3)	113.8 (13.3)

[1]Test scores scaled with a normative mean of 100, SD=15.

Table 2

Proportion of accept responses to filtered sentences

Age in years	Correct Accepts Mean (range)	False Positives Mean (range)	d′
5	.41 (0–.76)	.30 (0–.57)	0.42
6	.41 (0–.69)	.34 (.06–.66)	0.28
7	.55 (.15–.92)	.33 (.06–.54)	0.67
8	.54 (.08–.92)	.26 (.03–.69)	0.84
9	.66 (.15–.92)	.30 (.06–.66)	1.05
10	.65 (.31–.92)	.25 (.06–.54)	1.16
11	.69 (.54–.92)	.24 (.06–.45)	1.28
12	.72 (0–1.0)	.18 (0–.40)	1.70
13	.72 (.31–1.0)	.17 (0–.37)	1.64
14	.71 (.23–1.0)	.23 (.03–.57)	1.43
15	.68 (.23–.92)	.20 (.09–.43)	1.43
16	.69 (0–1.0)	.14 (0–.27)	1.80
17	.73 (.38–.92)	.14 (.02–.37)	1.83
18	.74 (.07–.92)	.13 (.03–.26)	1.93

Table 3

Regions of interest active during the prosody processing task.

Primary Location	Hemisphere	Talairach xyz Coordinates	Inferior-Superior Extent
Group Map			
inferior frontal gyrus	right	40, 18, 3	−10, 15
	left	−43, 16, 1	−10, 15
middle frontal gyrus	right	28, 47, 11	−5, 25
precentral sulcus	right	46, −3, 41	35, 45
	left	−46, −8, 41	35, 45
cingulate gyrus	midline	2, 25, 36	25, 45
superior temporal gyrus	right	51, −22, 3	−5, 10
	left	−53, −23, 6	−5, 15
anterior occipital	right	31, −77, 14	0, 30
	left	−32, −80, 10	0, 20
Age Map			
anterior insula	right	31, 20, 4	−5, 10
	left	−32, 21, 3	0, 10
precentral sulcus	right	41, 5, 33	25, 45
	left	−47, −5, 39	35, 45
superior temporal gyrus	right	51, −23, 5	−5, 10
	left	−55, −25, 6	−5, 15

Table 4

Regions of interest that predict behavioral performance and age (in months) for children ages 5 to 18 years.

Prediction	Region of Interest	Parameter estimate[*]	F value	P value
Correct Accepts (controlled for age)	**Full model**		**11.50**	**.0001**
	Left Precentral Sulcus	0.07	1.25	.2638
	Left Superior Temporal Gyrus	0.16	5.74	.0172
	Cingulate	−0.09	1.83	.1771
	Right Middle Frontal Gyrus	0.07	1.27	.2608
False Positives (controlled for age)	**Full model**		**15.43**	**.0001**
	Left Superior Temporal Gyrus	−0.23	13.92	.0002
	Right Anterior Insula	−0.14	3.64	.0575
	Right Precentral Sulcus	0.10	2.63	.0575
	Cingulate	0.09	1.92	.1666
	Right Middle Frontal Gyrus	0.07	1.15	.2848
Age in months (controlled for task performance)	**Full model**		**15.43**	**.0001**
	Left Anterior Occipital	0.14	4.52	.0344
	Right Anterior Occipital	−0.14	4.46	.0356
	Right Middle Frontal Gyrus	−0.18	11.30	.0009
	Left Precentral Sulcus	0.08	1.94	.1650
	Right Precentral Sulcus	0.25	16.04	.0001
	Left Superior Temporal Gyrus	0.10	3.39	.0665

[*] standardized