This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Anatomical correlates of dynamic auditory processing: Relationship to literacy during early adolescence

Mary Elizabeth Sutherland, Robert J. Zatorre, Kate E. Watkins, Pierre-Yves Hervé, Gabriel Leonard, Bruce G. Pike, Caroline Witton, Tomáš Paus

A B S T R A C T

Adults show great variation in their auditory skills, such as being able to discriminate between foreign speech-sounds. Previous research has demonstrated that structural features of auditory cortex can predict auditory abilities; here we are interested in the maturation of 2-Hz frequency-modulation (FM) detection, a task thought to tap into mechanisms underlying language abilities. We hypothesized that an individual’s FM threshold will correlate with gray-matter density in left Heschl’s gyrus, and that this function–structure relationship will change through adolescence. To test this hypothesis, we collected anatomical magnetic resonance imaging data from participants who were tested and scanned at three time points: at 10, 11.5 and 13 years of age. Participants judged which of two tones contained FM; the modulation depth was adjusted using an adaptive staircase procedure and their threshold was calculated based on the geometric mean of the last eight reversals. Using voxel-based morphometry, we found that FM threshold was significantly correlated with gray-matter density in left Heschl’s gyrus at the age of 10 years, but that this correlation weakened with age. While there were no differences between girls and boys at Times 1 and 2, at Time 3 there was a relationship between gray-matter density in left Heschl’s gyrus in boys but not in girls. Taken together, our results confirm that the structure of the auditory cortex can predict temporal processing abilities, namely that gray-matter density in left Heschl’s gyrus can predict 2-Hz FM detection threshold. This ability is dependent on the processing of sounds changing over time, a skill believed necessary for speech processing. We tested this assumption and found that FM threshold significantly correlated with spelling abilities at Time 1, but that this correlation was found only in boys. This correlation decreased at Time 2, and at Time 3 we found a significant correlation between reading and FM threshold, but again, only in boys. We examined the sex differences in both the imaging and behavioral data taking into account pubertal stages, and found that the correlation between FM threshold and spelling was strongest pre-pubertally, and the correlation between FM threshold and gray-matter density in left Heschl’s gyrus was strongest mid-pubertally.

© 2012 Elsevier Inc. All rights reserved.

Introduction

The ability to read and write – literacy – is largely dependent on a person’s knowledge of the speech sounds used in a particular language (phonology). This phonological awareness is dependent on basic auditory processes, such as following the acoustic modulation present in speech. These processes require good temporal resolution to track the small and rapid changes in sounds that characterize phonemes and alter the meanings of words. Reading ability correlates with auditory processing, particularly with tasks that require acuity in such temporal processing (Ahissar et al., 2000; Au and Lovegrove, 2007, 2008; Boets et al., 2008; Witton et al., 1998; Witton et al., 2002). Specifically, the ability to detect amplitude and frequency modulation (AM and FM respectively) of auditory stimuli predicts phonological awareness measured by non-word reading, in both normal and dyslexic readers, even though the latter group had higher detection thresholds (Witton et al., 1998).

When a carrier tone is modulated in frequency (FM) at a rate of 2 Hz, it results in a highly perceptible change in pitch that most listeners are able to track. When the modulation rate is raised to 40 Hz, the changes are perceived as roughness, and at 240-Hz FM people perceive a tone at the modulating frequency (i.e. 240 Hz).
Thus, perception of 2- and 40-Hz FM depend on the ability to process changes over time and for this reason are often used to determine sensitivity to the basic auditory processes thought to underlie phonological ability. In fact, a strong correlation between phonological awareness and 2-Hz FM detection threshold has been shown in children at the age of 10 years, even after controlling for intelligence and overall reading ability (Talcott et al., 2000). Detection threshold of FM was determined using a task that requires listeners to judge which of two pure tones with the same carrier frequency (i.e. perceived to have the same fundamental pitch) has been frequency-modulated, and the modulation depth was adjusted according to an adaptive staircase procedure. The correlation between performance on this task and literacy abilities has been hypothesized to lie in the nature of the 2-Hz FM processing task, which is dependent on temporal processing. These relationships must have an underlying neural substrate; however, the anatomical correlates of FM processing have not yet been investigated.

Functional imaging studies have suggested that the right and left hemispheres are specialized such that right auditory regions have higher spectral resolution while left auditory regions have higher temporal resolution (Zatorre and Belin, 2001; Zatorre and Gandour, 2008; Zatorre et al., 2002). This conclusion has been supported by a number of studies that use both speech and non-speech stimuli (Boemio et al., 2005; Brancucci et al., 2008; Brechmann and Scheich, 2005; Giraud et al., 2007; Jamison et al., 2006; Okamoto et al., 2009; Schonwiesner et al., 2005). These converging results suggest that auditory stimuli that are characterized primarily by their temporal variation are processed similarly, regardless of whether they convey any linguistic meaning (Zaehle et al., 2004). These functional hemispheric differences exist at a prelexical level of speech processing (Sandmann et al., 2007) and even extend to the auditory brainstem (Horneckel et al., 2009). Thus, if the 2-Hz FM detection task is related to literacy because it measures acuity in the temporal domain, then we would expect performance on this task to be linked to processing in primary auditory cortex in the left hemisphere. Primary auditory cortex is located on Heschl’s gyrus (Rademacher et al., 2001), and shows significant hemispheric asymmetries in anatomy (Penhune et al., 1996; Sigalovsky et al., 2006). Since the function of this region is thought to be related to its structure (Zatorre et al., 2002), we would expect the morphology of left Heschl’s gyrus to correlate with an individual’s FM detection threshold.

The link between structural variation in Heschl’s gyrus and behavioral abilities has already been demonstrated in adults: for example, processing of temporally modulated stimuli correlates with the volumes of left Heschl’s gyrus, whereas spectral processing correlated with the volume of right Heschl’s gyrus (Warrier et al., 2009). Also, the ability to learn to recognize different pitch contours is related to the size of Heschl’s gyrus (Wong et al., 2008). White-matter density in left Heschl’s gyrus predicts how well an individual learns to discriminate foreign speech sounds (Golestani et al., 2007).

All of the aforementioned studies have been carried out in adults. But in studying structure–function relationships, one also needs to take into account age, and in the case of children and adolescents, maturational processes. The studies that have shown behavioral correlations between literacy ability and 2-Hz FM detection have all tested children in primary school (Talcott et al., 2002), when the auditory cortex is not yet mature functionally (Ponton et al., 2000; Poulsen et al., 2009). Generally, adolescence is marked by global and local age-related variations in both white and gray matter (Paus, 2005). Thus, the relationship between FM threshold and its underlying neural correlates and, correspondingly, the relationship between FM threshold and literacy ability may change during this period of human development.

Here, we used a longitudinal design to explore the anatomical correlates of FM processing and their maturation through early adolescence; 63 healthy children were followed from the age of approximately 10 to 13 years. Our first aim was to replicate the previous finding in 10-year-olds that the ability to detect 2-Hz FM correlates with certain literacy skills dependent on phonological processing, such as spelling, non-word reading, and reading (Talcott et al., 2000). Secondly, we explored how the relationship between the 2-Hz FM detection threshold and literacy ability changed over the adolescent period. Thirdly, we used voxel-based morphometry (Ashburner and Friston, 2000) to test our hypothesis that the ability to detect 2-Hz FM correlates with gray-matter density in left Heschl’s gyrus. Finally, we explored how that relationship might evolve over time.

Materials and methods

Participants

Sixty-five healthy English-speaking 10-year-old children were recruited for a longitudinal study of normal adolescent development, but only 63 completed both imaging and behavioral measures (see Table 1). The study involved a 2-day visit during which brain imaging studies, psychometric assessment and behavioral testing were carried out. Children and parents also completed a number of questionnaires. The children were seen again 18 (Time 2) and 36 months (Time 3) after the initial visit (Time 1). All participants passed a neurological screening test. Written informed consent (parents) and assent (children) was obtained in accordance with the ethical approval granted by the Montreal Neurological Institute Research Ethics Board.

All participants completed the Puberty Developmental Scale (Petersen et al., 1988), an eight-item self-report measure of psychological development based on Tanner stages with separate forms for males and females. There are five categories of pubertal status used in this scale: (1) pre-pubertal, (2) beginning pubertal, (3) mid-pubertal, (4) advanced pubertal, (5) post-pubertal. Participants answer questions about their growth in stature and pubic hair, as well as menarche in females and voice changes in males. These self-report measures have been found to correlate significantly with physician ratings of pubertal development (Dorn et al., 1990).

Psychometric assessments

The children’s reading and spelling (literacy) abilities were assessed using Woodcock-Johnson (WJ) Achievement subtests (spelling, non-word reading, 1 and passage comprehension) at each time point (Woodcock et al., 2001). We calculated cognitive abilities using the Similarities and Digit Span subscales of the WISC (Wechsler Intelligence Scales for Children) III for Times 1 and 3 and WISC IV for Time 2 (to assess general verbal abilities) and the Matrices subtest of the WISC IV (to assess general non-verbal abilities). We chose to alternate between the WISC III and WISC IV to decrease the possibility of familiarity and practice effects influencing the scores. One of the differences between the WISC III and WISC IV is that the WISC III does not contain the Matrices subtest, so these data were only collected at Time 2. These three subtests were chosen to replicate the measures used by (Talcott et al., 2000), who chose them as they represent three of the four subtests commonly used to determine a child’s short form intelligence quotient. In our study, we administered the complete WISC III and IV, and we report the IQV (verbal intelligence quotient) and PIQ (performance intelligence quotient) in Table 1. These measures were computed from the full complements of the relevant subtests.

Psychophysical paradigm

All participants completed an audiometric hearing test before data collection at each time point, to ensure that they had normal hearing.

1 Non-word reading refers to the “Word Attack” subtest of the Woodcock Johnson reading tests and measures the ability to read nonsense words aloud.
This entailed reporting correctly all the sounds presented in the frequency range typically spanned by speech (1000–4000 Hz) in at least one ear at 25 dB of the audiometric threshold.

The detection threshold for a 2-Hz FM tone with a 1-kHz tonal carrier was measured using a two-interval, two-alternative forced-choice paradigm at each time point (Talcott et al., 2000). The target always contained the FM signal (defined in Eq. (1)) and the other tone was a pure tone (defined in Eq. (2)). The order of the two tones was randomized.

$$S_1(t) = A \sin(2\pi f_c t + \beta \sin(2\pi f_m t))$$  \hspace{1cm} (1)$$

$$S_2(t) = A \sin(2\pi f_c t)$$  \hspace{1cm} (2)$$

where \(S(t)\) is the signal as a function of time, \(A\) is the amplitude of the signal, \(f_c\) is frequency of the carrier (1 kHz in this case), \(f_m\) is modulation rate, \(t\) is time, \(\beta\) is the modulation index (the modulation depth expressed as the frequency deviation in Hz divided by the modulation rate). The onset and offset of the sounds were gated with a 20 ms Gaussian rise and fall times. Each sound lasted for 1000 ms and was separated from the other tone in the pair by 500 ms of silence. Stimuli were presented over Plantronics head-phones at a comfortable listening level.

FM testing was preceded by six practice trials that were used to familiarize the participants with the task and to adjust the starting point. The starting modulation depth was set at 6. This number refers to the extent of the sinusoidal deviation of the frequency from that of the carrier tone (in this case, 1 kHz), and represents a ratio of the peak deviation from the carrier divided by the modulating frequency (and therefore has no units). If a participant failed at this level, then the depth of the modulation was increased from 6 to 10. The data from the practice trials were not used to calculate the participant’s threshold of detection. The actual test followed an adaptive staircase procedure (Kaernbach, 1991). Six trials at the starting level were also randomly interspersed as “catch trials” to make sure that the participants were paying attention and doing the task accurately. Each trial began with a screen depicting two different colored birds, each of which corresponded to a tone (either the target FM tone or the pure tone). As a tone played, the birds moved. The children were instructed to click on the bird corresponding to the target tone. When they answered correctly, they were given feedback in the form of a brightly colored egg that appeared on the side of the screen next to the birds and the modulation index of the following trial was reduced by 1 dB (i.e. a factor of 1.122). When they answered incorrectly, the modulation index was raised by 3 dB (a factor of 1.412).

The staircase terminated after 10 reversals and each participant’s detection threshold was calculated using the geometric mean of the last 8 reversals. The average number of trials at Time 1 was 43 (mode = 39), with a range of 31–115. At Time 2, the average was also 43 (mode = 38) but the range was reduced to 32–61 trials. At Time 3, the average was 45 trials (mode = 40) with a range of 31–58 trials.

For the behavioral analyses, we used data from all participants assessed, including 2 children who were not scanned, which brought our total participant count to 65. It should be noted, however, that the numbers of participants contributing to different analyses at different time points varies. At Time 1, data from 4 participants were excluded from the FM detection task due to failure on the task (detection thresholds above the starting size of 6 coupled with a failure to detect more than 20% of the catch trials) and data obtained in a further 4 were corrupted or missing. Data from 13 participants were missing for the WJ passage comprehension subtest. Data were available for all participants on the WJ spelling subtest and the WJ non-word reading subtest. At Time 2, data were missing from 2 of the participants on the FM detection task, from 4 participants on the WJ passage comprehension subtest, and from 1 participant on the WJ spelling subtest. All participants (65) completed the WJ non-word reading subtest at Time 2. At Time 3, data were missing from 10 participants on the FM detection task, from 5 participants on the WJ passage comprehension subtest, from 3 on the WJ spelling subtest, and from 4 on the WJ non-word reading subtest.

The data for each behavioral test were first checked for normality of the distribution using the Kolmogorov–Smirnov test. Because the data were non-normally distributed, non-parametric statistics were used for further analyses. Specifically, we calculated the Spearman’s rank correlation coefficients for the auditory FM threshold and literacy scores at each time point separately. Separate correlations for boys and girls were calculated at each time point.

**Image acquisition, preprocessing, and analysis**

Magnetic resonance imaging (MRI) data were collected using a Siemens 1.5-T superconducting magnet. High-resolution T1-weighted (T1-W) anatomical images were acquired using the following sequence: 3D RF-spoiled gradient echo scan with 140–160 slices, 1-mm isotropic resolution, TR = 22 ms, TE = 9 ms, flip angle = 30°. After non-uniformity correction using N3 (Sled et al., 1998) and brain extraction (using Brain Extraction Tool from the FMRIB Software Library), each of the T1-W scans from Time 1 and Time 3 was co-registered to the corresponding Time 2 T1-W scan. This was done with a 9-
parameter linear transform, of which only the 6 parameters relating to the translations and rotations (i.e. rigid body) were kept. This allowed preservation of the totality of the morphological differences between scans, including the size difference, while centering the brains appropriately across scans within each individual. Then, for the group stereotactic average computation, each of T1-W scans of Time 2 was linearly registered (9 parameters) to the average registration template of Time 2 (which will be referred to as the SFIC template), obtained via several iterative non-linear registrations. The first iteration was performed with a different template, and the average of the obtained set of registered SFIC-Time 2 images was then used as the “target” of the second iteration. The registration parameters of Time 2 images to the stereotactic SFIC space were then combined with the rigid registration parameters of Time 1 and Time 3 data to the “native space” of the Time 2, so as to register the Time 1 and Time 3 T1-W structural scans to the SFIC template space. The Time 2 data were thus used as a “reference point”. We then applied a tissue classification algorithm (Cocosco et al., 2003) to generate gray matter (GM), white matter (WM) and cerebrospinal fluid images (CSF), to all the SFIC-template registered images from the three visits. These binary tissue images were then smoothed with a 10-mm Gaussian filter.

The smoothed gray-matter tissue maps were then analyzed statistically using the surfstat (Worsley et al., 2009) version of voxel-based morphometry (VBM), a technique that measures local differences in gray-matter density over the whole brain. For each time point, regression analysis using a general linear model was applied to each voxel in the stereotactic space to determine whether there was a relationship between the behavioral variable (2-Hz FM detection threshold) and gray-matter density. We had a strong a priori hypothesis that this relationship would be maximal in auditory cortex, in particular voxels in left Heschl’s gyrus (see Introduction). Therefore, two separate masks were created from probability maps of Heschl’s gyrus in each hemisphere (Penhune et al., 1996) and used to restrict the search volume to set a small-volume statistical significance threshold for the t-statistic map. Using this mask resulted in a t-threshold value of 3.67 at p<0.05, corrected for multiple comparisons within the restricted search volume. For display purposes, we selected the peak voxel at each time point from both right and left Heschl’s gyrus and plotted the gray matter values obtained at these voxels with performance on the FM detection task (see Figs. 2a–c). The Euclidean distance between the peak voxels in left Heschl’s gyrus was 2.8 mm (between Times 1 and 2) and 6.6 mm (between Times 2 and 3), and in right Heschl’s gyrus 5.6 mm (between Times 1 and 2) and 9.4 mm (between Times 2 and 3).

For the imaging analyses, data were available for 63 children. At Time 1, data from 18 of these participants were excluded due to one or a combination of the following: excessive head movement during scanning (4), image artifacts caused by dental braces (4), missing or corrupted imaging data (5) and incomplete behavioral data (7). The remaining 45 children (24 males, 21 females) had a mean age of 10.05 years (SD 0.38 years). At Time 2, all 63 children were scanned again. Data from 3 participants were excluded due to image artifacts caused by dental braces. The remaining 60 children (31 males, 29 females) had a mean age of 11.55 years (SD 0.4 years). At Time 3, 60 of the children were scanned. Data from 10 participants were excluded due to image artifacts caused by dental braces (4) or corruption of the imaging data (6). The remaining 50 children (27 males, 23 females) had a mean age of 13.01 years (SD 0.42 years).

Results

Behavioral results — Time 1

At Time 1, data were available for the 2-Hz auditory FM detection threshold, WJ spelling and non-word reading subtests in 57 of the 65 children tested (age: 10 years). We used one-tailed p-values as our hypotheses regarding these data had a clear directionality: lower thresholds (i.e. better FM detection) should be related to higher scores on the literacy subtests. We also applied a Bonferroni correction for multiple (18) comparisons, which changed our (uncorrected) significance level to p = 0.003 (corrected p = 0.05). Analysis of the whole-group data using this correction revealed no significant correlations between FM threshold and the scores on either the WJ spelling subtest (Spearman’s rho = −0.261; p = 0.025; Fig. 1a) or the WJ non-word reading subtest (rho = −0.228, p = 0.043; Fig. 1b). We also separated our sample by sex and recalculated these correlations. This revealed that the whole-group correlations were primarily driven by the data obtained in boys (n = 29), who showed strongly significant correlation for FM threshold with non-word reading (rho = −0.593, p = 0.0005), which was significantly different (determined using a Fischer r–z transform) from the correlation in the girls: rho = −0.015, p = 0.47 (n = 28). The boys also showed a strongly significant correlation between FM threshold and spelling (rho = −0.537, p = 0.0015), whereas the girls showed a very weak correlation (rho = −0.156, p = 0.215). In this case, however, the correlations between the boys and girls were not different significantly (p = 0.11). Data were available for the FM threshold test and WJ passage comprehension subtest in 46 children. The relationship between the scores on these two measures was not significant in the analysis of the whole group or for the analyses of the boys and girls separately (Fig. 1c). We report two-tailed p-values for our comparison of boys vs. girls as these tests were done post hoc; boys had lower median scores than girls on the literacy tests and FM threshold, but these differences were close to significance only for the spelling (Mann–Whitney U test, p = 0.059) and non-word reading (p = 0.061) subtests (see Table 1). The boys and girls did not differ in VIQ and PIQ or in their subtests scores on Similarities and Digit Span (see Table 1).

Behavioral results — Time 2

At Time 2, data were available for the FM detection threshold, WJ spelling and non-word reading subtests for 62 of the 65 children tested (age: 11.5 years). The correlation between FM threshold and spelling decreased (Fig. 1d), while the correlation between FM threshold and non-word reading stayed stable (Fig. 1e). Although there was a slight drop in the magnitude of the correlation between FM threshold and spelling at Time 2, this was not significant. Separate correlations for the boys revealed that the correlation between FM threshold and both spelling and non-word reading decreased (Spelling: rho = −0.325, p = 0.040; n = 30; non-word reading: rho = −0.390, p = 0.015, n = 31), but again, these decreases were not significant (Spelling: p = 0.34; non-word reading: p = 0.32). As seen at Time 1, these correlations in the girls were also not significant at Time 2 (Spelling: rho = −0.170, p = 0.175, n = 32; non-word reading: rho = −0.103, p = 0.29, n = 31), and also did not differ significantly from the Time 1 correlations (Spelling: p = 0.96, non-word reading: p = 0.749). Data were available for the FM threshold test and WJ passage comprehension subtest in 59 children. The relationship between the two was not significant at the whole-group level or for the boys and girls separately (Fig. 1f). There were no significant differences between boys and girls on any of the scores for the intelligence and literacy tests and detection thresholds at Time 2 (see Table 1).

Behavioral results — Time 3

At Time 3, data were available for FM threshold and WJ spelling in 55 children, for FM threshold and WJ non-word reading in 54 children and for FM threshold and WJ passage comprehension in 53 children (age: 13 years). The correlations between FM threshold and scores on any of the literacy subtests were not significant at Time 3 (see Figs. 1g–i) and not significantly different from the other Time points. Separate analyses for boys and girls revealed that all
The correlation between FM threshold and gray-matter density in left Heschl's gyrus did not differ according to sex (boys $r = -0.451$; girls $r = -0.561$; see Fig. 3a). The correlations on the right were not significant in either sex (see Fig. 3b).

**Gray-matter density — Time 2**

At Time 2, VBM was used to correlate gray-matter density and FM threshold in 62 children aged 11.5 years. There were no significant correlations between gray-matter density and 2-Hz FM detection threshold in Heschl's gyrus or elsewhere in the brain. Plotting the gray-matter density of the voxel in Heschl's gyrus with the highest t-value for the correlation against FM detection threshold revealed a reduced slope of the regression line on the left side ($r = -0.26$) compared with the Time 1 data, though a comparison of the correlation coefficients using an r-z transform showed no statistical difference between these slopes ($p = 0.114$). There was no slope on the right side ($r = 0.02$; see Figs. 3c and 3d), and no effect of sex on the correlation between FM threshold and gray-matter density in the left Heschl's gyrus (boys: $r = -0.285$; girls: $r = -0.197$, $p = 0.73$; see Fig. 3c).

**Gray-matter density — Time 3**

At Time 3, VBM was used to correlate gray-matter density and FM threshold in 50 children aged 13 years. There was no significant
globally different value than at Time 2. This may be due to partial volume effects: the mix of gray matter and white matter influences the outcome for the voxel value in such a way that it is not very meaningful to compare one location with the other. The numbers we are interested in are the correlation coefficients, relative values not dependent on the absolute value of the voxel, but rather on the relationship with FM threshold. Therefore, we cannot make any claims about this global difference in voxel values at Time 3.

To summarize, 2-Hz auditory FM detection ability correlates with gray-matter density most strongly at age 10 and in the left hemisphere (see Figs. 2 and 3). The strength of the relationships between FM threshold and gray-matter density decreased over time points when all participants (boys and girls) were analyzed as one group. At Time 3 only, a comparison of the correlation coefficients using an r-z transform showed the correlation between 2-Hz auditory FM threshold detection ability and gray-matter density in left Heschl’s gyrus to be significantly stronger in the boys than the girls (p < 0.01). The boys showed no change in correlation coefficient comparing Times 1 and 2 (p = 0.5), 2 and 3 (p = 0.17), or 1 and 3 (p = 0.53). The girls showed no significant change between Times 1 and 2 (p = 0.16) or 2 and 3 (p = 0.25) but they did show a change when comparing Time 1 with Time 3 (p = 0.02). These direct comparisons across time points confirm the cross-sectional correlation results given above, and demonstrate that the results for the boys are relatively stable, whereas the girls show a change in the structure–function relationship over time.

Further analyses examining sex differences

A possible reason for observing sex differences in the correlations of FM threshold with literacy measures and with gray-matter density in this age range could be that the males were less sexually mature than the females at each time point (see Table 1). We attempted to tease apart the effects of puberty and sex at Time 1 when the correlations between FM threshold and spelling and FM threshold and gray-matter density were the strongest. We ran separate correlations for groups of children with scores of 1 (pre-pubertal) and 3 (mid-pubertal) on the Puberty Development Scale (Petersen et al., 1988). For FM threshold and spelling, the correlation was stronger in the pre-pubertal group (r = −0.452, p = 0.035, n = 22, 6 females, 16 males) than for the mid-pubertal group (r = −0.114, p = 0.643, n = 19, 15 females, 4 males). For FM threshold and gray-matter density in left Heschl’s gyrus, the correlation was stronger and highly significant in the mid-pubertal group (r = −0.703, p = 0.005, n = 14, 11 females, 3 males) than the pre-pubertal group (r = −0.414, p = 0.07, n = 19, 5 females, 14 males). These two correlation coefficients did not differ significantly (p = 0.269), however. But note that at age 10, the number of females and males in the different puberty-based groups is very different: 6 females and 16 males in the pre-puberty group and 15 females and 4 males in the mid-pubertal group. Therefore, it is impossible to dissociate these effects from those related to the above sex differences in these structure–function relationships. The correlation between FM threshold and gray-matter density in left Heschl’s gyrus at Time 3 that showed a strongly significant effect in the boys but not the girls was also examined with respect to the puberty scale. In this case, separating the children according to whether they were in the mid-pubertal stage (5 females, 11 males) or advanced puberty (10 females, 4 males) did not reveal any significant correlations between FM threshold and gray-matter density. This suggests that this relationship at Time 3 is not driven by sexual maturity, though we cannot be sure as the non-normal distribution of the data prohibited us from directly comparing the two correlation coefficients.

We also tested the hypothesis that the sex difference we were seeing was the result of the peak voxel being located at a different place in Heschl’s gyrus in girls and boys; a VBM analysis of each group

relationship between the gray-matter density in left Heschl’s gyrus and FM threshold (see Fig. 3e). The scatterplot of gray-matter density at the voxel in Heschl’s gyrus with the highest t-value and FM threshold score. The peak in the left Heschl’s gyrus is marked with a cross. These data are overlaid on a T1-weighted average image template of participants at age 11.5.

**Fig. 2.** Correlation between gray-matter density in Heschl’s gyrus and 2-Hz FM detection threshold at age 10 years. The left side of the axial and coronal images corresponds to the left hemisphere. Colored areas represent the value of the t-statistic for voxels where gray-matter density significantly correlates with the auditory FM threshold score. The peak in the left Heschl’s gyrus is marked with a cross. These data are overlaid on a T1-weighted average image template of participants at age 11.5.

-2

-4.1
separately showed this not to be the case (i.e. the peak voxel in girls and boys occurred at approximately the same place at each time point for girls and boys: the Euclidean distance at Time 1 was 3.16 mm, 2.22 mm at Time 2, and 2.45 mm at Time 3).

**Discussion**

Our results provide evidence both for a correlation between literacy and auditory processing in the temporal domain, and for possible neuroanatomical underpinnings of this link: namely that gray-matter density in left Heschl's gyrus explains some of the variance in individual differences for auditory 2-Hz frequency modulation detection. These findings change depending on the age and the sex of the participants. In particular, gray-matter density in left Heschl's gyrus predicts the FM detection threshold for boys and girls at age 10 years but not at 11.5 years. At age 13 years, only the boys show a significant correlation between FM threshold and gray-matter density. The decrease in the relationship seen between gray-matter density in left Heschl's gyrus and FM threshold across time points when considering the whole group is reflected in the behavioral correlations between literacy measures and auditory 2-Hz FM detection threshold, which, especially in the case of spelling and non-word reading, also decrease with age in whole group analyses. For the behavioral correlations, however, the relationships between literacy measures and FM threshold are stronger in boys at all time points. The influence of puberty is difficult to tease apart from these sex differences, as the girls are more sexually mature at all ages. It seems, however, that the correlation found for the boys only at age 13 between the gray-matter density in left Heschl's gyrus and FM threshold is not related to sexual maturity.

**Literacy ability and FM threshold**

The relationship between spelling and non-word reading, both measures of phonological ability, and auditory temporal processing (measured via FM detection) is consistent with the literature (Boets et al., 2008). Our results suggest that while the ability to detect frequency modulation is correlated with phonological literacy ability, it is not correlated to the extent reported by Talcott et al. (2000). In fact, it seems that most of the studies investigating the relationship between literacy and various types of auditory processing tasks have found smaller correlations than those reported originally by Talcott et al. (2000), e.g. Talcott et al. (2002). Also, our results suggest that these relationships are subject to sex differences. Other research has suggested that while this relationship between literacy and FM detection threshold can be found, it is not causal, as only a subgroup of people with specific language impairment has trouble discriminating between different frequencies (Bishop et al., 1999). In fact, only some people with specific language disorders are impaired on auditory tasks that require temporal processing (Ahissar et al., 2000). Yet despite the question of directionality, our behavioral results provide correlations that justify our imaging hypothesis, namely that the structure of left Heschl's gyrus is related to FM threshold, which in turn is a reliable index of dynamic auditory processing, a skill necessary for various literacy tasks.

**The neuroanatomical correlates of FM processing**

Our study extends the research on these behavioral correlations by adding an imaging component, the results of which show that gray-matter density of regions previously shown to be engaged in auditory tasks that require temporal processing (e.g. Zatorre and Belin, 2001) can predict FM threshold at the age of 10 years. These results add to a growing body of literature suggesting that there is a relationship between neural structure, function and behavior. The causality and directionality of these relationships remains unknown, however. We still cannot confirm whether an initial difference in gray-matter density led to better FM detection thresholds, or increased exposure to auditory events that vary in time led to better FM detection threshold and increased gray-matter density. This question of causality is not unique to our study; it is also present, for example, in studies comparing musicians and non-musicians. In this latter domain, research suggests that while many neuroanatomical differences between musicians and non-musicians can be attributed to training-induced plasticity (Hyde et al., 2009), there may well be some inherent differences that are not due to musical training but influence musical ability (Foster and Zatorre, 2010). A similar combination of factors could contribute to our results: the differences in gray-matter density in...
left Heschl’s gyrus at age 10 might reflect both differing auditory experience as well as an initial difference in gray-matter density.

**Development**

We examined how the relationship between gray-matter density in left Heschl’s gyrus and FM detection threshold changed during early adolescence. The lack of correlation in both behavioral and imaging data seen at the later ages (11.5 and 13 years), especially in girls, may reflect a change in the neural resources needed to process rapid temporal auditory stimuli. In general, adolescence is marked by a continuous increase in white-matter volume, especially in male adolescents, and age-related decreases of gray matter in various cortical regions (Paus et al., 2008). The voxel values reported at Time 3 are lower than at Times 1 and 2, but that is due to the location of those voxels within left Heschl’s gyrus (6.6 mm distant from Time 2), not to a global decrease of gray matter. The FM threshold also improves (i.e. decreases) with age, though at age 10 better FM thresholds correlate with greater gray-matter density. It may be that our testing period encompasses the top of the maturation curve, and the subsequent gray matter “loss” (gray-matter density decreases as the participants age), which is most likely not constant between individuals, blurs the relationship between gray-matter density in left Heschl’s gyrus and FM threshold.

**Sex differences**

Another possible explanation for the change in relationship between gray-matter density in left Heschl’s gyrus and FM detection threshold is that they are associated with sex differences. Previous research has shown sex differences in the functional organization of language: males show engagement of the left inferior frontal gyrus during phonological tasks while females show engagement of both the left and right inferior frontal gyrus (Shaywitz et al., 1995). Females, compared with males, also appear to have greater interhemispheric connectivity when performing phonological tasks (Bitan et al., 2010). Such differences are not constrained to phonological processing, but rather have been found in other domains as well, for example, sex differences have also been found in singing (Hough et al., 1994), as well as in the neuroanatomical correlates of dyslexia (Sandu et al., 2008). Thus, it is not surprising that we also found sex differences in the correlation between gray-matter density in left Heschl’s gyrus and FM detection ability: overall, the relationship between gray-matter density in left Heschl’s gyrus and FM threshold was stronger in boys and was similar at Times 1 and 3 (though decreased at Time 2), while girls showed a weaker correlation at Times 2 and 3. Similar results were also seen for the relationship between literacy abilities (spelling and non-word reading) and FM threshold. We tested whether sexual maturation (as assessed using the Petersen scale of puberty) was correlated with the relationship between gray-matter density in left Heschl’s gyrus and FM threshold but our results pointed to no effect of puberty on this relationship (though, as expected, the girls were more mature than the boys at all time points). We also analyzed the whole-brain data for males and females separately, to determine whether the sex difference we were seeing was the result of the peak voxels occurring at a different position in boys versus girls, but this was not the case.

These observed sex differences are in line with longitudinal research showing that both overall brain size and cortical gray-matter follow sex-dependent trajectories, such that females show peaks in gray matter 1 to 2 years earlier than males (Lenroot et al., 2007). Males appear to have a steeper increase of white matter during adolescence, accompanied by decreases in the lobular values of magnetization-transfer ratio, an index of myelination (Perrin et al., 2009). These differences in neural development suggest that the differing correlation we observed between males and females may be due to expected dissimilarities in the developmental trajectory. There were no corresponding significant differences in literacy abilities, which is in line with research showing that sex-differences in brain “activation” are not necessarily associated with differing performance (Shaywitz et al., 1995). On the other hand, correlations between the “bulbosity” of the corpus callosum have been shown to correlate with better cognitive performance in women only, suggesting that the degree to which the hemispheres are connected have different implications for men and women (Davatzikos and Resnick, 1998). Yet we must be cautious in interpreting these differences, as they are morphological differences that do not explain a difference in behavior e.g. (Fine, 2010). This caution should extend to this study, though we found the correlation between FM threshold and literacy ability to differ between the sexes at all time points, including Time 1, which is not reflected in the imaging data (girls and boys show a similar correlation between gray-matter density in left Heschl’s gyrus and FM threshold).

**Conclusions**

We have demonstrated that the link between an individual’s auditory 2-Hz FM detection threshold and literacy ability exists on the background of the relationship between 2-Hz FM processing and gray-matter density in left Heschl’s gyrus, a region involved in the processing of temporally varying auditory stimuli. This relationship changes during early adolescence, and differs between the sexes. Our behavioral results are complementary: FM detection threshold is correlated with literacy abilities, but this correlation decreases with age and is different for boys and girls. Taken together, our results provide further support, at a neural level, for the hypothesis that literacy ability can be partially explained by FM detection threshold because the latter relies on temporal processing mechanisms, which, in turn, are linked to the structural properties of the left auditory cortex. We have also shown that this functional relationship, while quite stable at the age of 10 years, changes as the brain matures, though further studies are needed to assess the exact nature of this change.

**Acknowledgments**

This work was funded by the Sante Fe Institute Consortium, the Canadian Institutes of Health Research, the National Science Foundation, and the National Sciences and Engineering Research Council of Canada. We wish to thank Candice Cartier, Kristina Martini, Andree Hardy, Elissa Golden, Valerie Legge, and Line Gingras for assistance with the recruitment of participants and data collection.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2012.01.051.

**References**


