One year of musical training affects development of auditory cortical-evoked fields in young children

Takako Fujioka,1,3 Bernhard Ross,1 Ryusuke Kakigi,3 Christo Pantev4 and Laurel J. Trainor1,2

1Rotman Research Institute, Baycrest, University of Toronto, Toronto, 2Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Canada, 3Department of Integrative Physiology, National Institute for Physiological Sciences, Okazaki, Japan and 4Institute for Biomagnetism and Biosignal analysis, University of Münster, Münster, Germany

Correspondence to: Takako Fujioka, PhD, Rotman Research Institute, 3560 Bathurst Street, Toronto, Ontario, M6A 2E1, Canada
E-mail: tfujioka@rotman-baycrest.on.ca

Auditory evoked responses to a violin tone and a noise-burst stimulus were recorded from 4- to 6-year-old children in four repeated measurements over a 1-year period using magnetoencephalography (MEG). Half of the subjects participated in musical lessons throughout the year; the other half had no music lessons. Auditory evoked magnetic fields showed prominent bilateral P100m, N250m, P320m and N450m peaks. Significant change in the peak latencies of all components except P100m was observed over time. Larger P100m and N450m amplitude as well as more rapid change of N250m amplitude and latency was associated with the violin rather than the noise stimuli. Larger P100m and P320m peak amplitudes in the left hemisphere than in the right are consistent with left-lateralized cortical development in this age group. A clear musical training effect was expressed in a larger and earlier N250m peak in the left hemisphere in response to the violin sound in musically trained children compared with untrained children. This difference coincided with pronounced morphological change in a time window between 100 and 400 ms, which was observed in musically trained children in response to violin stimuli only, whereas in untrained children a similar change was present regardless of stimulus type. This transition could be related to establishing a neural network associated with sound categorization and/or involuntary attention, which can be altered by music learning experience.

Keywords: maturation; cortical plasticity; auditory cortex; musical training; magnetoencephalography

Abbreviations: AEF = auditory evoked magnetic field; AEP = auditory evoked potential; CI = confidence interval; MEG = magnetoencephalography; SEM = standard error of the mean

Received July 3, 2006. Revised August 10, 2006. Accepted August 11, 2006

Introduction

Music is an important part of human culture. Without training, culture-specific musical knowledge is acquired by infants (Trainor and Trehub, 1992, 1994). Just as vision (Lewis and Maurer, 2005) and language (Werker and Tees, 2005), musical acquisition is constrained by general developmental trajectories (Trainor, 2005). However, as opposed to language, musical abilities and literacy are emphasized far less and inconsistently included in the regular school curriculum. Consequently, musical training outside of school greatly enhances the spectrum of musical abilities. Becoming a successful musician often takes more than a decade of training commencing at an early age and includes regular practice for several hours daily (Ericsson et al., 1993; Sloboda and Davidson, 1996).

Recent neuroscience revealed that such intensive learning experiences involve changes in brain function and/or anatomy. Repeated practice optimizes neuronal circuits by changing the number of neurons involved, the timing of synchronization and the number and strength of excitatory and inhibitory synaptic connections. The effects of perceptual learning have been observed in animal electrophysiology across visual, sensory and auditory systems (see reviews, Edeline, 1999; Weinberger, 2004). In audition, behaviourally trained animals exhibit increased tonotopic organization of primary auditory cortex (Recanzone et al., 1993). The mere exposure to an enriched acoustic environment without training enhances auditory cortical responses and sharpens the tuning of auditory neurons,
even to un-experienced sounds in both young and old animals (Engineer et al., 2004), and such enhancements can last for days with different time courses of decay for different peak components. In summary, those studies show that training and acoustic environments have complex impacts on the auditory system.

The function of the human brain can be studied non-invasively by recording the electroencephalogram (EEG) or the magnetoencephalogram (MEG) extra-cranially, which both reflect the summation of electrical currents generated by the synchronous depolarization of neurons. In particular, the obligatory auditory cortical response to acoustic stimulation, termed the auditory evoked potential (AEP) or auditory evoked magnetic field (AEF), gives important information about when and where the sound is processed. In adults, long latency AEP/AEF typically consists of several positive and negative deflections such as the P1(m) (vertex positive AEP or AEF from upward oriented source at 50 ms in latency, ‘m’ stands for magnetic field response), N1(m) (100 ms), P2(m) (200 ms), occasionally N2(m) (250–300 ms) and the following sustained potential/field (SP/SF). These components are thought to have different origins and to reflect different types of processing. The recorded EEG and MEG contain complementary information (Hari et al., 1982). While EEG contains both laterally and tangentially oriented neural activities to the head surface from cortical and subcortical areas, distortion in signal due to volume conductions makes source localization difficult. In contrast, MEG measures only tangentially oriented neural activities from cortical areas and its source analysis is free from such distortion (Hämäläinen et al., 1993). This makes MEG especially suitable for investigating human auditory areas located in supratemporal plane and the supratemporal gyrus of which neurons are mostly perpendicular to the head surface.

The morphology of AEP/AEF takes 20 years to develop from the newborn to adult status (Courchesne, 1990; Ponton et al., 2000; Ponton et al., 2002). Typically, newborns exhibit a biphasic wave consisting of a positive peak at 200–300 ms followed by a negative peak at 400–600 ms, while an earlier positive peak P1 around 100 ms emerges in younger children (1–3 years of age) followed by two negative components at 250 ms and 450 ms (Barnet, 1975; Courchesne, 1990; Kushnerenko et al., 2002). After 3 years of age, an N1/P2 complex at 150 ms in latency emerges only with slow stimulation (Paetau et al., 1995; Bruneau et al., 1997; Sharma et al., 1997; Čeponiūnė et al., 1998). During school age, the latencies and amplitudes generally decrease, although different AEP/AEF components show different maturation rates, depending on the neuronal, synaptic and myelinafational maturation of the layer-specific neuronal populations that contribute to the processing reflected in each component (Ponton et al., 2000, 2002; Gomes et al., 2001; Tonnquist-Uhlen et al., 2003). Also, AEP/AEF development in children is different for speech and non-speech sound. Pang and Taylor (2000) reported that the N1 to speech sound matured between 13 and 14 years of age and the N1 to pure tones after 15–16 years of age. Wunderlich et al. (2006) described larger discrepancies in amplitude and latency in overall AEP components depending on stimulus types (speech, high- and low-frequency pure tones) in children from newborns to 3 year old. This implies that processing musical sounds might also mature differently from processing non-musical sound and that the maturational trajectory can be affected by intensive musical experience.

As effect of musical training on adult auditory cortical responses larger N1m and P3 (endogenous component peaking at 300 ms relate to target detection) have been found in musicians than non-musicians, especially when they commenced musical training at an earlier age (Pantev et al., 1998; Trainor et al., 1999). The timbre-specific enhancement of N1m indicated that this enhancement strongly depends on their training experience (Pantev et al., 2001). Enhanced P2 in musicians (Shahin et al., 2003) demonstrated training specific sensitivity of this AEF component. Endogenous evoked responses related to the automatic auditory processing have been shown to be enhanced in musicians when the stimuli involve processing of melodic, harmonic or temporal structure (Russeler et al., 2001; Tervaniemi et al., 2001; Koelsch et al., 2002; Fujioka et al., 2004, 2005; van Zuijen et al., 2004, 2005). These responses except P3 reflect pre-attentive processes for sound perception and do not require superior behavioural performance from musicians as prerequisite. These enhanced auditory responses in musicians are accompanied by their enlarged cortical areas such as medial part of Heschl’s gyrus (Schneider et al., 2002) and the anterior part of corpus callosum (Schlaug et al., 1995).

Because short-term training in auditory ‘non-musical’ discrimination tasks can also cause enhancement in adults’ AEP/AEF (Reinke et al., 2003; Tremblay et al., 2001; Tremblay and Kraus, 2002; Bosnyak et al., 2004), it is of interest to examine how the brain develops differently in young children undergoing musical training, and to determine which AEP components are sensitive to such a training experience at early age. These issues were addressed in a previous study (Shahin et al., 2004), where musically trained children of 4–5 years of age showed larger P1 and P2 to the piano tones compared with untrained children. However, the impact of training on developmental trajectory remained unclear because no significant difference was seen between the two measurements 1 year apart in either group.

Thus, the aim of the present study is to address three questions: (i) how do auditory responses in children mature within a year? (ii) What component of brain responses in each hemisphere matures differently to musical sound and non-musical sound? (iii) How does musical training affect normal maturation? Using a violin tone and a noise-burst stimulus, we recorded AEFs in four sessions over a year period from 4- to 6-year-old children. Half of the children...
were enrolled in a music programme called Suzuki methods, whereas the other half were not involved with any musical education outside of school.

Material and methods

Subjects

Twelve children (three females) of 4–6 years of age at the first MEG recording participated in this study. Four sessions were performed at intervals of 3–4 months so that the last session was conducted within 11–14 months after the first session (mean: 12.08). Six children (one female) had attended one of the Suzuki music schools in Toronto area within 3 months before the first measurement (five violin and one piano). The average duration of lessons before the first session was 1.5 months (range: 0.75–2.5). The other six had taken no music lessons outside of school. One participant in each subgroup was left-handed. One subject in each group missed the second session and another subject missed the third session. The total number of subjects and the average age at each session are given in Table 1.

Parents gave informed consent to participate after they were informed about the nature of the study. Parents also completed a detailed questionnaire concerning their child’s daily activities inside and outside of school, handedness and health status. None had reported learning or health problems. The time spent for outside-school activities ranged between 0.5 and 1.5 h daily in both groups. The variety of extra-curricular activities in the Suzuki children included gymnastics, hockey, karate, skate, soccer and swimming in addition to Suzuki study. In the untrained group the activities included dance, gymnastics, hockey, karate, soccer and swimming. The average time for listening to music was 4.4 h per week for both Suzuki and untrained children. Only one child in the Suzuki group had both parents who played musical instruments actively without formal training, and another had a parent who was trained in piano at an advanced level but was not playing currently. Four children in the untrained group had a parent who was trained to an advanced level, two of whom were currently active hobby musicians. None of the participants reported either having absolute pitch perception or playing/singing by ear. The Research Ethics Board at Baycrest approved all experimental procedures.

Musical education

The Suzuki method is a music pedagogy system developed by the Japanese violinist Shinichi Suzuki (1898–1998) in Japan in the 1940s. His association Talent Education has spread throughout Japan and worldwide from the 1960s. The first Canadian Suzuki programme was established in 1965 according to The Canadian Encyclopaedia (http://www.thecanadianencyclopedia.com).

We chose students in Suzuki training rather than other institutional or private music programmes for four reasons. First, the programme is provided by a certified teacher in an institution licensed by the association, which uses the Suzuki lesson books in order. Thus common exposure to musical materials across participants over time was ensured. It should also be noted that absolute pitch training is not a part of Suzuki method. Secondly, the programme strictly forbids selection of children according to their initial musical talent. This ensures that children enrolled in the programme are not selected because they are behaviourally gifted before the training. Thirdly, the parental involvement and social philosophy of this method allows us to assume that children in the programme have similar supports from family and peers regardless of their progress in the actual training. Finally, the absence of early training in reading musical notation in the programme allows us to examine the effects of training in auditory and sensorimotor modalities, providing a closer model to training-induced cortical change in behavioural neuroscience.

Stimuli

Two types of acoustic stimuli were used. One was a violin tone, produced by one down-bow stroke at the pitch of A4 (fundamental frequency of 440 Hz) with an open A string. The whole duration of the tone was 850 ms; the power of the acoustic wave decayed to 33% of the maximum at the 500 ms time point. The violin tone was played by a graduate student in the faculty of music, University of Toronto, and recorded in the faculty studio using a microphone (TLM 193, Georg Neumann, Berlin, Germany), mixer (GS3, Allen and Heath, Cornwall, UK), analogue–digital converter (Delta 1010, M-Audio, Irwindale, CA, USA) and a Pentium-4 processor computer system (Intel, Santa Clara, CA, USA). The other stimulus was a 500 ms noise burst, including 5 ms linear rising and falling ramps, created by a custom-built computer program. The simple time envelope for the noise rather than the same envelope as the violin sound was chosen to avoid a wind- or storm-like sound.

Both stimuli were stored and played in Windows wave file format using 16 bit and a 44 100 Hz sampling rate in STIM computer program (Neuroscan Inc., El Paso, TX, USA). Stimuli were presented in blocks of 100 repetitions at a stimulus–onset asynchrony (SOA) of 3 s, resulting in block duration of ~5 min. The order of stimulus blocks was counter-balanced within each group and each session. Each stimulus block was repeated if subjects agreed to extend the recordings.

The sound intensities were set in each ear to 60 dB above hearing threshold for the violin tone. Piloting test in 10 young adults with normal hearing revealed that thresholds for the noise were 3 dB higher on average than those for the violin tone. Therefore, the noise sound file was calibrated to be 3 dB above the violin tone. The hearing thresholds were assessed for each ear behaviourally as follows. First, the children were familiarized with the violin sound; then, they were accommodated properly into the MEG recording position and instructed to raise their hand whenever they heard the
subject’s compliance was verified by video monitoring. Most
movies), projected onto a screen placed in front of the face. The
ear was corrected in data analysis. The children were instructed
Grove Village, IL, USA). The 10 ms delay until the sound reaches
connected to the transducers (ER3A, Etymotic Research Inc., Elk
insert earphones, which consisted of foam insert tips
inside the MEG dewar (Fig. 1). The sound stimuli were delivered
lay down comfortably in supine position on a bed with their head
coordinate system was adjusted accordingly afterwards. Children
locations were recorded in the head shape data and MEG
subject, they were attached as close as possible to the correct points.
exact pre-auricular points owing to the extreme discomfort in the
on the scalp surface were recorded using a 3D digitizer (Polhemus,
MEG room. The coil positions and
auricular points, respectively, before subjects were brought to the
sensor were adhered at the nasion and to the left and right pre-
position on a bed with the head inside the whole-head dewar.
Fig. 1 Apparatus for MEG recording with a participant in a lying
sound, while the experimenter gradually increased its intensity in
5 dB steps from 0 dB using an audiometer (OB822, Madsen,
auditory response, resulting in the most representative
tangential sources from the auditory cortices. The tangential source
activities should characterize P1-N1-P2-N2 peaks in which most
resulting EEG activities should characterize P1-N1-P2-N2 peaks in which most
maturational change occurs (Ponton et al., 2002) (In their study N1
auditory areas, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
the vector normal to the x–y plane in the upward direction. Table 2 also shows the number of dipoles accepted
according to the following criteria: (i) goodness of fit was better
than 70%; (ii) source orientation upward ~100 ms; and (iii)
location was successfully estimated within the general auditory
eys were recorded in the head shape data and MEG
cordate system was adjusted accordingly afterwards. Children
lay down comfortably in supine position on a bed with their head
inside the MEG dewar (Fig. 1). The sound stimuli were delivered
through insert earphones, which consisted of foam insert tips
attached to plastic tubes (diameter: 5 mm, length: 1.5 m),
connected to the transducers (ER3A, Etymotic Research Inc., Elk
Grove Village, IL, USA). The 10 ms delay until the sound reaches
the ear was corrected in data analysis. The children were instructed
to stay awake and not to move their bodies. They watched a
soundless movie of their own choice (cartoons or child-friendly
movies), projected onto a screen placed in front of the face. The
subject’s compliance was verified by video monitoring. Most

Data analysis
Table 2 shows the numbers of successfully recorded blocks in each session. In most cases the recording of a repeated second block was feasible. AEF responses were averaged according to stimulus onset within each stimulus block after rejection of artefact (eye-blink, eye-movement and body movement, typically exceeding 1.0 pT) by visual inspection. The mean number of recorded trials per block was 97.2 out of maximum 100, while that of accepted trials was 85.5.

AF sources were modelled for each block as equivalent current dipoles in the left and right hemispheres using spatiotemporal fitting methods (Scherg and Von Cramon, 1986) within the time interval of 0–800 ms after stimulus onset. This fit algorithm gave an estimate of the centre of neural activity for the whole time-course of the auditory response, resulting in the most representative tangential sources from the auditory cortices. The tangential source activities should characterize P1-N1-P2-N2 peaks in which most maturational change occurs (Ponton et al., 2002) (In their study N1 is referred to as N1b to distinguish it from the radially oriented N1 components. Here we use N1 for simplicity as we do not measure radial sources). A single sphere model with the radius between 7.5 and 8.2 cm was estimated from the individual digitized head shapes, yielding almost similar sphere size to adult allowing dipole fitting without modification. The three fiducials defined the x–y plane of a Cartesian coordinate system with the origin at the midpoint between the pre-auricular points, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
defined by the vector normal to the x–y plane in the upward direction. Table 2 also shows the number of dipoles accepted
according to the following criteria: (i) goodness of fit was better
than 70%; (ii) source orientation upward ~100 ms; and (iii)
location was successfully estimated within the general auditory
areas, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
defined by the vector normal to the x–y plane in the upward direction. Table 2 also shows the number of dipoles accepted
according to the following criteria: (i) goodness of fit was better
than 70%; (ii) source orientation upward ~100 ms; and (iii)
location was successfully estimated within the general auditory
areas, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
defined by the vector normal to the x–y plane in the upward direction. Table 2 also shows the number of dipoles accepted
according to the following criteria: (i) goodness of fit was better
than 70%; (ii) source orientation upward ~100 ms; and (iii)
location was successfully estimated within the general auditory
areas, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
defined by the vector normal to the x–y plane in the upward direction. Table 2 also shows the number of dipoles accepted
according to the following criteria: (i) goodness of fit was better
than 70%; (ii) source orientation upward ~100 ms; and (iii)
location was successfully estimated within the general auditory
areas, the posterior–anterior axis (x) pointing toward the nasion and the medial–lateral axis (y)
pointing from right to left. The inferior–superior axis (z) was
defined by the vector normal to the x–y plane in the upward direction.
in the first session. Sensor topographies are shown for four different time points, corresponding to the positive peak at 100 ms (termed P100m), the negative peak at 250 ms (N250m), the positive peak at 300 ms (P320m) and the negative peak at 450 ms (N450m). The polarity of these components is positive for an upward oriented source that corresponds to a vertex positive potential in the EEG.

From source waveforms, peak amplitude and latency for P100m, N250m, P320m and N450m were identified in each individual data set. When a peak was not clearly identifiable, that point was treated as missing. The data of four participants who missed either the second or third session of recording were estimated by the average within each subject. The peak amplitude was measured from the baseline for all components.

The data were statistically examined by repeated-measures analyses of variance (ANOVA) using three within-subject factors: Sound (violin, noise), Hemisphere (left, right) and Session (1st, 2nd, 3rd, 4th). An additional between-subjects factor, Group (Suzuki, untrained), was examined. When Mauchly’s sphericity test was significant, the appropriate adjustment in degrees of freedom [Greenhouse–Geisser’s epsilon (ε)] was applied. Post hoc comparisons were conducted with Bonferroni corrections.

### Behavioural tests

Music and digit span tests were conducted at the first and fourth sessions to assess children’s musical discrimination and working memory abilities. The music test consisted of a sequence

### Table 2

<table>
<thead>
<tr>
<th>Session</th>
<th>Subjects</th>
<th>Violin</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recorded blocks</td>
<td>n dipole</td>
<td>GOF(%)</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>23</td>
<td>19</td>
</tr>
</tbody>
</table>

**Fig. 2** Mean dipole locations for each session (Sessions 1, 2, 3 and 4) of violin and noise stimuli are shown in two-dimensional planes for (top) the inferior–superior and the medial–lateral directions, and (bottom) the posterior–anterior and the medial–lateral directions. Error bar in each plot indicates standard error of the mean (SEM).
of same–different discrimination tasks in three subcategories of harmony, rhythm and melody, containing 10, 8 and 8 trials, respectively. The further details of each section are described in a previous study (Anvari et al., 2002). This test took ~30 min including instructions. Digit span used in the present study is a subtest from the Wechsler Intelligence Scale for Children—Third Edition (WISC-III) (Wechsler, 1991) and took ~10 min to conduct. Since most subjects were under 6 years old at the first session, normative data are not available for WISC-III. Also, the normative scores of the music test for different age groups are not available. Thus, for both tests the raw scores were submitted to statistical analysis. The mean difference between groups at each of two of the sessions as well as between sessions in each group were examined by the 95% confidence interval (CI) computed by a bootstrap resampling (iteration = 10 000) to examine the significance level of 5% because of the small number of the samples.

Results
General maturation of auditory evoked responses
AEF waveforms
Clear AEs and its source waveforms were obtained from all subjects for each stimulus in each session. Figure 4A shows the data from one subject in all four sessions. The P100m, N250m, P320m and N450m peaks and the following SF were seen consistently from most of the subjects for both stimuli. Half of the subjects exhibited an additional negative–positive component between the P100m and the N250m (named N150m and P200m). Each subgroup had one subject with bilateral N150m peak, one of which is shown in Fig. 4B. Four other subjects exhibited N150m only in the right hemisphere. Figures 3 and 4 demonstrate that the responses to the two stimuli were different within individuals and changes across sessions varied depending on the AEF component. Also, responses from the left and right hemisphere differed in terms of peak amplitudes and latencies.

Peak amplitude
Peak amplitude values of the P100m, N250m, P320m and N450m components were examined with repeated-measures ANOVA using three within-subject factors: Sound (violin, noise), Hemisphere (left, right) and Session (first, second, third, fourth). Figure 5A–D shows the mean values and standard errors of the mean for all components.

For P100m amplitude (Fig. 5A), the main effect of Sound was significant \( F(1,11) = 10.27, P = 0.008 \), reflecting a larger response for violin (mean: 24.3 nAm) than for noise (19.8 nAm). As well, Hemisphere was significant \( F(1,11) = 14.42, P = 0.003 \) owing to the larger response in the left (26.9 nAm) than in the right hemisphere (17.4 nAm). Session was not significant and there were no significant interactions.

For N250m amplitude (Fig. 5B) Session was significant \( F(3,30) = 7.89, P = 0.0005 \) owing to a decrease in peak negativity over time. Post hoc tests showed significant differences between sessions first and fourth \( (P = 0.0002) \), second and fourth \( (P = 0.0003) \) and third and fourth \( (P = 0.003) \), where significance was accepted for \( \alpha = 0.0083 \) according to the Bonferroni correction. The Sound \( \times \) Session interaction was not significant.
interaction was significant \(F(3,30) = 4.57, P = 0.009\), reflecting a greater decrement over time in the N250m for violin than for noise.

For P320m amplitude (Fig. 5C), the main effect of Hemisphere was significant \(F(1,9) = 5.47, P = 0.04\) with a more positive peak in the left (−6.3 nAm) than in the right hemisphere (−12.1 nAm). It may be noted that the P320m peak is below the baseline (i.e. the amplitude value is less than zero), although the peak polarity is positive. Session was also significant \(F(3,27) = 5.99, P = 0.003\) owing to a larger amplitude of the fourth session than those of the other sessions. Post hoc comparisons revealed significant differences only between the third and fourth sessions (\(P = 0.0003, \alpha = 0.0083\)).

For N450m (Fig. 5D), Sound was the only significant effect \(F(3,27) = 5.99, P = 0.003\), reflecting larger negativity for violin (−36.9 nAm) than for noise (−25.8 nAm).

In summary, violin tones produced larger P100m and N450m components than did noise bursts. N250m and P320m changed across sessions, indicating rapid developmental changes over the longer latency components in contrast to the more stable P100m component. Only N250m showed differential change across sessions for violin tones and noise bursts. P100m and P320m showed hemisphere differences, with consistently larger peaks on the left than on the right. No interactions between Hemisphere and the other factors were found for any of the components, indicating that hemispheric differences in amplitude were stable over time.

**Peak latency**

Peak latencies of the P100m, N250m, P320m and N450m components, shown in Fig. 5E–H, were examined with the same ANOVA as applied to the peak amplitudes.

A main effect of Sound for P100m latency (Fig. 5E) \(F(1,11) = 7.37, P = 0.020\) reflects longer latency for violin (mean: 110.6 ms) than noise (105.5 ms). Hemisphere was also significant \(F(1,11) = 19.19, P = 0.001\), with a longer latency in the left (112.5 ms) than in the right hemisphere (103.6 ms).

For N250m (Fig. 5F), there was a marginal effect of Session \(F(3,30) = 2.69, P = 0.0635\), reflecting a decrease in latency over sessions in general, although post hoc comparisons did not reach significance. The factors Sound × Hemisphere × Session showed a tendency for interaction \(F(3,30) = 3.71, P = 0.06\). This seems to be due to the fact that the latency for the violin tone decreased more...
rapidly in the right than in the left hemisphere, whereas
the latency for the noise burst decrease in the left hemisphere
was not linear. However, post hoc tests were not significant.

The P320m component (Fig. 5G) was significantly
affected by Sound $[F(1,9) = 27.21, P = 0.0006]$ with a
much shorter latency for violin (308.6 ms) than for noise
(339.0 ms). Session was also significant $[F(3,27) = 13.77,
P < 0.0001]$, revealing a considerable decrease of latency
over time. There were significant differences between the
first and third sessions ($P = 0.001$) and the first and fourth
sessions ($P < 0.0001$). In addition, there were significant
interactions Sound $\times$ Hemisphere $[F(1,9) = 5.58, P = 0.042]$ and Sound $\times$ Session $[F(3,27) = 7.18, P = 0.001]$. The former
was caused by the fact that the latency in the right
hemisphere was 11.5 ms shorter on average than in the left
hemisphere for violin tone stimulation ($P = 0.10$, n.s.),
but the latency for the noise burst in the left hemisphere
was shorter than in the right by 15 ms ($P = 0.17$, n.s.).

Fig. 5 (A–H) Mean and SEM of amplitude of P100m, N250m, P320m and N450m components to violin and noise stimuli from left and right hemisphere in each session.
The latter interaction was due to the rapid decrease of latency in the right hemisphere over session for violin but not for noise ($P < 0.0001$).

For N450m latency (Fig. 5H), Sound was significant \[ F(1,10) = 6.99, P = 0.025 \] with the latency for violin 35 ms shorter than for noise. This difference across stimuli was similar to that observed for the P320m component, suggesting that there might be no additional delay specific to Sound on the N450m component. Session was also significant \[ F(3,30) = 4.82, P = 0.03 \], with a general decrement in latency across session. A significant difference was found between the first and fourth sessions ($P = 0.0008$). The interaction Hemisphere $\times$ Session approached significance \[ F(3,30) = 3.74, P = 0.07 \], reflecting a greater decline of latency in the right compared with the left hemisphere across session (n.s., $P = 0.12$).

To summarize, the noise burst evoked an earlier P100m component than the violin tone but the violin tone evoked earlier P320m and N450m components. For the N250m component, different stimulus-specific effects were seen in the right and left hemisphere. In line with the amplitude data, all components except the P100m exhibited latency decreases across session. The left P100m exhibited longer latencies than the right regardless of stimulus. Other components showed interactions between Hemisphere and Sound or Session, revealing complex effects of stimulus-specificity and hemispheric asymmetry with developmental change.

**Musical training effects**

**Behavioural tasks**

The individual scores and group data of the music test are illustrated in Fig. 6A. At the first session, the mean score was 13.7 points in the Suzuki group and 14.5 points in the untrained group, but this initial difference was not significant as the CI ranged from $-2.9$ to $1.5$ including zero point. The scores at the last session were 19.2 and 16.3 in Suzuki and untrained group, respectively, which was not significant ($-0.5 < CI < 5.8$). However, Suzuki group improvement by 5.5 points after 1 year was significant ($2.6 < CI < 8.0$), while the improvement of 1.8 points in the other group was not significant ($0.7 < CI < 4.6$).

As for digit span shown in Fig. 6B, two groups were not significantly different either in the first session (Suzuki: 8.5 points, untrained: 8.7 points, $3.5 < CI < 3.3$) or in the last session (Suzuki: 12.2 points, untrained: 10.0 points; $-0.3 < CI < 5.4$). The improvement in Suzuki children by 3.7 points, however, was significant ($1.5 < CI < 5.8$), although this was not the case for the increase of 1.8 points in the untrained children ($-2.6 < CI < 5.2$).

**Morphological change in AEF waveforms**

Figure 7 shows grand averaged source waveforms in each group, separately for the left and right hemisphere, session and sound type. In order to examine when in the waveforms significant change across sessions occurred, the standard deviation across the four sessions at each time point in the grand averaged source waveforms was calculated (Fig. 7, below each auditory evoked response). The pronounced peak at 100–400 ms in the time course of the standard deviation covers the onset of N250m and offset of P320m.

Furthermore, group difference became obvious in the time-courses of variance over 1 year of time. In the Suzuki-trained children, the change was clearly present only in the response to the violin tone, and had left-hemispheric dominance, while there was little change for the noise burst. On the contrary, in the untrained group, the change was similar for both stimuli and hemispheres.

**Laterality and stimulus specificity on N250m peak**

The group difference described above was assessed further on the N250m by another repeated-measures ANOVA with additional one between-subjects factor Group (Suzuki, untrained). The interaction Sound $\times$ Group $\times$ Session was close to significant both in amplitude ($P = 0.09$) and in latency ($P = 0.08$). Consequently, further analysis was conducted for the different sounds separately.

For the violin tone, a significant Group $\times$ Hemisphere interaction was found in amplitude \[ F(1,9) = 10.48, P = \]
reflecting more pronounced negativity in the left hemisphere in the Suzuki than in the untrained group, but no difference across groups in the right hemisphere (Fig. 8A).

The main effect of Session was significant \[ F(3, 27) = 12.86, \ P < 0.0001 \], and the interaction between Group \( \times \) Hemisphere \( \times \) Session was close to significant \( \( P = 0.09 \), reflecting a larger change of amplitude in the left than right hemisphere in the Suzuki group but no difference in the untrained group. For latency, the main effect of Group was significant \[ F(1, 9) = 8.21, \ P = 0.019 \] owing to a shorter latency in the Suzuki group (mean: 217.2 ms) of >30 ms compared with the untrained group (250.8 ms) as seen in Fig. 8B. As well, the interaction Session \( \times \) Group approached significance \( P = 0.08 \) because a large decrease of latency occurred only in the Suzuki group between the first and second session.

In contrast to the results for the violin tone, the noise burst produced no significant effects or interactions involving group. There was a significant overall effect of Session on the amplitude of the N250m, similar to that seen for the violin tone \[ F(3, 27) = 3.60, \ P = 0.03 \]. This indicates that the decrease sessions in the amplitude of the N250m for both stimulus types represents a general decrease in the amplitude of this component with increasing age, while the main effects and interactions involving Group and Hemisphere for the violin tone most likely reflect the effects of musical experience on the developing brain.

Fig. 7 Grand averaged source waveforms at each session with its standard deviation across all sessions in each hemisphere in response to violin and noise stimuli from (left) Suzuki-music-trained subjects and (right) non-musically trained subjects.
Discussion

The present study revealed morphological changes in the late components of the AEF over the period of 1 year in 4- to 6-year-old children, particularly in the 100–400 ms time window. In addition, larger P100m and N450m amplitude as well as more rapid changes of the N250m amplitude and latency were associated with the violin sound rather than the noise stimuli. The positive P100m and P320m components were larger over the left than the right hemisphere. Of most importance, the effects of Suzuki musical training over a 1-year period were seen in response to musical sound but not noise and were accompanied by enhanced behavioural performances in musical discrimination and non-musical working memory.

Development of AEF

The shorter peak latencies across sessions were observed for the N250m, P320m and N450m, and amplitudes decreased in the N250m and P320m during the time-course of 1 year. Decrease in P100 latency and amplitude with age have been reported in developmental studies (Molfese et al., 1975; Fuchigami et al., 1993; Tonqust-Uhlen et al., 1995; Oades et al., 1997; Sharma et al., 1997; Rojas et al., 1998; Cunningham et al., 2000; Ponton et al., 2000). We did not observe a significant P100 latency change within our 1-year window, which is consistent with other studies examining this age group (Korpilahti and Lang, 1994; Cunningham et al., 2000). In particular, Cunningham et al. (2000) reported that the P100 latency was relatively stable from 5 to 10 years but decreased abruptly after 12 years. Data from Ponton et al. (2002) also agree with this by showing a slower pace of P1 latency change between ages 5 and 11 than in the later ages after 12.

We interpret N150m and N250m in our data as corresponding to adult N1 and N2, respectively. For the latter, we are only concerned with those elicited from ‘non-oddball’ paradigm in children (Paetau et al., 1995; Ponton et al., 2000, 2002; Takeshita et al., 2002) and in adults (Hari et al., 1982; Mäkelä et al., 1993; Paetau et al., 1995; Picton et al., 1974). The distinction of these two components in maturational trajectory has been clearly indicated by a cross-age comparison by Ponton et al. (2002), although some earlier studies suggested child N250m as corresponding to adult N1. Developmental data show that N150m emerges in older children and increases in amplitude and decreases in latency with age while N250m amplitude decreases with age and is very weakly present in adults (Courchesne, 1990; Sharma et al., 1997; Cunningham et al., 2000; Pang and Taylor, 2000; Ponton et al., 2000, 2002). Our results are consistent with these trends in that the N150m was only visible in about half of the subjects and N250m decreased in amplitude over the year. In addition, unilateral N150m presence in the right hemisphere was found in one-third of our subjects and this appears to be concordant with the observation by Paetau et al. (1995), in which the right hemisphere exhibited much shorter N150m latency than the left in children of 6–8 years old. While the decrease of N250m latency observed here is in line with the other studies, Ponton et al. (2002) reported the increase. This seems likely owing to their rapid stimulation using clicks since a faster stimulation is reported to evoke larger N250(m) (Karhu et al., 1997; Takeshita et al., 2002), which might affect the degree of overlapping components.

The P320m and N450m responses have been described in only a few studies of children between 3 and 9 years of age (Paetau et al., 1995; Čeponienè et al., 1998) and in newborns (Kushnerenko et al., 2002); thus comparison throughout development is not available. Important findings in our data were that P320m and N450m decreased in latency at a similar rate while only the P320m exhibited an increase in amplitude with no change found in the N450m, and that the violin stimulus was more important to the P320m latency change than the noise stimulus. Thus, our data show large changes in these components over a single year, suggesting...
that it is important for future studies to further examine how these components change with development and specific stimulus conditions.

In general, latency decrements are thought to reflect an increase in myelination and synaptic efficiency (Eggermont, 1992; Moore and Guan, 2001), while decreases in amplitude are thought to reflect a decline in the number of pyramidal cell synapses contributing to post-synaptic potentials that occurs with synaptic pruning (Huttenlocher and Dabholkar, 1997) as more specialized and effective cortical networks develop. Our data indicate that these processes can be seen in AEF changes over the course of a single year. In addition, it is worth noting that these developmental changes are clearly seen in the source waveforms based on modelling the MEG signals with equivalent current dipoles. Successful estimation of bilateral sources within auditory areas with dipole modelling of paediatric data has been shown in AEP (Albrecht et al., 2000; Ponton et al., 2002) and AEF studies (Paetau et al., 1995; Rojas et al., 1998; Takeshita et al., 2002; Heim et al., 2003). Overall the data demonstrate that MEG recordings and source analysis are good methods for longitudinal studies of children’s brain function.

**Stimulus effects on AEF**

The violin sound evoked larger P100m and N450m than did the noise across our subjects. Similar stimulus effects were reported in previous studies. Speech stimuli evoked larger P100 and N250 compared with pure tones in newborns and young children of 1–6 years of age (Wunderlich et al., 2006) and in children at 8–10 years of age (Čeponienė et al., 2001). The N450 was also larger for syllables than for non-phonetic stimuli (Čeponienė et al., 2005). These studies including ours carefully calibrated the sound intensity to exclude possible confounding intensity effects. The observed stimulus effect appears to correspond to adult N1m (Diesch and Luce, 1997; Tütthin et al., 1999) and P2 (Shahin et al., 2005), which are enlarged according to acoustic complexity. Since N1m and P2m are at least partly generated from tonotopically organized areas (Pantev et al., 1995; Lutkenhöner and Steinsträter, 1998), it seems reasonable to assume that spectral complexity is an important factor on how many neurons with different characteristic frequencies respond. Although the functional relationship between adult’s N1/P2 and children’s components is not clear, it seems reasonable that neural activities in auditory cortex varies on the basis of required acoustical processing and/or subsequent perceptual processing.

The more rapid developmental changes in N250m related to the violin sound is consistent with a study in which faster maturation of the developing brain for right-handed subjects (10 out of 12 of our subjects were right-handed). The planum temporale (PT) has a greater volume than the right PT from the earliest age studied (3 years) (Preis et al., 1999). Even the congenitally deaf population exhibit a larger volume for the auditory cortex in the left hemisphere than in the right as revealed by morphometric MRI analysis (Esmorey et al., 2003). This study further showed that the volume of grey matter is similar in both hearing and deaf subjects while the white matter volume is smaller in the deaf group. This suggests that listening experience affects the myelination and the number of fibres projecting in and out of the auditory cortices. Another study found that the frontotemporal pathway develops predominantly in the left hemisphere.
Cortical maturation and musical training

Musical training effects on AEF and cognitive functions

Perhaps the most interesting finding in the present study is that in children with Suzuki music training, the morphological change at 100–400 ms was only present in response to the violin tone but not for the noise. Suzuki group also showed a larger N250m to violin tones in the left hemisphere compared with the untrained group. This AEF difference was accompanied by significant improvements in musical discrimination and in the digit span task.

The stimulus-specific enhancement of AEF in Suzuki students is in line with reports from adult musicians, although the effects in adults were, again, found for earlier components such as N1m (Pantev et al., 1998; Pantev et al., 2001) and P2 (Shahin et al., 2003). Although functional difference between N150m and N250m is suggested from the sensitivity to stimulation rate, the origins of both components are located close together in the supratemporal plane (Takeshita et al., 2002). It is possible that the stimulus-sensitive neurons, which initially contribute to N250m contribute to N150m as age increases.

As discussed above, functional relevance of adult’s obligatory N2 to stimulus category, hemisphere and experience remains unclear owing to the sparseness of the relevant literature. Still it is possible that adult N2b, a subcomponent of N2, which is strongly enhanced by attention, can be related to N250 in children to some extent. Using oddball paradigm, an enlarged N250 in children has been reported for rare targets (Johnstone et al., 1996; Oades et al., 1997). Thus, the violin sound might have captured the involuntary attention of children who have been taking music lessons, resulting in an enhanced N250m.

Alternatively, it has been suggested that auditory-related abilities correlate with the N250 response in children. Tononquist-Uhle´n (1996) found that the obligatory N250 component in children with language impairment showed a longer latency and smaller amplitude than that of normal children. In addition, the N250 latency was found to correlate significantly with behavioural performance in a fine-graded phonological discrimination task in children between 5 and 15 years of age (Cunningham et al., 2000). Oddball paradigms containing rare targets of moving sound also evoked enhanced N250 responses in ‘good listeners’ compared with that in ‘poor listeners’ where ‘good’ and ‘poor’ were defined by performance in various auditory tasks before the AEP recording (Wambacq et al., 2005). Therefore, we speculate that the N250m may index some part of stimulus categorization, in combination with attentional modulation, and that taking music lessons may affect this component selectively.

Finally, these results contribute to the evidence that musical training affects AEF development in children. Although the groups differed at the first measurement on the left N250m (their early musical environment probably differed), they changed differently over the course of the year, suggesting the effect of musical training. Specifically, the N250m response to the violin tone changed more rapidly in the Suzuki group but not in the untrained group. Also, the latency of this component decreased over sessions only in the Suzuki group. It is of course possible that these effects are due to some other factors such as home environment; however, both groups engaged in similar amounts and varieties of extra-curricular activities. Furthermore, the differences between groups were specific to the violin tone; the groups did not differ in response to the noise stimulus. Thus the increasing differences between groups over time are probably related to the musical activities.

This is the first study to show that brain responses in musically trained and untrained children change differently over the course of year. Previously, enhanced P100, N150 and P200 responses to piano tones were found in children who were taking piano lessons compared with untrained children (Shahin et al., 2004). However, their data did not show a difference before and after training, and they did not examine later components such as N250. In fact, Fig. 3 in Shahin et al. (2004) appears to show an enhanced N250 similar to ours in the trained group, but the data they collected did not allow examination of changes over time. In addition, the responses in their data evoked by the violin tones exhibited very small or non-existent N150 and P200 components consistent with our data. This is probably due to the slow onset of the violin tones compared with the piano tones (Elfner et al., 1976). Although a number of
studies have shown stimulus-specific enhancement in auditory evoked responses between adult musicians and non-musicians (Pantev et al., 1998, 2001; Shahin et al., 2003), our study demonstrates the early onset of these differences during the first year of musical training.

While improvement in musical tasks is not surprising after 1 year of Suzuki music lessons, enhanced performance in digit span gives additional evidence for transfer effect between music and non-musical abilities such as literacy (Anvari et al., 2002), verbal memory (Ho et al., 2003), visuospatial processing (Costa-Giomi, 1999), mathematics (Cheek and Smith, 1999) and IQ (Schellenberg, 2004). In particular, Schellenberg (2004) randomly assigned children to either music lessons or drama lessons. This design eliminates differences between groups that could pre-exist between children whose parents enrol them in music classes compared with those who do not. It should be noted in this regard that although we did not have random assignment, and the groups differed at the first measurement on one of the AEF components, they changed differently over the course of the year, which strongly suggests the effects of musical training. Interestingly, with respect to brain structure measured by morphometric MRI analysis, children (5–7 years) planning to take musical lessons did not differ from those with no such plans (Norton et al., 2005). While it appears that music practice might have impacts on ‘general intelligence’, the underlying neural mechanism remains unclear. Our finding of enhanced performance after a year of musical training on the digit span task suggests that this experience affects working memory capacity, at least to some extent. Alternatively, it might be also related to the more advanced perseverence skills and/or the ability to sustain focused attention observed in Suzuki children compared with children enrolled in other activity such as creative movement classes (Scott, 1992). This interpretation is also consistent with the enhanced changes on N250m component in the Suzuki group, which could be related to attentional modulation.

**Conclusion**

Maturational changes in AEFs can be observed within a year in 4- to 6-year-old children using individual waveforms of cortical source activity. Source space analysis also reveals different stimulus and hemispheric effects for different components. Furthermore, musical training resulted in specific developmental changes in the responses to musical sounds but not to responses to noise stimuli, probably reflecting the development of neuronal networks specialized for important sounds experienced in the environment.

**Acknowledgements**

This research is supported by International Music Foundation awarded to L.J.T. and C.P.; Canadian Institutes for Health Research to T.F. and B.R.; and The Sound Technology Promotion Foundation to T.F. and R.K. The authors thank Dr Terence W. Picton for advice and discussion, Judy Vendramini and Haydeh Shaghati for their assistance in data acquisition and Patricia Van Roon, Tim Bardouille, Andreas Wollbrink and Uril Yakubov for their technical support.

**References**


Bosnyak DJ, Eaton RA, Roberts LE. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. Cereb Cortex 2004; 14: 1088–99.


Engineer ND, Percaccio CR, Pandya PK, Moucha R, Rathbun DL, Kilgard MP. Environmental enrichment improves response strength, threshold,
Kushnerenko E, Cˇ eponien ´ _ 23: 5545–52.


