Music Training for the Development of Speech Segmentation

Clément François[1], Julie Chobert[2], Mireille Besson[2,3] and Daniele Schön[1]

[1]Institut de neurosciences des systèmes, INSERM and Aix-Marseille University, Marseille, France [2]Laboratoire de Neurosciences Cognitives, CNRS and Aix-Marseille University, Marseille, France and [3]Cuban Neuroscience Center, La Habana, Cuba

Address correspondence to Clément François. Email: clement.francois@incm.cnrs-mrs.fr

The role of music training in fostering brain plasticity and developing high cognitive skills, notably linguistic abilities, is of great interest from both a scientific and a societal perspective. Here, we report results of a longitudinal study over 2 years using both behavioral and electrophysiological measures and a test-training-retest procedure to examine the influence of music training on speech segmentation in 8-year-old children. Children were pseudo-randomly assigned to either music or painting training and were tested on their ability to extract meaningless words from a continuous flow of nonsense syllables. While no between-group differences were found before training, both behavioral and electrophysiological measures showed improved speech segmentation skills across testing sessions for the music group only. These results show that music training directly causes facilitation in speech segmentation, thereby pointing to the importance of music for speech perception and more generally for children's language development. Finally these results have strong implications for promoting the development of music-based remediation strategies for children with language-based learning impairments.

Introduction

Music engages a wide range of processing mechanisms, from sound encoding to higher cognitive functions such as sequencing, attention, memory, and learning. These functions, which are shared with several other human abilities (e.g., language), might in turn be shaped by music training. Thus, musicians are a privileged population for studying brain plasticity as well as for investigating the intriguing possibility that musical expertise transfers to other domains such as language. It is now well-established that music training induces functional and structural changes in the auditory and sensori-motor systems, making musicians more efficient and more sensitive in music-related tasks than nonmusicians. For instance, brainstem and primary auditory cortex responses to synthetic or instrumental sounds show more robust pitch encoding for musicians than for nonmusicians (Shahin et al. 2003, 2005; Musacchia et al. 2007; Wong et al. 2007). Also, musicians discriminate deviant chords or detect omitted sounds better than nonmusicians (Koelsch et al. 1999; Rüsseler et al. 2001; Brattico et al. 2009). These functional differences can be accompanied by morphological differences in terms of grey matter volume and density in the auditory cortex (Schlaug et al. 1995; Keenan et al. 2001; Bermudez and Zatorre 2005). Moreover, there is growing evidence that music training benefits linguistic skills such as dynamic acoustic analysis, pitch and lexical stress processing, phonological awareness, reading, and second-language proficiency (e.g., Tallal and Gaab 2006; Kraus and Chandrasekaran 2010). Some of these

findings have also been extended to children. Musically trained children better detect pitch changes in speech (Magne et al. 2006; Moreno et al. 2009; Kraus and Chandrasekaran 2010) and show increased verbal and reading abilities than children who did not receive music training (Moreno et al. 2009, 2011), thereby providing evidence for music to language transfer effects (Besson et al. 2011).

Turning to speech segmentation, the ability to extract words from continuous speech, there is evidence that infants, children, and adults can use the statistical properties of auditory input to discover words and sound patterns (Saffran et al. 1996, 1999; Aslin et al. 1998; Kuhl 2004; Gervain et al. 2008; Teinonen et al. 2009). In speech, the conditional probability of syllable Y happening given syllable X will be higher for syllables that follow one another within a word than for those at word boundaries (e.g., in "pretty music", the probability of "ty" given "pre" is higher than that of "mu" given "ty"). Thus, the statistical structure of a language seems to greatly contribute to speech segmentation. Interestingly, we recently found that musical expertise facilitated speech segmentation of an artificial language in adults (François and Schön 2011). Participants were familiarized with a stream of 5 artificial trisyllabic sung pseudo-words and then presented with 2-alternative forced choice tests with trisyllabic spoken pseudo-words or 3 tones melodies played on a piano timbre. On each trial, participants had to choose which of 2 sequentially presented items sounded more familiar. Musicians were more accurate than nonmusicians in both the musical and linguistic tests, although this difference remained a trend ($P = 0.11$). However, analyses of the event-related potentials (ERPs) to both musical and linguistic items revealed that a fronto-central negative component was significantly larger for unfamiliar than for familiar items for musicians only. This result was taken as evidence that musical practice facilitated stream segmentation.

However, cross-sectional studies comparing musicians and nonmusicians demonstrate correlations but not causality (Schellenberg 2004). Here, we used the longitudinal approach to test for causality. We conducted a longitudinal study spanning over 2 school years and we followed the developmental dynamics of music to speech transfer effects. We controlled for any pre-existing predispositions for music by using a test-training-retest procedure with 8-year-old children pseudo-randomly assigned to 1 of 2 training groups (music or painting) without self-selection. In the first "test session", 24 8-year old children listened to 5 min of an artificial sung language (Schön et al. 2008) built by random concatenation of 4 trisyllabic meaningless pseudo-words. Syllables were always sung using a fixed syllable-pitch mapping (Fig. 1A). After this familiarization phase, children were presented with 2 spoken items and had to decide which item sounded more familiar (32 trials). Importantly, all items in the test were
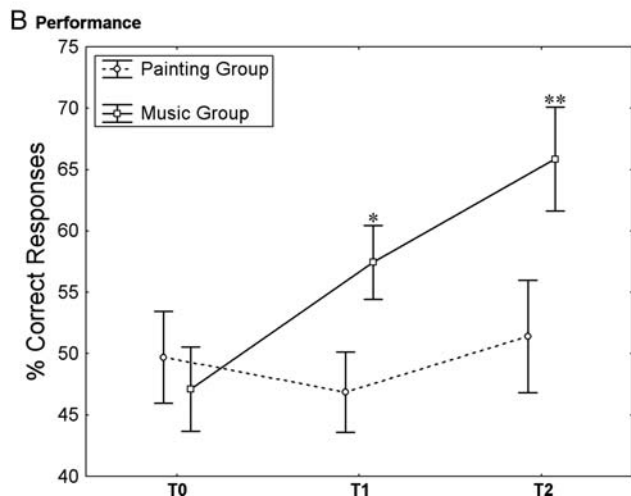
Figure 1. Effects of music and painting training on speech segmentation. (*A*) Experimental design used in the 3 test sessions. (*B*) Percentage of correct responses in the 3 test sessions (T0, T1, and T2) for the music (solid line) and painting (dashed line) groups. Error bars indicate the standard error of the mean. Asterisks show significant differences between groups (*$P = 0.05$; **$P < 0.01$).

spoken and not sung. Both EEG and behavioral responses were recorded during the task. Children were then pseudo-randomly assigned to 2 training groups (controlling for age, school level, sex, socio-economic background, and musical expertise and for the level of performance in several neuropsychological tests assessing reasoning, memory, and attentional processing). One group of children took music and the other painting classes for 45 min, twice a week in year 1 and once a week in year 2. "Test sessions" 2 and 3 (T1 and T2) were identical to "test session" 1 (T0) and took place approximately after 1 and 2 years. We hypothesized that children in the music group would improve their speech segmentation abilities across the test sessions more than children in the painting group. Based on our previous findings with adults (François and Schön 2011) and considering that we simplified the stimuli for children (4 pseudo-words rather than 5), we expected a behavioral facilitation together with larger ERP differences between familiar and unfamiliar items over frontal regions in the musically trained compared with the painting-trained children.

Materials and Methods

Participants

A total of 37 8-year-old nonmusician children were enrolled in these experiments. Thirteen children were excluded from final analysis either because they moved away during either the first (5) or the second year (3) or due to inattentive behavior and impulsiveness (5), thus leading to a final group of 24 children (mean age = 8, standard deviation = 0.45, 19 right-handed, 14 boys, normal hearing, no known neurological problems). None of the children had taken part to such an experiment before this project. Moreover, none of them took

music or painting lessons privately either before or during the project. All children were French native speakers. Parental informed consent was obtained for each child and the data were analyzed anonymously. This study was approved by the CNRS and was conducted in accordance with national norms and guidelines for the protection of human subjects.

Longitudinal Study: Procedure

Children were tested before training (at T0), after approximately 1 year (at T1), and 2 years (at T2). At T0, T1, and T2, children were tested individually in a quiet room of their school in 2 separate sessions that included neuropsychological assessments and electrophysiological tests, respectively. Each session lasted for 2 h and was separated by 4 or 5 days. Results were used for the pseudo-random assignment of children to the music or painting groups and as a baseline at T0 to evaluate the impact of the training programs at T1 and T2. Pseudo-random assignment to music or to painting training group was based on results of several neuropsychological tests issued from the WISC-IV and NEPSY batteries (Verbal Comprehension, Perceptual Reasoning, Working Memory, and Attention) as well as on age, school level, sex, and socio-economic background. This was done to ensure that no significant differences existed between the 2 groups before training. Children had similar socio-economic backgrounds ranging from middle to low social class according to the criteria of the National Institute of Statistics and Economic Studies. None of the children and none of their parents had formal training in music or painting. Moreover, children enrolled in the music group did not have their instruments at home to prevent any additional practice outside the music classes.

Two teachers professionally trained in music or painting were specifically hired for this project from October to May in year 1 and in year 2. Music training was based on a combination of Kodaly and Orff approaches (http://www.iks.hu/; http://www.orff.de/en.html). Painting training was based on the approach developed by Arno Stern (http://www.arnostern.com/). The teaching activity was coordinated by the research group and care was taken to ascertain that both groups were similarly motivated and stimulated.

Speech Segmentation Experiment: Design and Procedure

During the familiarization phase, children were asked to listen carefully to a continuous stream of sounds. During the following test, children had to choose, by pressing 1 of 2 response buttons, which of 2 items (first or second) most closely resembled what they just heard in the stream. In the test, items were spoken (i.e., flat contour). In each test trial, one item was a pseudo-word from the artificial language (i.e., gimysy, pogysi, pymiso, sipygy) while the other was built by merging the last syllable of one pseudo-word with the first 2 syllables of another (e.g., Sisipy and Sypymi) or the last 2 syllables of one pseudo-word with the first syllable of another (e.g., Gysigi and Pygygi). Pseudo-words and partial pseudo-words did not have any meaning in French. The mean transitional probabilities (TP) were 0.8 for pseudo-words (ranging from 0.6 to 1) and 0.4 for partial pseudo-words (ranging from 0.32 to 0.6). Each pseudo-word was presented with each partial pseudo-word, making up 16 pairs and repeated twice in a quasi-random order (32 trials). Stimuli were presented via headphones. Learning phase and test lasted 5 min each.

Material

The artificial language was built using 9 syllables combined to give rise to 4 trisyllabic pseudo-words (gimysy, pogysi, pymiso, sipygy: i.e., non-lexical vocables respecting the phonotactic constraints of French). Each of the 9 syllables was associated with a distinct tone. Therefore, each pseudo-word had a unique melodic contour (gimysy C3 D3 F3, pymiso B3 E4 F4, pogysi D4 C4 G3, sipygy G3 B3 C4; Fig. 1). The language stream was built by a random concatenation of the 4 pseudo-words (without repetition of the same item twice in a row) and synthesized using Mbrola (http://tcts.fpms.ac.be/synthesis/mbrola.html). Each pseudo-word was repeated 100 times in the stream to give rise to a 5 min stream.

Data Acquisition

As children were not given speed instructions, analyses were conducted on the percentage of correct responses. EEG was recorded before training (at T0) and after 2 years (at T2) from 32 scalp electrodes located at standard positions (International 10/20 system sites) during the behavioral task. The data were then re-referenced offline to the algebraic average of the mastoids. Trials containing artifacts were excluded. Artifacts were first detected by eye-balling and then by using a 75 μV maximum amplitude criterion (less than 10% of the trials). Two extra participants were discarded from ERP analyses only due to major EEG artifacts. The EEG was amplified by Biosemi amplifiers with a band-pass of 0–102.4 Hz and was digitized at 512 Hz.

Data Analyses

Neuropsychological data were analyzed using repeated-measures multivariate ANOVAs with group (music vs. painting) as a between-subject factor, "Test Session" (T0 vs. T1 vs. T2) as a within-subject factor, and the score at the test as the dependent variable. Behavioral data in the 2-alternative force choice test were analyzed using repeated-measures ANOVAs (RM-ANOVAs) to compare the percentage of correct responses across groups and testing sessions. The Tukey tests were used for post hoc comparisons. Average performance was also compared with chance level using 2-tailed one-sample t-tests. Finally, further analyses of the behavioral data modeled the effect of the items. This was done using a 2×4 RM-ANOVA including group (music and painting) a between-subject factor and Items (4 words) as within-subject factors, respectively, after the first and the second period of training.

ERP Data Analyses

ERP data for familiar items (averaged across the 4 familiar items and across children) and for unfamiliar items (averaged across the 4 unfamiliar items and across children) were analyzed by computing the mean amplitudes in successive non-overlapping 50 ms windows from 0 to 1000 ms post-stimulus onset. RM-ANOVA was used for statistical assessment that included group (music vs. painting) and Familiarity (familiar vs. unfamiliar items). Moreover, to test for the distribution of the effects, the model included the anterior–posterior (frontal, central, and parietal) and hemisphere factors (left and right). P-values were adjusted using the Greenhouse–Geisser correction. Because of the increased likelihood of type I errors, only effects that reached significance ($P < 0.05$) in at least 2 consecutive 50 ms windows were considered significant.

Results

Neuropsychological Data

Insofar as children were pseudo-randomly assigned to the 2 different training groups taking into account the results at the neuropsychological tests, the 2 groups did not differ before training in any of the tests used (the two groups did not differ before training on: Reading age, $P = 0.31$; digit span direct, $P = 0.79$; digit span indirect, $P = 0.79$; digit span total, $P = 0.95$; similitudes, $P = 0.26$; symbols, $P = 0.48$; PM47, $P = 0.75$; visual attention, $P = 0.62$; arrows, $P = 0.68$; auditory attention, $P = 0.23$; orientation, $P = 0.49$; visuomotor precision, $P = 0.42$; irregular words reading, $P = 0.81$; regular words reading, $P = 0.39$; pseudo-words reading, $P = 0.34$; phoneme suppression, $P = 0.65$; phoneme fusion, $P = 0.84$; logatom repetition, $P = 0.26$). The level of performance in both groups improved from T0 to T1 and T2. An analysis of neuropsychological data (WISC-IV and NEPSY, see the Materials and Methods section) across the 3 test sessions showed a main effect of session (RM-MANOVA: $F_{2,21} = 85$, $P < 0.001$). This pattern of results was expected given that children were approximately 1 year

older at T1 and 2 years older at T2. Importantly, this improvement was of similar size in both the music and the painting training groups (main effect of group: $F < 1$; group by session interaction: $F < 1$).

Behavioral Data

Figure 1B clearly shows that the level of performance steadily increased for the music group across testing sessions while it did not change for the painting group (group by session interaction $F_{2,44} = 3.4$, $P = 0.04$). Most importantly, while the level of performance in the music group was at chance at T0 ($P = 0.40$, one-sample t-test), it was higher than chance after 1 (T1) and 2 years (T2) of music training ($P = 0.02$ and 0.004, respectively). Moreover, the benefit due to the second period of training was similar to the benefit of the first period ($P = 0.80$). In contrast, the level of performance in the painting group remained at chance level (0.5) at T0, T1, and T2 (P always >0.40).

Results were further analyzed by taking into account the statistical structure of each of the familiar items. Familiarity accuracy at T1 and T2 was higher in trials containing items with high TP than in trials containing items with low TP (Fig. 2, main effect of Items at T1: $F_{3,66} = 3.1$; $P = 0.03$; main effect of Items at T2: $F_{3,66} = 2.9$; $P = 0.04$).

This fine-grained analysis showed several important results (Fig. 2). First, children in the music group performed better than children in the painting group on almost all items (main effect of group: $F_{1,22} = 5.3$; $P = 0.03$ at T1 and $F_{1,22} = 5.4$; $P = 0.02$ at T2). Secondly, while children in the music group did not succeed at recognizing the pseudo-word with the lowest TP at T1 (performance was at chance, $P = 0.27$, one-sample t-test), they did succeed at T2 ($P = 0.01$, one-sample t-test). Thirdly, although children in the painting group were still at chance level with all items at T1 (P's > 0.16, one-sample t-test), they were above chance level on the 2 pseudo-words with the highest TP at T2, although this did not reach significance (P's > 0.14, one-sample t-test).

Electrophysiological Data

The difference between ERPs to familiar and unfamiliar items was tested with a 4-way RM ANOVA including group (music vs. painting) as a between-subject factor and familiarity (familiar vs. unfamiliar), antero-posterior (frontal, central and parietal) and hemispheres (right and left) as within-subject factors. At T0, and in both groups, unfamiliar items elicited a larger negativity than familiar items between 450 and 550 ms post-stimulus onset over frontal regions (familiarity by antero-posterior interaction: $F_{2,40} = 10.5$, $P < 0.001$). Post hoc analyses revealed that ERPs were more negative for unfamiliar than for familiar items over frontal regions only (−2.6 μV of effect size; $P = 0.008$). The hemisphere and group factors were not significant in the main effects or in the interactions with the other factors (all P's > 0.29).

At T2, and in both groups, the familiarity effect was significant over frontal regions in the 200–300 ms and in the 450–550 ms ranges (familiarity by antero-posterior interaction: $F_{2,40} = 4.8$; $P = 0.03$ and $F_{2,40} = 5.0$, $P = 0.03$, respectively). Post hoc comparisons revealed that the familiarity effect between 200 and 300 ms was maximal over frontal regions, but this difference did not reach significance (1.7 μV of effect size; $P = 0.09$). In contrast, unfamiliar items elicited significantly
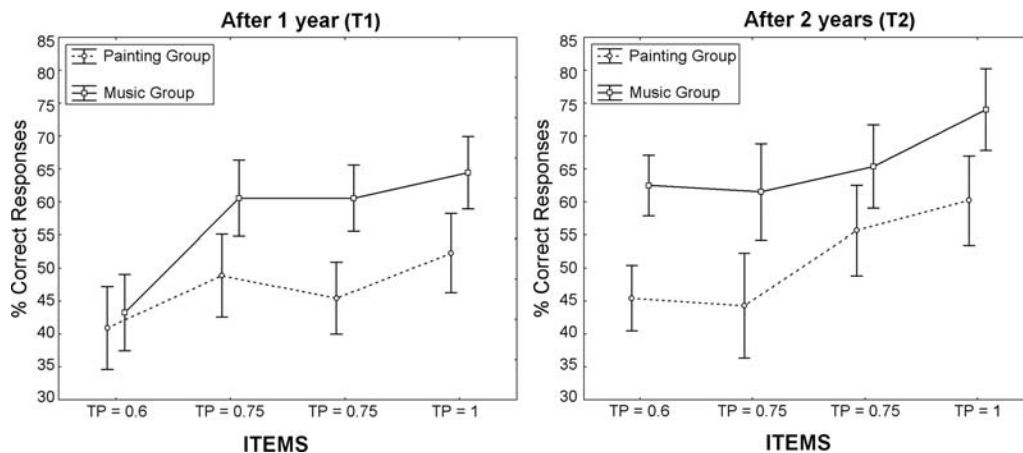
Figure 2. Effects of music and painting training on the sensitivity to TP of each item (given along the *x*-axis). Percentage of correct responses is given for each of the 4 familiar items after 1 year (T1) and after 2 years (T2) of training for the music (solid line) and painting (dashed line) groups.

more negative ERPs than familiar ones over frontal regions between 450 and 550 ms (2.4 µV of effect size; $P = 0.007$). Most importantly, the familiarity effect in the 450–550 ms latency window was larger after 2 years in the music group (2 µV of effect size; $P = 0.002$) than in the painting group (0.1 µV of effect size; $P = 0.99$; familiarity by group interaction: $F_{1,20} = 7.9$, $P = 0.01$). In both time windows, the main effect of hemisphere or the interactions involving this factor were not significant (all P's > 0.14), except for a main effect of hemisphere in the 200–300 ms latency band ($P = 0.06$).

Discussion

The main findings of the present study can be summarized as follows. Children with musical training improved their speech segmentation abilities while children in the painting group did not. Moreover, while the electrophysiological responses were different for familiar and unfamiliar words in both groups, this difference was greater in the music group than in the painting group.

The data reported here extend previous findings showing that in adults, musical expertise facilitates speech segmentation (François and Schön 2011). In this previous experiment, behavioral results showed a trend for a musical practice advantage and electrophysiological data revealed a significantly larger fronto-central negative component for unfamiliar than for familiar items in musicians only. Interestingly, behavioral data in children showed a clear advantage of the music training group. This slight discrepancy between adults and children behavior might be due to the fact that the stream and test used with adults were more complex. An alternative explanation could be that both adult musicians and nonmusicians are already skilled enough at stream segmentation while this ability is still developing in 8 year-old children, thus allowing to observe training-related differences. Concerning electrophysiological data, the morphology and topography of the negative component were similar in adults and children. Indeed, as it was the case in adult musicians, a fronto-central negative component was sensitive to the degree of familiarity of the items in the children music group only. This music training advantage could emerge, possibly via increased

efficiency of general mechanisms involved in regularity extraction and sequence learning (Janata and Grafton 2003).

Music training thus fosters brain plasticity and facilitates speech segmentation. This facilitation may result from several (but not exclusive) processes. Music training may improve general auditory encoding abilities encompassing the brainstem and auditory regions that, in turn, facilitate speech segmentation (Tallal and Gaab 2006; Kraus and Chandrasekaran 2010). Alternatively, music training may facilitate the emergence of more stable memory traces via a more efficient working memory and sequencing processes integrating pitch and syllabic structures, through anatomical and/or functional modifications going beyond the auditory regions. Finally, music training may reduce the effect of interference of adjacent syllables/items (Pechmann and Mohr 1992; Berti et al. 2006), possibly via more efficient temporal dynamic processing (Tallal and Gaab 2006), focusing of attention (Baumann et al. 2008), or executive functions (Moreno et al. 2011).

In this respect, 2 results are of particular interest. First, accuracy was significantly higher with items having high TP than with items having low TP in the music group at T1 and T2. In contrast, this difference did not reach significance in the painting group, although there was a trend at T2 (Fig. 2). Thus, children were sensitive to TP and not simply to differences in the frequency of occurrence when choosing between the "pseudo-word" and the "partial pseudo-word" in a given trial ("pseudo-words" being heard 3 times more often than "partial pseudo-words" during the learning phase).

Secondly, the scalp distribution of the familiarity effect is frontal, in line with previous ERPs and functional magnetic resonance imaging data showing activity in the inferior and middle frontal gyri taken to index the implicit detection of word boundaries in adults and children (McNealy et al. 2006, 2010, 2011; Cunillera et al. 2009). Interestingly, Slumming et al. (2002) reported an increased grey matter density and volume in the left inferior frontal gyrus of musicians (Slumming et al. 2002). Therefore, while music training certainly influences the functional organization of the auditory subcortical and cortical network, it seems that its impact on brain plasticity goes beyond the auditory system tapping onto the dorsal and ventral pathways which seem to play an important

role in language acquisition and higher order processes (Scott and Wise 2004; Hickok and Poeppel 2007; Rodriguez-Fornells et al. 2009).

Importantly, the longitudinal approach, coupled with pseudo-random assignment to 1 of the 2 training groups, controls for possible pre-existing predispositions for music and ascertains that music is the cause of the observed changes (e.g., Schellenberg 2004; Moreno et al. 2009). Note also that the effect described here is not a general effect due to higher motivation or arousal in the music class, since children in both groups improved equally well in the neuropsychological tests. Intriguingly, we did not replicate the results reported by Schellenberg (2004) who showed significant improvement on intelligence tests after 1 year of musical training. However, one should note that Schellenberg used the IQ-full scale including 5 subtests for each type of IQ. Because our longitudinal study did include several experiments, we only used a subset of performance and verbal IQ in order to keep the testing session duration reasonable. Moreover, other important differences are related to the age of the children and to the sample sizes. Children enrolled in Schellenberg's study were younger (6 years old) than the ones enrolled in the present study (8 years old). Musical training at this younger age may act as the general school environment and improve IQ as typically seen at the start of schooling (Schellenberg 2004). Also, the sample size was twice as large in Schellenberg's study than in ours. This factor seems to play an important role since Moreno et al. (2011) recently tested 64 children and reported improved verbal ability after 20 days of musical training but not of visual arts training. In contrast, in a previous study of our group, conducted with another sample of 37 8-year-old children (Moreno et al. 2009), the group by session interaction for the full-scale IQ was not significant ($P > 0.20$). Also Hyde et al. (2009) conducted a similar longitudinal experiment with 31 children and failed to find any significant differences in general intelligence measures while they observed structural changes in auditory and motor brain areas after 15 months of training. They suggested that larger groups may be better suited to confirm the Schellenberg results. Thus, the similar general improvement in the 2 groups is most likely driven by maturation and repetition effects. In contrast, both the higher level of performance in the speech segmentation task and the larger ERP familiarity effect in the music group were driven by musical training (Fig. 3).

Our findings provide new evidence that music training can play an important role in children's language development by facilitating speech segmentation, a building block of language acquisition. Importantly, speech segmentation abilities are known to be closely linked to other speech abilities in typically developing children and to be impaired in children with speech disorders. Recent studies in typically developing infants and school-age children point to a strong link between speech segmentation abilities and more general linguistic proficiency such as expressive lexicon (Newman et al. 2006) and foreign language proficiency (McNealy et al. 2011). Moreover, children with language-based learning impairments not only have difficulties in speech segmentation tasks (Evans et al. 2009) and an impoverished perception of speech rhythms (Abrams et al. 2009; Goswami et al. 2011), but also have a poorer performance than typically developing children in tasks involving musical metrical structures (Huss et al. 2011).
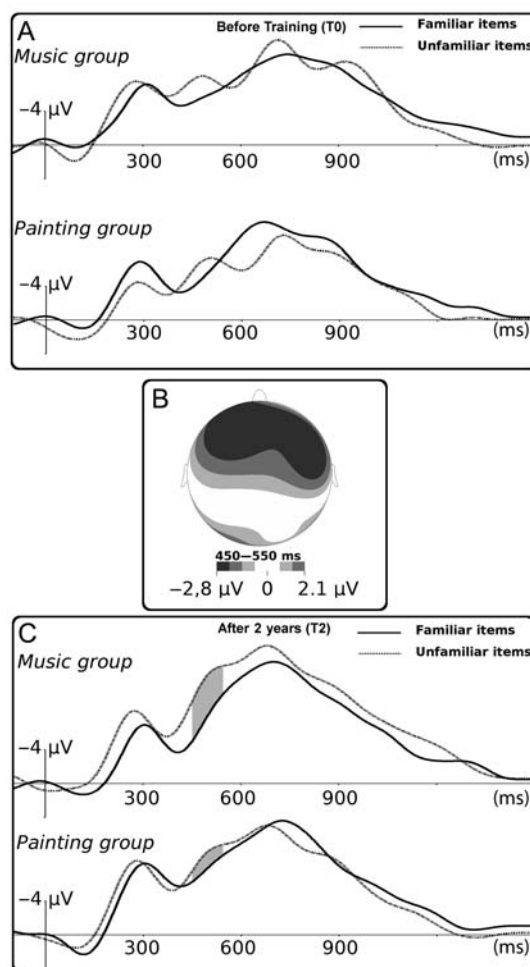

Figure 3. (A) Grand-averages event-related potentials to familiar (thick) and unfamiliar (dotted) items (averaged across items) recorded during the behavioral task before training (T0) in the music group (top) and in the painting group (below). (B) Map showing the distribution of the effect (unfamiliar–familiar) in the 450–550 ms range window (averaged across sessions and groups). (C) Grand-averages event-related potentials to familiar (thick) and unfamiliar (dotted) items (averaged across items) recorded during the behavioral task after 2 years (T2) in the music and in the painting groups (F3 electrode, $n = 12$). Shaded area shows a significant larger familiarity effect size in musically trained children compared with painting-trained children.

This strongly supports the view that musical training, by fostering rhythm perception and production, may have an important role for the development of language skills in children. Thus, taken together, these results favor the idea that by developing both perceptual and cognitive functions, music training shapes individual development (Patel 2010).

References

Abrams DA, Nicol T, Zecker S, Kraus N. 2009. Abnormal cortical processing of the syllable rate of speech in poor readers. J Neurosci. 29(24):7686–7693.

Aslin RN, Saffran JR, Newport EL. 1998. Computation of conditional probability statistics by 8-month-old infants. Psychol Sci. 9 (4):321–324.

Baumann S, Meyer M, Jäncke L. 2008. Enhancement of auditory-evoked potentials in musicians reflects an influence of musical expertise but not selective attention. J Cogn Neurosci. 20 (12):2238–2249.

Bermudez P, Zatorre RJ. 2005. Differences in gray matter between musicians and nonmusicians. Ann N Y Acad Sci. 1060:395–399.

Berti S, Münzer S, Schröger E, Pechmann T. 2006. Different interference effects in musicians and a control group. Exp Psychol. 53 (2):111–116.

Besson M, Chobert J, Marie C. 2011. Transfer of training between music and speech: Common processing, attention, and memory. Front Psychol. 2:94.

Brattico E, Pallesen KJ, Varyagina O, Bailey C, Anourova I, Järvenpää M, Eerola T, Tervaniemi M. 2009. Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. J Cogn Neurosci. 21(11):2230–2244.

Cunillera T, Càmara E, Toro JM, Marco-Pallares J, Sebastian-Galles N, Ortiz H, Pujol J, Rodriguez-Fornells A. 2009. Time course and functional neuroanatomy of speech segmentation in adults. Neuroimage. 48(3):541–553.

Evans JL, Saffran JR, Robe-Torres K. 2009. Statistical learning in children with specific language impairment. J Speech Lang Hear Res. 52(2):321–335.

François C, Schön D. 2011. Musical expertise boosts implicit learning of both musical and linguistic structures. Cerebral Cortex. 21 (10):2357–2365.

Gervain J, Macagno F, Cogoi S, Peña M, Mehler J. 2008. The neonate brain detects speech structure. Proc Natl Acad Sci USA. 105:14222–14227.

Goswami U, Wang HL, Cruz A, Fosker T, Mead N, Huss M. 2011. Language-universal sensory deficits in developmental dyslexia: English, Spanish and Chinese. J Cogn Neurosci. 23(2):325–337.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nat Rev Neurosci. 8:393–402.

Huss M, Verney JP, Fosker T, Mead N, Goswami U. 2011. Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. Cortex. 47 (6):674–689.

Hyde KL, Lerch J, Norton A, Forgeard M, Winner E, Evans AC, Schlaug G. 2009. Musical training shapes structural brain development. J Neurosci. 29(10):3019–3025.

Janata P, Grafton ST. 2003. Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. Nat Neurosci. 6:682–687.

Keenan JP, Thangaraj V, Halpern AR, Schlaug G. 2001. Absolute pitch and planum temporale. Neuroimage. 14(6):1402–1408.

Koelsch S, Schröger E, Tervaniemi M. 1999. Superior pre-attentive auditory processing in musicians. Neuroreport. 10(6):1309–1313.

Kraus N, Chandrasekaran B. 2010. Music training for the development of auditory skills. Nat Rev Neurosci. 11(8):599–605.

Kuhl PK. 2004. Early language acquisition: Cracking the speech code. Nat Rev Neurosci. 207(4427):203–205.

Magne C, Schön D, Besson M. 2006. Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. J Cogn Neurosci. 18(2):199–211.

McNealy K, Mazziotta JC, Dapretto M. 2011. Age and experience shape developmental changes in the neural basis of language-related learning. Dev Sci. 14(6):1261–1282.

McNealy K, Mazziotta J, Dapretto M. 2006. Cracking the language code: Neural mechanisms underlying speech parsing. J Neurosci. 26(29):7629–7639.

McNealy K, Mazziotta JC, Dapretto M. 2010. The neural basis of speech parsing in children and adults. Dev Sci. 13(2):385–406.

Moreno S, Bialystok E, Barac R, Schellenberg EG, Cepeda NJ, Chau T. 2011. Short-term music training enhances verbal intelligence and executive function. Psychol Sci. 22(11):1425–1433.

Moreno S, Marques C, Santos A, Santos M, Castro SL, Besson M. 2009. Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. Cerebral Cortex. 19 (3):712–723.

Musacchia G, Sams M, Skoe E, Kraus N. 2007. Musicians have enhanced subcortical auditory and audiovisual processing in speech and music. Proc Natl Acad Sci USA. 104(40):15894–15898.

Newman R, Ratner NB, Jusczyk AM, Jusczyk PW, Dow KA. 2006. Infant's early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. Develop Psychol. 42(4):643–655.

Patel AD. 2010. In: Bailar M, editor. Music, biological evolution, and the brain. Emerging disciplines. Houston: Rice University Press. p. 91–144.

Pechmann T, Mohr G. 1992. Interference in memory for tonal pitch: Implications for a working-memory model. Mem Cognit. 20 (3):314–320.

Rodriguez-Fornells A, Cunillera T, Mestress-Misse A, De Diego Balaguer R. 2009. Neurophysiological mechanisms involved in language learning in adults. Phil Trans Roy Soc B. 364:3711–3735.

Rüsseler J, Altenmüller E, Nager W, Kohlmetz C, Münte TF. 2001. Event-related brain potentials to sound omissions differ in musicians and non-musicians. Neurosci Lett. 308:33–36.

Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month old infants. Science. 274:1926–1928.

Saffran JR, Johnson E, Aslin RN, Newport EL. 1999. Statistical learning of tone sequences by human infants and adults. Cognition. 70 (1):27–52.

Schellenberg EG. 2004. Music lessons enhance IQ. Psychol Sci. 15 (8):511–514.

Schlaug G, Janckë L, Huang Y, Steinmetz H. 1995. In vivo evidence of structural brain asymmetry in musicians. Science. 267 (5198):699–701.

Schön D, Boyer M, Moreno S, Besson M, Peretz I, Kolinsky R. 2008. Song as an aid for language acquisition. Cognition. 106(2):975–983.

Scott SK, Wise RJ. 2004. The functional neuroanatomy of prelexical processing in speech perception. Cognition. 92(1-2):13–45.

Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J Neurosci. 23(13):5545–5552.

Shahin A, Roberts LE, Pantev C, Trainor LJ. 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport. 16(16):1781–1785.

Slumming V, Barrick T, Howard M, Cezayirli E, Mayes A, Roberts N. 2002. Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. Neuroimage. 17(3):1613–1622.

Tallal P, Gaab N. 2006. Dynamic auditory processing, musical experience and language development. Trends Neurosci. 29(7):382–390.

Teinonen T, Fellman V, Näätänen R, Alku P, Huotilainen M. 2009. Statistical language learning in neonates revealed by event-related brain potentials. BMC Neurosci. 13:10–21.

Wong PCM, Skoe E, Russo NM, Dees T, Kraus N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci. 10:420–422.