Music in the first days of life

*D. Perani[1,2,3,4], *M.C.Saccuman[1,2], P. Scifo[2,3], D. Spada[7],
G. Andreolli[1], R. Rovelli[6], C. Baldoli[2,3], S. Koelsch[8]

[1] Vita-Salute San Raffaele University and [2] San Raffaele Scientific Institute, Milan, Italy
[3] C.E.R.M.A.C. San Raffaele Scientific Institute, Milan, Italy
[4]National Neuroscience Institute, Italy
[5] Department of Neuroradiology, San Raffaele Scientific Institute, Milan, Italy
[6]Department of Neonatology, San Raffaele Scientific Institute, Milan, Italy
[7] Psychology Institute, School of Medicine, Universita'degli Studi, Milan, Italy
[8]Max-Planck-Institute for Neuropsychology, Leipzig, Germany

* These authors equally contributed to the work

In adults, specific neural systems with right-hemispheric weighting are necessary to process pitch, melody and harmony, as well as structure and meaning emerging from musical sequences. To which extent does this neural specialization result from exposure to music or from neurobiological predispositions? We used fMRI to measure brain activity in 1 to 3 days old newborns while listening to Western tonal music, and to the same excerpts altered, so as to include tonal violations or dissonance. Music caused predominant right hemisphere activations in primary and higher-order auditory cortex. For altered music, activations were seen in the left inferior frontal cortex and limbic structures. Thus, the newborn's brain is able to plenty receive music and to figure out even small perceptual and structural differences in the music sequences. This neural architecture present at birth provides us the potential to process basic and complex aspects of music, a uniquely human capacity.

Human musical activity covers an extremely wide spectrum, ranging from a non-musician's effortless appreciation of music in everyday life to the highly specialized abilities of expert performers and composers. If human beings have learnt to create and appreciate music, this means that our brains developed along the evolution specific neural systems to process the basic elements of music as well as its complexity. The perception of music requires sophisticated cognitive machinery for the decoding of pitch, rhythm, and timbre, as well as for the computation of strings of auditory elements that represent forms that convey musical meaning. Neuropsychological, neurophysiological and neuroimaging studies have described a network for music perception in adult non-musicians, involving the superior temporal gyrus, as well as inferior frontal and parietal areas, with a dominance of the right over the left hemisphere (1). It is still unclear, however, to which extent these neural correlates are due to adaptation of the brain resulting from exposure to the musical environment and to neurobiological constraints. Behavioural data from infants of 2 months of age, with casual everyday exposure to music shows that babies are surprisingly skilled at perceiving subtle aspects of musical stimuli, such as consonant versus dissonant intervals, the relational processing of pitch and tempo, and variations in meter, timbre, tempo and duration (2). This points to the existence of a neural predisposition for music processing and raises the hypothesis of a specificity of the underlying neural substrate. Yet, it is unknown whether a neuronal architecture that serves the processing of music is already present at birth.

We used functional magnetic resonance imaging (fMRI) to investigate the neural correlates of music processing in newborn infants. Our subjects were 18 healthy, full-term, non-sedated newborns, within the first three days of life. We collected fMRI images as babies heard 21-second musical stimuli alternating with 21-second blocks of silence (the noise produced by the MRI scanner was strongly attenuated by active noise cancellation).

Three sets of musical stimuli were used (see fig. 1, A): Set 1 ("Music") consisted of original instrumental (piano) excerpts drawn from the corpus of major-minor tonal Western music of the 18th and 19th century. From these excerpts, two further sets ("Altered Music") were created that were identical to the original excerpts with regards to rhythm, tempo and melodic contour, but in which the musical structure was manipulated. For Set 2 ("Key Shifts"), all voices were infrequently shifted one semitone up- or downwards, thus infrequently shifting the tonal centre to a key, which was harmonically only distantly related to the preceding harmonic context (e.g., from *C* major to *C#* major). In Set 3 ("Dissonance"), the upper voice (i.e. the melody) of the musical excerpts was permanently shifted one semitone upwards, rendering the excerpts more dissonant, and making the tonal key ambiguous. Acoustically, the altered stimulus sets differed from the original excerpts in that the infrequent shifts in tonal key (Set 2) introduced a subtle sensory dissonance (see supplementary materials), and that the dissonant excerpts (Set 3) had a higher degree of acoustic roughness. Such acoustic cues are regarded as important for children to identify and learn music-syntactic irregularities (3). Music-syntactic irregularities, which occur without such cues, are hardly detected by newborns, due to a lack of musical experience. In terms of musical structure, both altered stimulus versions differed from original excerpts in that the extraction of a tonal center was more difficult. Because extraction of a tonal center is a pre-requisite for the music-syntactic processing of harmonies, presentation of Sets 2 and 3 enabled us to investigate infants' first steps into music-syntactic processing.

Infants listened to two 7-minute sequences, with music blocks alternating between conditions and silence in pseudo-random order (Music – Silence – Altered Music – Silence) (see fig. 1, B), with each of the two sequences containing only one kind of music alteration. MR images were acquired on a Phillips 1.5 Tesla Intera scanner.

Functional images were processed within the framework of the general linear model in AFNI (4). Scans were inspected for signal outliers and motion artifacts. Each infant's usable sequences were registered to one stable scan for within-subject motion correction and realignment. Scans that required a shift of more than 3 mm or 3 degrees were ignored in subsequent analysis steps. Images were registered to an infant template created from the group's functional scans. The blood oxygenation level dependent (BOLD) response to each stimulus condition compared to baseline was estimated with multiple regressions, in a model that included the output from the motion-correction algorithm as regressors of no interest. Regressors of interest were obtained by convolving a square wave for each stimulation block of that condition with a gamma variate function approximating the hemodynamic response. A second level random-effects model was used to create group maps in the common template space. The estimates obtained through the regression were entered in two-way mixed-effects ANOVAs performed on each voxel in template space, with stimulus type as fixed variable, and subjects as random variables.

RESULTS

Analyses for ***Music vs. Silence***, showed an extended right hemispheric activation cluster focused in the superior temporal gyrus, from the primary and secondary auditory cortex (transverse temporal gyrus), extending anteriorly towards the planum polare, posteriorly to the planum temporale, the temporoparietal junction and the inferior parietal lobule. In addition, activation was observed in the right insula. In the left hemisphere, there was an activated cluster centred in the primary and secondary auditory cortices (Fig 2 A, first row). The right-hemispheric predominance of activations was confirmed by a Region of Interest analysis ($p < 0.05$, see below).

We then addressed the differences in activation between the two altered conditions, namely – ***Key Shifts vs. Silence*** and ***Dissonance vs. Silence***. Results showed that both these contrasts differed from the contrast Music vs. Silence. We have indeed found in case of both comparisons a less extended activation in the right temporal regions and, noteworthy, a left hemispheric involvement, with activation clusters in the superior and middle temporal regions, the inferior frontal gyrus, amygdala and ventral striatum. The *direct comparison* between the two altered conditions showed no significant differences, apart for the left amygdala and ventral striatum in which BOLD signal changes were more significant for the *Key Shifts*.

Thus, for the following steps of analysis, both altered versions were pooled and contrasted to silence, as well as to the original music excerpts. Note that pooling the two Altered Music Sets resulted in an equal amount of scans for original and altered excerpts (see methods).

For the ***pooled altered excerpts ("Altered Music") vs. Silence***, analyses revealed an activation of limited extent in the right hemisphere with a cluster of activated voxels including the primary auditory cortex and the posterior part of the superior temporal gyrus up to the temporoparietal junction. On the left, the primary auditory cortex was active, together with additional activations in the inferior frontal gyrus, amygdala and ventral striatum (Fig.2 A, second row).

The ***direct comparison between Music and Altered Music (pooled)*** confirmed the more significant activation for *Music* in the right superior temporal gyrus, whereas in the left hemisphere, the inferior frontal gyrus, the amygdala and ventral striatum were significantly more activated by the altered stimuli.

A ***region of interest (ROI) approach*** was used to specifically address the level of activity in the right and left primary and secondary auditory cortex. For this analysis, oval ROIs comprising the primary and secondary auditory cortices were drawn on the group template. Each subject's average percent signal change for each condition in each ROI was then entered into random-effects ANOVA with hemisphere (left/right) and stimulus type (Music and Altered Music) as within-subjects factors. This analysis showed a significant interaction between hemisphere and stimulus type (F = 4.74, p = 0.03), reflecting a difference in activation for the two stimulus types in the right hemisphere, with more activation for Music (Multiple Comparison Tukey HSD, p = 0.05). In the left hemisphere, no difference in activation was observed for the two stimulus types (Fig. 2 B).

DISCUSSION

Our results reveal a predisposition for music processes in the brains of newborns. A hemispheric functional asymmetry for music is indeed present at birth, with clear right hemisphere predominance in the primary, secondary and higher order auditory cortices (Figure 2, A and B). In adults, primary auditory cortex, particularly in the right hemisphere, is important for basic aspects of pitch processing, such as decoding of pitch height, pitch croma and pitch direction (5)(6). Right hemispheric areas of auditory cortex outside the primary zone are specialized in the processing of pitch patterns, the encoding and recognition of melodies, as well as in auditory working memory mechanisms (5)(6). Previous functional neuroimaging studies with adults have shown that superior temporal gyrus and planum polare become increasingly involved with increasing melodic complexity (7), and here we found in newborns activations beyond the Heschl's gyrus towards the right planum polare, as well as in the anterior superior temporal gyrus and planum temporale, for music. It has been suggested that the basic features of individual notes, such as pitch, timbre and intensity, are analyzed in the pathway up to and including the auditory cortices, while higher-order musical patterns formed by those features are analyzed by distributed networks in the temporal lobe (7). Similarly, the left temporal lobe seems to be anatomically specialized for phonological, lexical, and sentence-level processing (8).

There is an agreement in considering a hierarchical organization in the brain corresponding to the processing of music and language (7)(9)(10). The speech-specific and music-specific neural processing seems to emerge from differential demands on auditory and higher order processes for the different domains. In humans, the right and left primary auditory cortex has been proposed to be associated with complementary specializations, with the left performing better for temporal resolution, and the right for spectral resolution (11)(12). Fine temporal resolution is crucial for the phonetic analysis of speech, while fine spectral resolution is critical for pitch processing. These functional hemispheric differences correspond to neuroanatomical findings of gray and white matter differences in the left and right auditory cortices (13). The left hemisphere hosts larger fibre bundles and larger fourth layer neurons, leading to faster signal transmission. In the right hemisphere, more tightly packed neurons result in more fine-grained tonotopy.
Asymmetrical cortical patterning appears to occur very early in ontogeny. Gross anatomical asymmetries around the Sylvian fissure have been observed in fetuses from the middle of gestation, and in infants (14)(15), together with early asymmetry of gene transcription in embryonic human left and right cerebral cortex (16). In older infants (2 to 3 months old), functional asymmetry with left-hemispheric dominance has been observed for language processing (17). Our results are the first to show that a neural specialization for the analysis of musical information including variation of pitch patterns is present at birth, in spite of relative auditory cortex immaturity (18).

The second major result obtained is that we observed a significant difference in activation when babies were exposed to musical stimuli that were structurally altered. We saw a reduced BOLD signal in the right primary and secondary auditory cortex in comparison to the normal excerpts and, noteworthy, additional left-hemispheric activations in associative temporal and inferior. This extended left network renders it unlikely that just auditory/acoustical processing is taking place. Previous research in adults has implicated bilateral inferior frontal lateral cortex in the processing of music-syntactic information (19). The altered stimulus sets required more complex music-syntactic processing because the representation of a tonal center is more difficult to extract: in the dissonant versions the tonal center was ambiguous, and each key shift required the establishment of a new tonal center (note that the altered stimulus sets were identical to the original music in terms of melodic contour, tempo, and rhythm).
Interestingly, our data in combination with those previously reported for language processing in babies (17) suggest that babies engage the left inferior frontal cortex for the processing of more complex aspects of music and speech. This is also consistent with the notion that children use the prosodic aspects of speech (that is, the musical features of speech) to learn language, and that children make less difference between music and language than adults do (19).
It might be worth pointing out that the newborn auditory periphery (inner ear and brain stem) is capable of reacting differently to music and to acoustic deviances, such as those present in our altered stimuli. Such capability is crucial for the acquisition of

representations of music-syntactic regularities during early childhood, because music-syntactic irregularity and harmonic distance often co-occurs with acoustic deviance, such as reduced acoustic similarity and increased sensory dissonance. Hence, mechanisms underlying acoustic deviance detection provide sometimes information about the irregularity of chord-functions. Such information aids the detection of music-syntactic regularities, and the build-up of a structural model of a sequence of musical events.

Our data show that babies engage neural resources located in the left hemisphere in response to acoustical deviances that give rise to an extraction of a tonal centre. Such mechanism might represent the starting point for music-syntactic processing.

Interestingly, this left hemispheric system will be of fundamental importance for language-syntactic processing later on in the development. It is inextricably linked to the establishment and maintenance of representations of the complex acoustic environment, and thus to the analysis of music and language auditory scene.

The newborns also showed increased activation of limbic regions (amygdala, ventral striatum) with left side prevalence during the presentation of the altered dissonant music pieces. For adults, consonant music has a different emotional valence than dissonant music (20)(21), and adult listeners perceive changes in key, comparable to those used in the present study, as unpleasant compared to musical sequences with no key change (22).

We cannot exclude that the activations of limbic structures in the small infants reflect emotional responses to the altered musical stimuli. It might also be linked to the activated superior temporal system, due to the presence early established, reciprocal fibre tract connections (23).

Likewise, it appears that, ontogenetically, infants' first steps into language are based on prosodic information, and that musical communication in early childhood (such as maternal language) plays a major role for emotional, cognitive, and social development in children. Infants acquire considerable information about word and phrase boundaries (possibly even about word meaning) through different types of prosodic cues (i.e., the musical cues of language such as speech melody, metre, rhythm, and timbre)(24). Young children are surprisingly skilled at perceiving subtle aspects of musical stimuli. With respect to music processing, they mostly behave like musically untrained adults in perceiving consonant versus dissonant intervals, in the relational processing of pitch and tempo and in their sensitivity to variations in meter, timbre, tempo and duration. Human infants are inherently musical and music modulates their attention and arousal levels, and seems to provide pleasure or distress, as in the case of dissonant music (25) (2).

Our results show that a neuronal architecture that serves the processing of music is already present at birth. The newborn's brain is able to perceive music and to respond to small perceptual and structural differences in the music sequences. The present findings substantially contribute to our understanding of the neural basis of a universal and uniquely human capacity that is music.

METHODS

Subjects. 18 healthy full-term newborns (8 girls, 10 boys) within the first three days of life participated in the study. Gestation and birth histories were normal for all subjects. Data from 3 other newborns were not used, due to fuzziness resulting in large-scale movements throughout scanning. The newborns were not sedated.
Subjects' immediate family members were in the majority right-handed (88% right-handed), and parents reported no significant family history of learning disabilities or psychiatric and neurological disorders. Parents were in the majority of monolingual Italian background (one bilingual English-Italian), and they were not musicians. Mothers reported casual exposure to popular and classical music during pregnancy. One mother sang in a choir. Parents gave written consent in accordance with the procedures approved by the San Raffaele Hospital Ethical Committee.

Stimuli. Three sets of stimuli were created. Set 1 (Music) consisted of 10 instrumental tonal music excerpts, drawn from the corpus of major-minor tonal ("Western") classical music. The excerpts were taken from the following pieces:

1. J.S. Bach - French Suite No.1 in D- BWV812, Allemande
2. Scarlatti - Sonata in E, K.162, L.21
3. Schubert - Six Moments Musicaux, D.780, Op.94, No.3 in F-
4. Schubert - Andante in C, D.29
5. Schubert - 34 Valses sentimentales, D.779, Op.50
6. J.S. Bach - Two-Part Inventions, No.5 in Eb, BWV776
7. J.S. Bach - Two-Part Inventions, No.11 in G-, BWV782
8. Scarlatti - Sonata in D-, K.10, L.370
9. Scarlatti - Sonata in D, K.21, L.363
10. Mozart - Piano Sonata in G, K.283

Each excerpt was 21 seconds long. Sets 2 and 3 (Altered Music) were obtained by manipulating the excerpts of Set 1. For Set 2 all notes were shifted one semi-tone higher or lower at the end of cadences. For Set 3, the leading voice (the upper part of the excerpt) was shifted one semi-tone higher.

Imaging protocol. MR images were acquired on a Phillips 1.5 Tesla Intera scanner (Philips Medical Systems, Best, The Netherlands). Functional MRI was performed using an optimized EPI Gradient Echo pulse sequence with the following acquisition parameters: TR/TE=3000/40msec, voxel size= 3.75 x 3.75 x 3mm^3, 23 slices for each volume; 140 volumes for each scan. T2 weighted clinical images were reviewed by a licensed pediatric neuroradiologist (C.B.)

Procedure. Sounds were presented via piezo-electric, CE-certified, MRI compatible headphones, custom-made to fit newborns' ears (http://www.mr-confon.de/en/), and incorporating an active gradient noise-suppression mechanism in addition to passive deadening, leading to noise reduction in the order of 30 to 40 dB above 600 Hz. The sound presentation was adjusted at a comfortable volume level allowing the music to be clearly audible above residual scanner noise.

When infants were quiet, they were swaddled in a soft blanket and placed in a padded, custom-made cradle that fit inside the head coil. The infants were fitted with the piezo-electric headphones, and a foam pad was placed around their heads for additional noise dampening. Infants' behavior during scanning was monitored by a camera and microphone placed inside the magnet bore. To avoid startle reflexes at the onset of the functional scans, and to accustom the infant to MR sounds, we started acquisition with less noisy sequences: scout, T2 weighted clinical imaging and T1 weighted imaging. Scanning was immediately interrupted if the infants became restless.

A block design was used to maximize statistical power with 21 sec blocks alternating between conditions (Music – Silence – Altered Music -Silence) for a total scan time of 7 minutes. Each sequence contained only one kind of altered music. Two 7-minute sequences were presented, in alternate order to each successive infant: one containing Music and Key Shifts alterations, and one containing Music and Dissonance alterations. The two kinds of alteration were obtained from two different sets of 5 excerpts, to avoid repeating the same non-altered stimulus twice.

Data Analysis (See Supporting Material).

ACKNOWLEDGEMENTS
This work was supported by a grant of the Fondazione Mariani to Profs. Daniela Perani and Stefan Koelsch, and by the EU Project BrainTuning FP6-2004 NEST-PATH-028570

---

1. Peretz, I., Zatorre, R.J., Annu Rev Psychol 56, 89-114 (2005)

2. Trehub, S.E., Ann N Y Acad Sci 999, 402-413 (2003)

3. Koelsch S, Psychophysiology (2008) in press

4. Cox R.W Res 29, 162-173 (1996)

5. Zatorre, R., Oxford, pp. 231-246 (2003)

6. Zatorre RJ, Gandour JT, Philos Trans R Soc Lond B Biol Sci. 12;363(1493):1087-104 (Mar 2008) Review.

7. Patterson Roy D., Uppenkamp Stefan,. Johnsrude Ingrid S and Griffiths Timothy D. Neuron,Vol. 36, 767–776 (November 14, 2002)

8. Indefrey P. and Cutler A., Gazzaniga, Editor, *The new cognitive neurosciences III*, MIT Press,Cambridge, MA (2003)

9. Zatorre, R.J., Belin, P., and Penhune, V.B. Trends Cogn. Sci. *6*, 37–46 (2002).

10. Price C., Thierry G. and Griffiths T..TRENDS in Cognitive Sciences Vol.9 No.6 (June 2005)

11. Liegeois-Chauvel C., De Graaf J.B., Laguitton V., Chauvel P.Cereb Cortex 9, 484-496 (1999)

12. Zatorre R.J., Belin P., Cereb Cortex 11, 946-953 (2001)

13. Penhune V.B., Zatorre R.J., MacDonald J.D., Evans A.C. Cereb Cortex 6, 661-672 (1996)

14. Wada J.A., Clarke R., Hamm A., Arch Neurol 32, 239-246 (1975)

15. Chi J.G., Dooling E.C., Gilles F.H., Arch Neurol 34, 346-348 (1977)

16. Sun, T., Patoine, C., Abu-Khalil, A., Visvader, J., Sum, E., Cherry, T.J., Orkin, S.H., Geschwind, D.H., Walsh, C.A., Science 308, 1794-1798 (2005)

17. Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L., Science 298, 2013-2015 (2002)

18. Moore D. R: *British Medical Bulletin* 63:171-181 (2002)

19. Koelsch S. and Siebel W. A. TRENDS in Cognitive Sciences Vol.9 No.12 (December 2005)

20. Blood AJ, Zatorre RJ, Bermudez P, Evans AC, Nat Neurosci.;2(4):382- 7 (Apr 1999)

21. Koelsch S, Fritz T, V Cramon DY, Muller K, Friederici AD, Hum Brain 27(3):239-50 Mapp.(2006)

22. Koelsch S, Kilches S, Steinbeis N, Schelinski S. PLoS-ONE (2008) in press

23. Dubois J., Dehaene-Lambertz G., Perrin M., Mangin J.-F., Cointepas Y.,. Duchesnay E, Le Bihan D. and Hertz-Pannier L. Human Brain Mapping 29:14–27 (2008)

24. Jusczyk P.W. The discovery of spoken language. A Bradford Book The MIT Press, Cambridge, MassachussettsLondon, England **(**1997**)**

25.Trehub, S.E. Ann N Y Acad Sci 930, 1-16 (2001)

26.Leman M., Lesaffre M., Tanghe K., Introduction to the IPEM Toolbox in: Proceedings of the XIII Meeting of the FWO Research Society on Foundations of Music Research, Ghent, Belgium (March 2, 2001).

27.Altman N.R., Bernal, B., Radiology 221, 56-63 (2001)

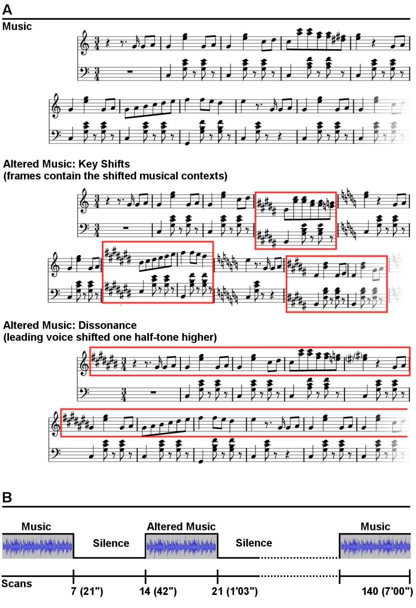
28.Anderson A.W., Marois R., Colson E.R., Peterson B.S., Duncan C.C., Ehrenkranz R.A., Schneider K.C., Gore J.C., Ment L.R., Magn Reson Imaging 19, 1-5 (2001)

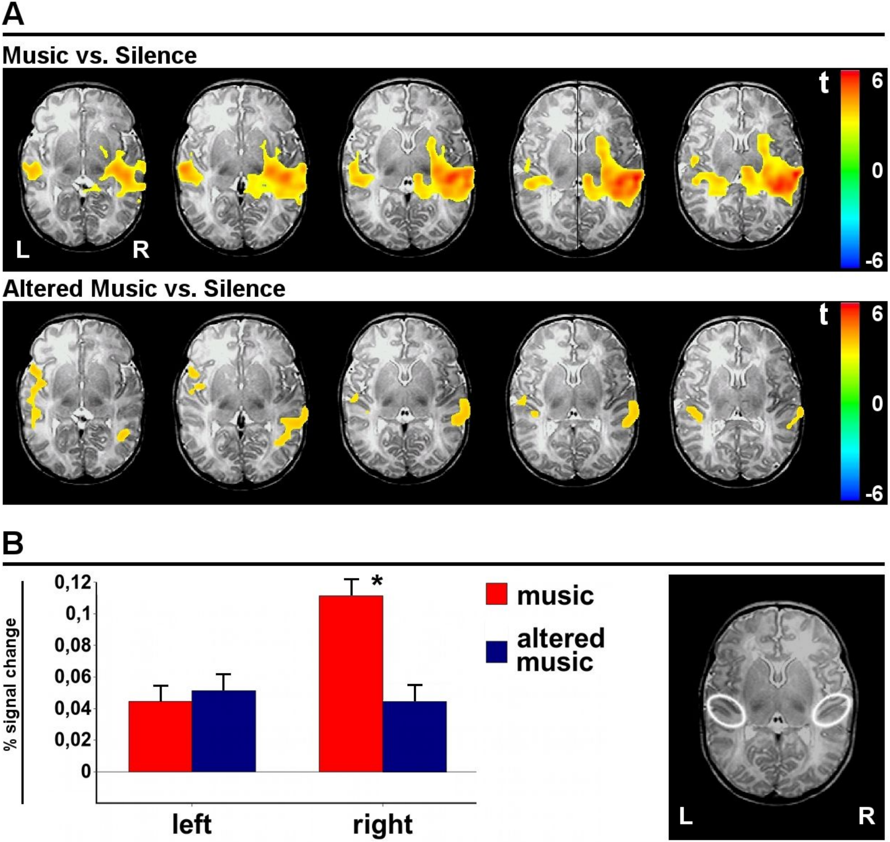
Legends for figures

Figure 1. Samples of musical stimuli presented in a block design
Musical excerpts were presented in as in B, with Classic Music and Altered Music alternating with periods of Silence.
In (A) are shown the three Sets of stimuli: Set 1 (Classic Music): part of one of the original excerpts (in this case, Schubert - Andante in C, D.29); Set 2 (Altered Music: Key-Shifts): the same excerpt of Set 1 altered in order to introduce shifting of the key (the notes in the red frames are shifted one semitone up- or downwards, as indicated by the alterations on the staves and by the shifting of the staves, upwards when the key shift was one semitone upwards, and downwards when key shift was one semitone downwards); Set 3 (Altered Music: Dissonance): the same excerpt of Set 1 altered in order to introduce continuous dissonance (the notes in the red frames (upper stave, leading voice) are shifted one semitone upward, thus creating dissonant harmonic intervals with the notes of the accompaniment that are in the original key).
Note that in the two Altered Music condition each one of the two kinds of alteration (Key-Shift and Dissonance) was obtained starting from a different subset of musical excerpts, in order to avoid the repetition of the unaltered stimuli. In the figure the same excerpt has been used for more clearness.

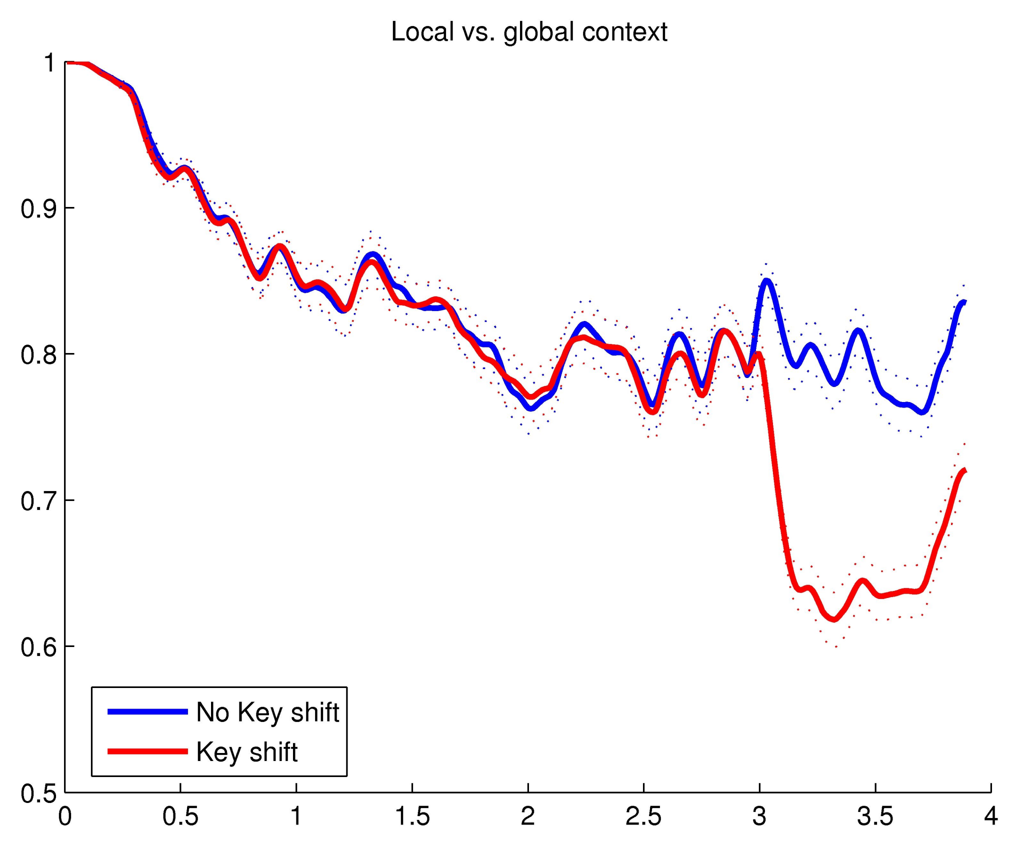
Figure 2. BOLD activations for musical stimuli in healthy non-sedated newborns (n=18)

(A) Mean activation for Classic Music relative to Silence (first row) and Altered Music relative to Silence (second row) in a random effect group analyses (voxel $p>0.0002$, cluster $p<0.05$ corrected). Five axial slices show the active clusters in temporal cortices and frontal regions. Activations are overlapped to a T2 weighed image from a single newborn subject. (B) Changes in activation for Classic Music and Altered Music in primary and secondary auditory cortices, defined by oval ROIs. The histograms show the results of the group ANOVA, with significant interaction between hemisphere and stimulus type and the significant activation for Classic Music in the right hemisphere ($F = 4.74$, $p = 0.03$). In the left hemisphere, the signal was not modulated by stimulus type. The error bars show standard error of the mean.


A
Music
Altered Music: Key Shifts
(frames contain the shifted musical contexts)
Altered Music: Dissonance
(leading voice shifted one half-tone higher)
B
Music
Silence
Altered Music
Silence
Music
Scans
7 (21")
14 (42")
21 (1'03")
140 (7'00")



A
Music vs. Silence
L
R
t
6
0
-6
Altered Music vs. Silence
t
6
0
-6
B
% signal change
0,12
0,1
0,08
0,06
0,04
0,02
0
*
left
right
music
altered music
L
R


Local vs. global context
1
0.9
0.8
0.7
0.6
0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
No Key shift
Key shift