SCIENTIFIC REP⚙RTS

Early auditory processing in musicians and dancers during a contemporary dance piece

Hanna Poikonen[1], Petri Toiviainen[2] & Mari Tervaniemi[1,3]

The neural responses to simple tones and short sound sequences have been studied extensively. However, in reality the sounds surrounding us are spectrally and temporally complex, dynamic and overlapping. Thus, research using natural sounds is crucial in understanding the operation of the brain in its natural environment. Music is an excellent example of natural stimulation which, in addition to sensory responses, elicits vast cognitive and emotional processes in the brain. Here we show that the preattentive P50 response evoked by rapid increases in timbral brightness during continuous music is enhanced in dancers when compared to musicians and laymen. In dance, fast changes in brightness are often emphasized with a significant change in movement. In addition, the auditory N100 and P200 responses are suppressed and sped up in dancers, musicians and laymen when music is accompanied with a dance choreography. These results were obtained with a novel event-related potential (ERP) method for natural music. They suggest that we can begin studying the brain with long pieces of natural music using the ERP method of electroencephalography (EEG) as has already been done with functional magnetic resonance (fMRI), these two brain imaging methods complementing each other.

In neuroscience, the disclosure of the riddle behind why music has such a strong and unique influence on our mind[1,2] began by studying individual sounds and sound streams[3]. Step by step the musical stimuli and the test settings in the brain laboratories became more complex and involved changing keys, vibrant chords and violated harmonies[4–6] as well as musical imagination and improvisation[7–9]. More recently, a big leap in the brain research of music was made when Alluri *et al.* studied the cerebral processing of individual musical features extracted from a whole musical piece played in a functional magnetic resonance imaging (fMRI) scanner[10]. Indeed, music as a whole activates the brain widely[11], but different musical features are processed in different brain regions[10]. Groovy beat travels from the ear into the specific brain structures via different pathways than the sentimental sounds of a violin. While the beat activates movement-related areas, such as the basal ganglia and the supplementary motor area[12], the calming melodic sound decreases the activation in amygdala thus increasing the activation in other limbic regions[13–15].

But how are these musical features processed on a shorter time scale, which is out of measurable reach of the temporal resolution of an fMRI? Is there an immediate difference in the processing of musical features between professional musicians and laymen? How is the hearing system tuned to perceive the musical features among professional dancers who also constantly use music in their work and creation? How does a simultaneously presented dance choreography influence to the auditory responses of the musical features? We chose to approach these thrilling questions utilizing the method of event-related potential (ERP) for electroencephalography (EEG). As we have shown before[16], rapid changes in the musical features of brightness, root mean square (RMS) amplitude, zero-crossing rate and spectral flux during the listening of natural music evokes ERP responses similar to the responses elicited while listening to a series of simple individual sounds.

We chose several long excerpts from the composition Carmen by Bizet-Shchedrin to be presented to professional musicians, professional dancers and a group of participants without any professional background in either music or dance. The musical excerpts were presented as an auditory stimulus, and as an audio-visual entity with a contemporary dance choreography of Carmen. We expected the ERP responses for the musical features to be attenuated and sped up when music was accompanied with concordant dance similar to the results gained

[1]Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, P.O. Box 9, FI-00014, Finland. [2]Department of Music, University of Jyväskylä, PL 35(M), FI-40014, Finland. [3]Cicero Learning, University of Helsinki, P.O. Box 9, FI-00014, Finland. Correspondence and requests for materials should be addressed to H.P. (email: hanna.poikonen@helsinki.fi)
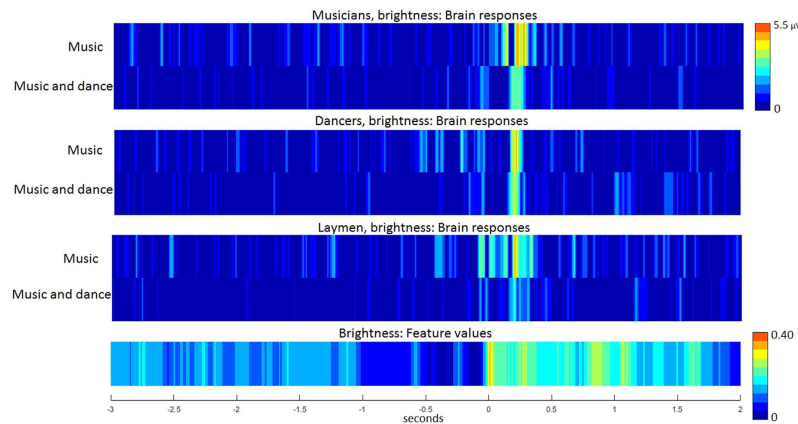
Figure 1. Brain responses of rapid increase in the musical feature brightness in musicians, dancers and laymen during auditory (music) and audio-visual (music and dance) condition. The absolute values of the amplitudes of the EEG epochs are presented over the 16 electrodes in the fronto-central region with the EEG epochs from −3 seconds to +2 seconds from the stimulus onset, and the temporal evolution of the musical feature brightness for the same 5-second time window. The stimulus onset is defined by the end of the Preceding Low-Feature Phase (PLFP) period.

by more simple multimodal stimuli[17,18]. Since professional background in music has been shown to facilitate the brain processes for individual sounds compared to laymen[19,20], we hypothesized that these kinds of changes would also be detected during continuous music listening. Further, the comparison of dancers and musicians may help in defining whether these changes are influenced by personal history in intense listening of music or in active music-making. Indeed, dancers have a different approach to music than musicians - for dancers the music is a tool for the kinesthetic expression whereas for musicians the music is the essence itself.

Results

The musical features under interest evoked auditory brain responses resembling those recorded in traditional ERP paradigms. Figure 1 shows the grand-average ERPs in the auditory and audio-visual conditions of the musical feature brightness for musicians, dancers and laymen. Figure 2 shows the recapitulation of the grand-average ERPs in the auditory and audio-visual conditions of brightness, RMS, zero-crossing rate and spectral flux for musicians, dancers and laymen. Scalp maps of the P50, N100 and P200 responses in the auditory and audio-visual condition of brightness for musicians, dancers and laymen are presented in Fig. 3. Statistical evaluation of the data indicated that most but not all of the P50 and N100 responses differed from the zero baseline while all the P200 responses did (see Table 1 for the t-tests of P50 response and Table 2 of N100 response).

In the repeated measures ANOVA, Group (musicians, dancers, control group) was set as the between-subject factor and Modality (auditory, audio-visual stimulus) and Musical feature (brightness, spectral flux, RMS, zero-crossing rate) were set as the within-subject factors.

P50 response. For the P50 response, neither the amplitude nor the latency showed a significant main effect for the factor Group. For the P50 latency, Modality showed a significant main effect with the Greenhouse-Geisser (GG) adjustment, $F(1, 51) = 8.41$, $pGG = 0.0055$ resulting from the latencies of auditory (mean latency 62.5 ms) and audio-visual stimulus (57.1 ms). For P50 amplitude, Musical feature showed a significant main effect, $F(3, 153) = 8.11$, $p = 0.00020$ (mean amplitude of brightness 1.79 μV, RMS 3.58 μV, spectral flux 2.04 μV, zero-crossing rate 1.57 μV). For P50 amplitude the Group*Musical feature interaction $F(6, 153) = 2.67$, $pGG = 0.026$ was caused by the difference between dancers (2.97 μV) and laymen (1.11 μV), $p = 0.014$, and between dancers and musicians (1.28 μV), $p = 0.030$, in the feature brightness revealed by multiple comparison of Group for the musical feature brightness with the critical value of Bonferroni. In addition, P50 amplitude had a significant Musical feature*-Modality interaction $F(3, 153) = 3.57$, $pGG = 0.037$ rising from the difference of the auditory (1.27 μV) and the audio-visual (2.31 μV) stimulus of brightness, $p = 0.047$ and of zero-crossing rate, $p = 0.0044$, with the amplitudes of 2.71 μV and 0.42 μV, respectively, revealed by multiple comparison of Modality with the critical value of Bonferroni. The amplitudes that did not differ significantly, for the auditory stimulus RMS 3.81 μV and spectral flux 2.31 μV, and for the audio-visual stimulus RMS 3.35 μV and spectral flux 1.77 μV.

N100 response. For the N100 latency the main effects for the factor Modality ($F(2, 51) = 11.35$, $pGG = 0.0014$, auditory (98.4 ms) and audio-visual stimulus (86.3 ms)) and for the factor Musical feature ($F(3, 153) = 5.69$, $pGG = 0.0025$, the mean latency of brightness 97.5 ms, RMS 85.7 ms, spectral flux 88.1 ms, zero-crossing rate 98.1 ms) were significant. For N100 amplitude, the interaction Group*Musical feature was significant, $F(6, 153) = 2.31$, $pGG = 0.046$, rising from the difference between dancers (−2.04 μV) and laymen (−4.69 μV) for the musical feature brightness, $p = 0.023$, revealed by multiple comparison of Group for the musical feature brightness with the critical value of Bonferroni. With the mean amplitude of −4.43 μV, musicians did not differ significantly from the other groups. Also, for the N100 amplitude, the main effects of Modality ($F(1, 51) = 5.85$, $pGG = 0.019$, auditory (−3.17 μV) and audio-visual (−2.41 μV) stimulus) and Musical feature ($F(3, 153) = 14.88$, $pGG = 0.00000014$,
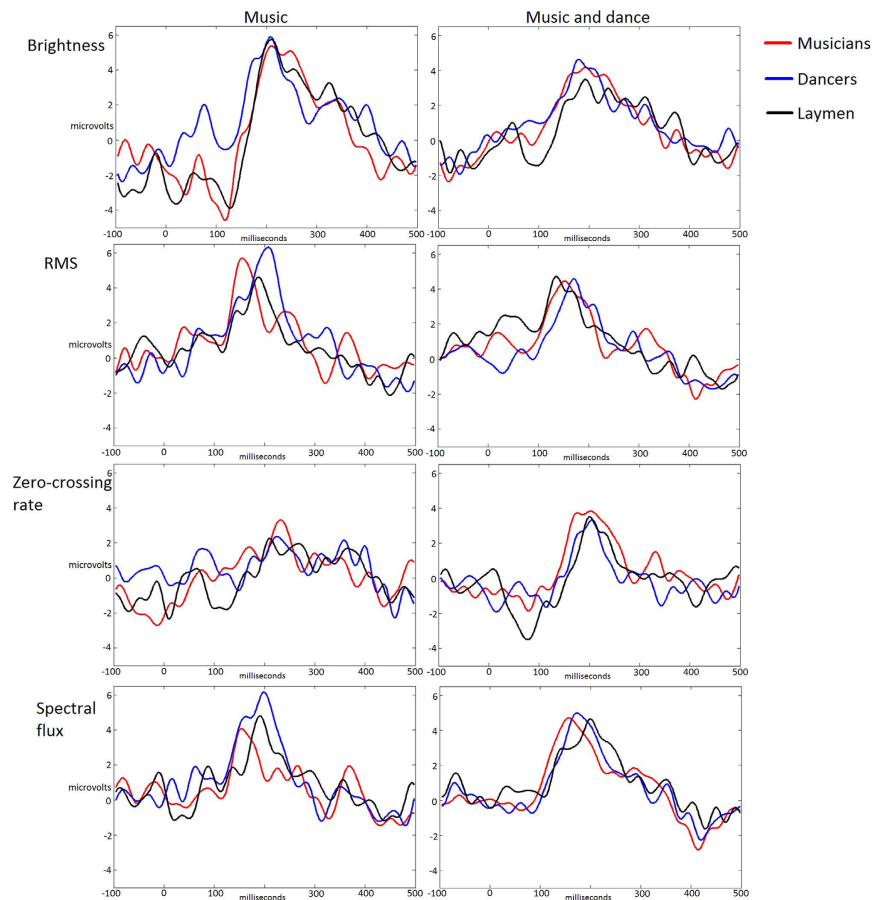
Figure 2. ERPs of the mean value over the averaged signal of 16 electrodes for the rapid changes in the musical features brightness, RMS, zero-crossing rate and spectral flux during the presentation of the auditory stimulus only (music; graphs in the column on the left) and during the stimulus of audiovisual entity (music and dance; graphs in the column on right). In each graph three groups of participants are compared: Musicians, dancers and control group. For brightness, RMS, zero-crossing rate and spectral flux the amount of extracted epochs for each test subject were 9, 8, 8 and 10, respectively, excluding a minimal amount of epochs rejected due to noisy data.

brightness mean $-3.72\,\mu V$, RMS $-1.60\,\mu V$, spectral flux $-1.96\,\mu V$, zero-crossing rate $-3.87\,\mu V$) were significant as well as the interaction of Musical feature*Modality, $F(3, 153) = 8.44$, pGG = 0.00015 caused by the difference of the auditory $(-5.15\,\mu V)$ and the audio-visual $(-2.29\,\mu V)$ stimulus of brightness, $p = 0.000036$, revealed by multiple comparison of Modality with the critical value of Bonferroni. The amplitudes that did not differ significantly, were for the auditory stimulus RMS $-1.90\,\mu V$, spectral flux $-2.13\,\mu V$ and zero-crossing rate $-3.50\,\mu V$, and for the audio-visual stimulus RMS $-1.30\,\mu V$, spectral flux $-1.80\,\mu V$ and zero-crossing rate $-4.25\,\mu V$.

P200 response. For the P200 response, neither the amplitudes nor the latencies differed significantly between the groups. For P200 latency, the main effect of Musical feature ($F(3, 153) = 13.80$, pGG = 0.0000012, the mean latency of brightness 207.7 ms, RMS 177.5 ms, spectral flux 185.1 ms, zero-crossing rate 206.6 ms) and Modality ($F(1, 51) = 6.04$, pGG = 0.017, auditory (200.2 ms) and audio-visual stimulus (188.3)) were significant. For P200 amplitude, the main effect of Musical feature ($F(3, 153) = 5.65$, pGG = 0.0059, the mean latency of brightness $7.33\,\mu V$, RMS $7.08\,\mu V$, spectral flux $6.80\,\mu V$, zero-crossing rate $5.56\,\mu V$) and Modality ($F(1, 51) = 5.63$, pGG = 0.021, auditory $(7.08\,\mu V)$ and audio-visual stimulus $(6.30\,\mu V)$) were significant as well as the Musical feature*Modality interaction ($F(3, 153) = 4.79$, pGG = 0.0056 rising from the difference of the auditory $(8.16\,\mu V)$ and the audio-visual $(6.51\,\mu V)$ stimulus of brightness, $p = 0.0064$ and of RMS, $p = 0.0066$, with the amplitudes of $7.90\,\mu V$ and $6.26\,\mu V$, respectively, revealed by multiple comparison of Modality with the critical value of Bonferroni. The remaining P200 amplitudes, which did not differ significantly between the modalities, were for the auditory stimulus spectral flux $6.87\,\mu V$ and zero-crossing rate $5.40\,\mu V$, and for the audio-visual stimulus spectral flux $6.72\,\mu V$ and zero-crossing rate $5.71\,\mu V$.

Discussion

Our results suggest that preattentive processing of changes in timbral brightness of continuous music is improved in dancers compared to musicians and laymen. In addition, brain responses to fast changes in musical features are suppressed and sped up in dancers, musicians and laymen when music is presented with concordant dance.
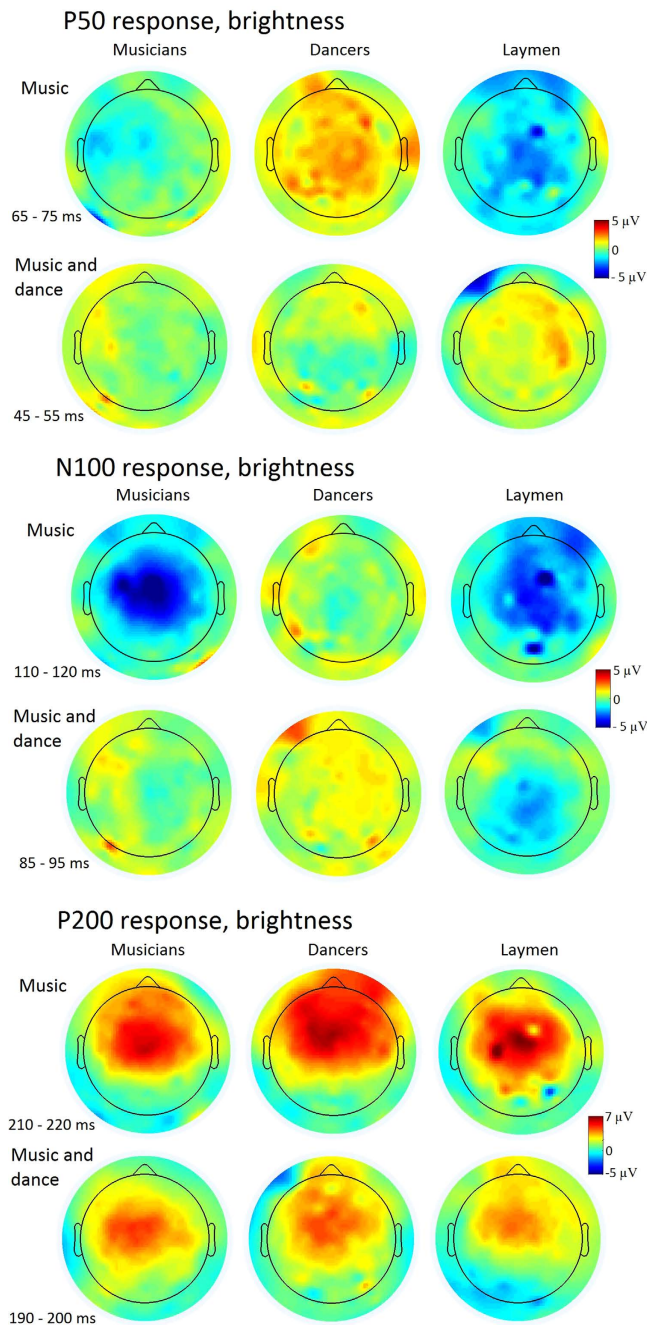
Figure 3. Scalp maps for the P50 (above), N100 (middle) and P200 responses (below) of brightness for musicians, dancers and laymen in auditory (music) and audio-visual (music and dance) condition.

Professional expertise in music can dramatically modulate the auditory processing in the brain[11,20,21]. Our results gained with continuous polyphonic music extend these earlier results obtained by using simple tones and short sound sequences. Also, our results shed light on how individual characters of a complex sound scene are processed in the brain. Indeed, fast and large changes in particular features of natural music evoke ERP responses corresponding to those evoked by simple sounds. Simultaneous presentation of a dance choreography with music makes our paradigm even more unique in ERP research. In the field of multimodal processing, our paradigm is an upgrade to the earlier studies of ecologically valid audio-visual stimuli[17,18].

Following the interdisciplinary trend of brain imaging using natural stimuli in order to meet the demands of ecological validity[10,22–25], the music research with ERPs can be upgraded in this respect as well. In addition to the complexity of the physical sound waves, also the human cognition and emotion become much more versatile with the natural musical stimulus. ERP research is necessary to complement the fMRI research because of their fundamental differences in temporal resolution and in the bioelectric origin of the signal.

t-test	t_{17}	p
Musicians		
Brightness		
Auditory stimulus	0.63	0.54
Auditory-visual stimulus	3.97	0.0010
RMS		
Auditory stimulus	5.71	0.000025
Auditory-visual stimulus	3.43	0.0032
Zero-crossing rate		
Auditory stimulus	2.75	0.014
Auditory-visual stimulus	1.67	0.11
Spectral flux		
Auditory stimulus	3.40	0.0034
Auditory-visual stimulus	1.74	0.10
Dancers		
Brightness		
Auditory stimulus	4.82	0.00016
Auditory-visual stimulus	3.49	0.0028
RMS		
Auditory stimulus	4.38	0.00041
Auditory-visual stimulus	3.45	0.0031
Zero-crossing rate		
Auditory stimulus	6.12	0.000011
Auditory-visual stimulus	0.28	0.79
Spectral flux		
Auditory stimulus	0.77	0.45
Auditory-visual stimulus	2.06	0.055
Laymen		
Brightness		
Auditory stimulus	0.44	0.66
Auditory-visual stimulus	3.91	0.0011
RMS		
Auditory stimulus	4.28	0.00050
Auditory-visual stimulus	3.91	0.0011
Zero-crossing rate		
Auditory stimulus	3.30	0.0043
Auditory-visual stimulus	−0.13	0.90
Spectral flux		
Auditory stimulus	3.09	0.0067
Auditory-visual stimulus	3.61	0.0021

Table 1. P50 response (time window from 30 milliseconds to 90 milliseconds of the stimulus onset). T-tests over the averaged signal of the 16 electrodes in the fronto-central region for musicians, dancers and laymen in the auditory and audio-visual condition of the musical features brightness, RMS, zero-crossing rate and spectral flux.

ERPs in processing multimodal information. In our study, the auditory N100 and P200 responses were suppressed and sped up in dancers, musicians and laymen during the audio-visual stimulus of a dance choreography compared to the unimodal presentation of the music of the choreography. Previously, Stekelenburg and Vroomen showed how the auditory N100 and P200 responses were suppressed and sped up only if the visual stimulus was synchronized with the auditory event and reliably predicted the sound[17]. As stimuli, they used natural human actions such as the pronunciation of a letter or a hand clap. In their study, N100 amplitude decreased when the visual cue reliably predicted the onset of the sound reducing the temporal uncertainty. In contrast, the P200 amplitude decreased when the content of the visual cue and the sound were coherent, such as the pronunciation of the same letter in voice and in the video. Therefore, N100 likely reflects the multisensory integration related to coherent timing of all the unimodal elements whereas P200 is rather related to the associative and semantic coherence of them[17]. Thus, suggested by the results of the earlier studies[17,18,26], dance movement has elements which reliably predict both temporally and associatively fast changes in the musical features reducing the surprise of the sudden change in music. Importantly, neither dancers nor musicians were shown to be more sensitive than laymen to these movement cues suggesting that processes underlying multisensory integration are not modified by the training of music and movement.

t-test	t_{17}	p
Musicians		
Brightness		
Auditory stimulus	−5.82	0.000021
Auditory-visual stimulus	−3.00	0.0089
RMS		
Auditory stimulus	−2.32	0.033
Auditory-visual stimulus	−1.43	0.17
Zero-crossing rate		
Auditory stimulus	−4.30	0.00048
Auditory-visual stimulus	−5.90	0.000017
Spectral flux		
Auditory stimulus	−5.78	0.000022
Auditory-visual stimulus	−1.76	0.096
Dancers		
Brightness		
Auditory stimulus	−2.72	0.015
Auditory-visual stimulus	−1.93	0.070
RMS		
Auditory stimulus	−2.34	0.032
Auditory-visual stimulus	−3.30	0.0042
Zero-crossing rate		
Auditory stimulus	−5.00	0.00011
Auditory-visual stimulus	−4.34	0.00044
Spectral flux		
Auditory stimulus	−1.80	0.089
Auditory-visual stimulus	−3.34	0.0038
Laymen		
Brightness		
Auditory stimulus	−7.12	0.0000017
Auditory-visual stimulus	−4.61	0.00025
RMS		
Auditory stimulus	−3.34	0.0039
Auditory-visual stimulus	−1.05	0.31
Zero-crossing rate		
Auditory stimulus	−5.45	0.000043
Auditory-visual stimulus	−6.13	0.000011
Spectral flux		
Auditory stimulus	−3.52	0.0026
Auditory-visual stimulus	−2.36	0.031

Table 2. N100 response (time window from 50 milliseconds to 150 milliseconds of the stimulus onset). T-tests over the averaged signal of the 16 electrodes in the fronto-central region for musicians, dancers and laymen in the auditory and audio-visual condition of the musical features brightness, RMS, zero-crossing rate and spectral flux.

In the studies of Stekelenburg and Vroomen[17,18] the audio-visual interaction might have facilitated the auditory processing[27] by amplifying the signal intensity in the unimodal sensory cortices[28]. Optionally, the visual cue could evoke sensory gating on the auditory cortex[29] by reducing the novelty and surprise of the sound. The sensory gaiting is shown to suppress P50, N100 and P200 responses in a paired-sound paradigm[30,31]. Professional musicians have a reduced paired-sound P50 suppression[32], yet their N100 is reduced in a manner comparable to that of controls.

Musical features and group differences. In our study, early cortical processing of music differed in dancers compared to both musicians and laymen. P50 to brightness was larger in dancers than in musicians and laymen. In contrast to the P50, the N100 to brightness in laymen was larger than in dancers, which might be a counter effect of the strong P50 of dancers. In the P200 response the group differences are already diminished.

The processes involved in movement-related imagination could be more active in dancers during their listening to music[33,34], possibly increasing the sensitivity to the fast changes in brightness. Optionally, intense and versatile physical training with music could improve cerebral processes which enhance the early reaction to these changes. Fine temporal changes in music are essential for dancers to create precise rhythmical movement which

could, after years of exposure, lead to sensitization in the early auditory processes without concomitant sensitization of the longer-latency responses. Indeed, all large changes of the musical features in the millisecond-scale occur with respect to the temporal structure of music. In addition, pitch, which is an important but not the only factor for brightness, and temporal structure are suggested to be largely integrated in auditory-motor transformations[35].

Functional integration in the cortico-basal ganglia loops that govern motor control and integration is suggested to be enhanced in dancers compared to laymen[36]. Basal ganglia project not only to the motor cortex but are highly interconnected with widespread areas on the cerebral cortex. Thus, they also play an important role in non-motor cognitive and sensory functions and in a wide range of learning challenges[37]. In vision, cortico-basal ganglia loop participates in action selection in response for a visual stimulus[38]. The auditory cortico-basal ganglia network is less studied but there is evidence for a similar network as in visual domain[39]. Cortico-basal ganglia loop is crucial in the voluntary attentive movements whereas basal ganglia-brainstem loop is involved in the involuntary movements, such as breathing, swallowing and maintaining the body posture. In Parkinson's and Huntington's diseases the function of both cortico-basal ganglia loop and basal ganglia-brainstem loop is suggested to be violated[40]. The whole-body movement training of professional dancers seems to modify the cortico-basal ganglia network[36]. When compared to laymen, musicians show modulation on the cortical areas related to sound and movement, especially on the dominant hand of the instrument, and increased connectivity strength in motor-related regions[41–43]. However, it might be the improved cortico-basal ganglia loop of dancers which plays a key role in the enhancement of the preattentive auditory processing of dancers. Similarly to sportsmen, whose motor-related brain areas are sensitized to sports sounds[44], auditory-motor processes of dancers may be sensitized to musical cues such as rapid changes in brightness. Furthermore, continuous music, which is generally used in dance training, might be a unique stimulus in enhancing top-down controlling of the basal ganglia to the auditory cortex in dancers.

The dance style, in which each dancer was specialized, may have an influence on the early auditory processing of changes in the musical features due to familiarity with the composition or with the musical genre in general[45]. Such specialization of brain functions and structure has previously been shown in musicians[21,46,47]. Also, a strong background in dance improvisation, and thus possibly enhanced movement imagery during listening to music even without an association to a learned choreography, may have an influence to the preattentive auditory processing by augmenting the sensitivity to the musical cues. The composition used in our study was played with string instruments with occasional percussion. Thus, the musicians specialized in string instruments, might have had enhanced brain responses to the fast changes in the musical features compared to the musicians with biography in non-string instruments[48].

By means of non-musical stimuli, it could be studied whether this sensitization is related to the musical sounds only or to the auditory information in general. However, it is increasingly common to use non-musical sounds, such as environmental sounds or digital sounds, in the creation of contemporary dance. Familiarity with the composition or with the dance style used in our study could modify the early auditory processing[33,49]. Our participants had a versatile background in dance. Thus, a follow-up study in which expertise in specific dance styles are compared, would be important to analyze the effect of familiarity of sound space and of movement language to the early auditory responses.

Musical features and ERPs evoked by unimodal vs. bimodal stimuli. The musical features were processed differently between the groups of participants as well as between the sensory modalities: During audio-visual presentation of a dance piece, N100 and P200 of brightness and P200 of RMS are attenuated in dancers, musicians and laymen when compared to the auditory presentation. Similarly to our earlier study[16], the musical feature brightness evoked the strongest ERP responses. Thus, our results suggest that the brain is tuned better to detect the changes in timbral brightness rather than the changes in intensity, harmony or the musical dynamics in general reflected by RMS, zero-crossing rate and spectral flux, respectively. Interestingly, the preattentive P50 response of zero-crossing rate is suppressed but that of brightness enhanced during the audio-visual stimulus when compared to the auditory one. The increased P50 response of brightness is contrary to the results gained with multimodal auditory N100 and P200 responses[17,18]. Indeed, the N100 and P200 amplitudes of brightness are suppressed. Possibly, the dance movement anticipates changes in timbral brightness both temporally and associatively. In addition, the intensity-related RMS evokes a suppressed P200 response during the audio-visual stimulus, suggesting that the dance movement predicts associative rather than temporal changes in the intensity of the sound.

Our results propose that long-term activities with music sensitizes the sensory auditory processes despite the music not being produced by oneself. Further research is needed to discover whether this sensitization is due to increased anticipation, attention or some other factors possibly related to the coupling of the auditory and motor systems as discussed above. We did not find differences between the participating groups in the suppression of the ERP responses evoked by a multimodal presentation. In contrast, musical features seem to be processed in the brain along diverging pathways producing variability in the ERP responses of the study groups and of the sensory modalities.

Conclusions

Our P50, N100 and P200 brain responses suggest that continuous overlapping auditory stimulus such as natural music is processed in the brain at least partly similarly to the simplified sounds traditionally used in ERP research. In contrast, Hasson *et al.* report that, in the visual modality, the brain processes visual stimuli differently in a more ecological setting than in conventional controlled settings[50]. Importantly, the musical features of our study are classified as lower level features evoking bottom-up neural processes. Due to the novelty of the current test paradigm, the musical stimulus could not be optimized beforehand. To evoke clear ERP components in future

studies, we recommend to use music which has large changes in the low-level musical features within a short time window. With a replication study of fMRI, Burunat et al.[23] showed constant results in the processing of low-level features whereas the results in the processing of high-level features were not stable. High-level features related to rhythm and melody contour require context-dependent information and evoke top-down processes over a longer time-span[10,23]. In addition, the processing of such high-level features may be more sensitive to the state and traits of the listeners, as well as of their background in music[23]. While we analyzed only the post-stimulus cortical processing within a relatively short time window, both further processing of these low-level features as well as the processing of higher level musical features may be different to the conventional simplified sound stimuli. However, our results of cortical sound processing indicate that natural music evokes stronger brain responses than the various traditional simplified stimuli. In fact, with single sounds it has already been shown that with spectrally rich sounds and synthetized sounds mimicking natural instrumental sounds, the brain responses are larger than with pure sinusoidal tones[51–53]. The brain seems to be more sensitive to the stimuli from the real-life environment. Therefore, natural stimuli of continuous music are ideal for applied studies, for example in estimating the depth of coma[54], the prognosis of vegetative state[55], comparing the efficiency of medical treatment in psychotic disorders[56] and estimating the efficiency of expressive therapies such as music and dance/movement therapy.

Methods

Participants. 20 professional musicians, 20 professional dancers and 20 people without a professional background in either music or dance participated in the experiment. However, two participants from each group were left out from the data analysis since their EEG data lacked several electrodes around the brain area of our interest. Thus, in the groups of musicians and dancers there were 13 female and 5 male participants and in the control group 12 female and 6 male participants. The background of the participants was screened by a questionnaire of music and dance related to both professional and every-day level. Professional background of musicians varied from singing to various instruments, such as piano, violin or saxophone. The professional background of dancers was versatile from ballet and contemporary dance to street dance. Several musicians reported expertise in more than one instrument and several dancers in more than one dance style. The age of the participants ranged from 21 to 31 years (25.4 on average) among musicians, from 23 to 40 years (29.1 on average) among dancers and from 20 to 37 years (25.3 on average) among laymen. Two participants in each of three groups included in the data analysis were left-handed. No participants reported hearing loss or history of neurological illnesses. All subjects gave written informed consent. The experiment protocol was conducted in accordance with the Declaration of Helsinki and approved by the University of Helsinki review board in the humanities and social and behavioural sciences.

Stimuli. Long excerpts of Carmen composed by Bizet-Shchedrin were used as stimuli. The version of the composition Carmen was performed by Moscow Virtuosi Chamber Orchestra and published by Melodiya, Moscow 1987. Many participants reported being familiar to the composition. The total length of the musical stimulus was approximately 15 minutes, which was cut to 20 trials, the duration of each trial being between 15 and 63 seconds (44.5 seconds on average). Music without visual stimulus, silent dance as well as music and dance as an audiovisual entity were presented to the participants. During the presentation of music only, the participants were advised to listen to the music eyes open although there was no visual stimulus on the screen. The excerpts were chosen from the composition based on their musical and emotional versatility. Some excerpts were musically full and complex whereas the other parts were monotonic and simple. Also, the emotional content varied significantly, some excerpts transmitting a joyful atmosphere, others anger or devastating sadness. The dance choreography presented was based on the contemporary ballet choreographed by Mats Ek. However, the contemporary dancer who performed the dance excerpts for our research purposes, had an artistic freedom to create solo versions to suit her own expression. Thus, the dance choreography was not familiar to any of the participants.

Equipment and procedure. The stimuli were presented to the participants with the Presentation 14.0 program. Each set of trials contained 20 excerpts of the same sensory modality/modalities and these sets were presented in a random order via a monitor and headphones with the intensity of 50 decibels above the individually determined hearing threshold. Randomization of the presentation order of the stimuli is a standard procedure in experimental psychology which is suggested to reduce the influence of individual differences in other simultaneous cognitive processes. The distance of the monitor from the participant was 110 cm. The participants were advised to listen to the music and watch the dance video as still as possible. The playback of each trial was launched by the researcher. From time to time, between the stimuli, the researcher had a short conversation with the participant via microphone to make sure the participant felt comfortable during the test procedure. The total length of the experiment material was 60 minutes. With pauses and conversations based on the individual needs of each participant, the whole test session lasted about 70–80 minutes.

The data were recorded using BioSemi bioactive electrode caps with 128 EEG channels and 4 external electrodes placed at the tip of the nose, left and right mastoids and under the right eye. The offsets of the active electrodes were kept below 25 millivolts in the beginning of the measurement and the data were collected with a sampling rate of 1024 Hz. The beginning and the end of each musical piece was marked with a trigger into the EEG data.

Feature extraction with MIRtoolbox. We used MIRtoolbox (version 1.3.1) to computationally extract the musical features. MIRtoolbox is a set of MATLAB functions designed for the processing of audio files[57] and is used for the extraction of different musical features related to various musical dimensions identified in psychoacoustics and sound engineering as well as traditionally defined in music theory. In addition to the dimensions

of dynamics, loudness, rhythm, timbre and pitch also high-level features related to meter and tonality, among others, can be processed. Low-level features are those that are perceived in a bottom-up fashion without a need for domain-specific knowledge. For instance, loudness, pitch and timbre processing automatically recruit sensory mechanisms, and are performed rapidly in very short-time spans. On the other hand, rhythm and melody contour encapsulate context-dependent aspects of music and recruit perceptual processes that are top-down in nature, and require a longer time-span. Since our interest was to study the early auditory processing evoked by fast changes in music, we chose to analyze the following low-level features: Brightness, root mean square (RMS) amplitude, zero-crossing rate and spectral flux. Each one of these features captures a different perceptual element in music.

Brightness was computed as the amount of spectral energy above a threshold value fixed by default in MIRtoolbox at 1500 Hz for each analysis window[57]. Therefore, high values in brightness mean that a high percentage of the spectral energy is concentrated in the higher end of the frequency spectrum. Thus, brightness is influenced by both the pitch of the sound and the characteristic spectrum of the instrument with which the sound is created. Root mean square (RMS) is related to the dynamics of the song and defined as the root average of the square of the amplitude[57]. Louder sounds have high RMS values whereas quieter ones have low RMS values. The zero-crossing rate, known to be an indicator of noisiness, is estimated by counting the number of times the audio waveform crosses the temporal axis[57]. Higher zero-crossing rate indicates that there is more noise in the audio frame under consideration. The noise measured by zero-crossing rate refers to noise as opposed to harmonic sounds rather than to noise as distortion of clean signal. Spectral flux represents the Euclidian distance between the spectral distributions of successive frames[57]. If there is a large amount of variation in spectral distribution between two successive frames, the flux has high values. Spectral flux curves exhibit peaks at transition between successive notes or chords. These musical features were obtained by employing short-time analysis using a 25-millisecond window with a 50% overlap, which is in the order of the commonly used standard window length in the field of Music Information Retrieval (MIR)[58]. Overlapping of windows is recommended in the analysis of musical features to detect fast changes in the features and their possible inactive periods with a precise time resolution.

Preprocessing. The EEG data of all the participants were first preprocessed with EEGLAB[59] (version 9.0.2.2b). The external electrodes of the left and the right mastoid were set as a reference. The data were high-pass filtered at 1 Hz and low-pass filtered at 30 Hz.

Setting the Triggers. The triggers related to the musical features extracted with MIRtoolbox were added to the preprocessed EEG data. In continuous speech, the best ERP-related results are gained when the triggers are set into the beginning of the word[60,61]. Long inter-stimulus interval is shown to increase the amplitude of the N100 response[62]. Additionally, strong stimulus intensity has been shown to enhance ERP responses[63,64]. Previous knowledge from the individual sound processing was utilized in our study of continuous music, in which the individual sounds are connected to each other in an overlapping and dynamic manner.

Approximately 10 triggers per each feature were set. We used the same MATLAB algorithm for the search of time points with rapid increase of a musical feature as was used in the study of Poikonen *et al.* for defining the time points of the trigger[16]. The algorithm was tuned using specific parameter values adapted to each musical feature. In our study, the time period with low feature values preceding the rapid increase in the value of the musical feature corresponds to the inter-stimulus interval (ISI) of the previous literature. However, in our study, the intervals are not between individual stimuli anymore nor are the intervals completely silent, and thus this ISI-type of period is called the Preceding Low-Feature Phase (PLFP) in this paper.

The length of the PLFP was modified and the rapid increase was required to exceed a value called magnitude of the rapid increase (MoRI). The mean values of all the segments of each one of the 20 sound excerpts and each musical feature were calculated and the magnitude of the change from the lower threshold value V_{n-} to the higher threshold value V_{n+} was defined based on the mean value (MV_n) in each particular sound excerpt for each musical feature. The largest changes in the musical features were when the V_{n-} remained under -20% of MV_n and V_{n+} increased above $+20\%$ of MV_n. The smallest changes in the musical features were when the V_{n-} remained under -15% of MV_n and V_{n+} increased above $+15\%$ of MV_n. Valid triggers were preceded by a PLFP whose magnitude did not exceed the lower threshold V_{n-}. The length of PLFP with values below V_{n-} was 625 milliseconds minimum and 1 second maximum. In all cases, valid triggers had an increase phase that lasted less than 75 milliseconds during which the feature value increased from V_{n-} to V_{n+}.

Procedure of the ERP analyses. After adding the triggers into the preprocessed data, the data were treated with Independent Component Analysis (ICA) decomposition with the runica algorithm of EEGLAB[59] to detect and remove artifacts related to eye movements and blinks. ICA decomposition gives as many spatial signal source components as there are channels in the EEG data. Thus, the amount of components was 128 in 22 participants. In the remaining 32 participants, several noisy channels each were removed in preprocessing and therefore less than 128 ICA components were decomposed in them. Typically, 1 to 5 ICA components related to the eye artifacts were removed. Noisy EEG data channels of the abovementioned 32 participants were interpolated. The average number of interpolated channels among these 32 participants was 3.1 channels, the actual number of interpolated channels varying from one per person up to 8 per person. The continuous EEG data were separated into epochs according to the triggers. The epochs started 500 milliseconds before the trigger and ended 1000 milliseconds after the trigger. The baseline was defined according to the 500-millisecond time period before the trigger. To double check the removal of the eye artifacts, the epochs with amplitudes above $\pm$ 100 microvolts were rejected.

The statistical analyses were conducted with MATLAB version R2015b. In the statistical analysis, 16 electrodes (B1, B21, B22, B32, C1, C2, C11, C22, C23, C24, D1, D2, D13, D14, D15 and D18 of the 128-channel BioSemi

EEG gap) were averaged as one signal. Cz was not included among the averaged channels because it was not recorded from five participants due to a broken electrode. Each participant had only 8–10 trials for each musical feature in each sensory modality due to the need to minimize the duration of an experimental session, which was already 60 minutes long. To improve the S/N ratio of the signal, we averaged the signal over several electrodes.

According to the Shapiro-Wilk test, 75.0% of the P50 responses, 87.5% of the N100 responses and 75.0% of the P200 responses were normally distributed. Thus, the repeated measures ANOVA was used in the statistical analysis.

The repeated measures ANOVA was calculated for both amplitude and latency of the P50, N100 and P200 responses. A time window from 30 ms to 90 ms was chosen for the statistical analyses of the P50 response, a time window from 50 ms to 150 ms for the N100 response and a time window from 100 ms to 280 ms for the P200 response.

References

1. Masataka, N. & Perlovsky, L. the efficacy of musical emotions provoked by Mozart's music for the reconciliation of cognitive dissonance. *Sci. Rep.* **2,** 694 (2012).
2. Masataka, N. & Perlovsky, L. Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. *Sci. Rep.* **3,** 2028 (2013).
3. Sutton, S., Braren, M., Zubin, J. & John, E. R. Evoked-potential correlates of stimulus uncertainty. *Science* **150,** 1187–1188 (1965).
4. Brattico, E., Tervaniemi, M., Näätänen, R. & Peretz, I. Musical scale properties are automatically processed in the human auditory cortex. *Brain Res.* **1117(1),** 162–174 (2006).
5. Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R. & Pantev, C. Automatic encoding of polyphonic melodies in musicians and onmusicians. *J. Cogn. Neurosci.* **17(10),** 1578–1592 (2005).
6. Steinbeis, N., Koelsch, S. & Sloboda, J. A. The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *J. Cognitive Neurosci.* **18(8),** 1380–1393 (2006).
7. Zatorre, R. J. & Halpern, A. R. Mental concerts: Musical imagery and auditory cortex. *Neuron* **47(7),** 9–12 (2005).
8. Berkowitz, A. L. & Ansari, A. Expertise-related deactivation of the right temporoparietal junction during musical improvisation. *NeuroImage* **49,** 712–719 (2010).
9. McPherson, M. J., Barrett, F. S., Lopez-Gonzalez, M., Jiradeivong, P. & Limb, C. J. Emotional intent modulates the neural substrates of creativity: An fMRI study of emotional targeted improvisation in jazz musicians. *Sci. Rep.* **6,** 18460 (2016).
10. Alluri, V. *et al.* Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage* **59(4),** 3677–3689 (2012).
11. Herholz, S. C. & Zatorre, R. J. Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron* **76(3),** 486–502 (2012).
12. Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19(5),** 893–906 (2007).
13. Brown, S., Martinez, M. J. & Parsons, L. M. Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport* **15(13),** 2033–2037 (2004).
14. Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A. & Zatorre, R. J. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* **14,** 257–262 (2011).
15. Koelsch, S. Towards neural basis of music perception – A review and updated model. *Front. Psychol.* **2,** 110 (2011).
16. Poikonen, H. *et al.* Event related brain responses while listening to entire pieces of music. *Neuroscience* **312,** 58–73 (2016).
17. Stekelenburg, J. J. & Vroomen, J. Neural correlates of multisensory integration of ecologically valid audiovisual events. *J. Cogn. Neurosci.* **19(12),** 1964–1973 (2007).
18. Vroomen, J. & Stekelenburg, J. J. Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J. Cogn. Neurosci.* **22(7),** 1583–1596 (2009).
19. Seppänen, M., Hämäläinen, J., Pesonen, A.-K. & Tervaniemi, M. 2012. Music training enhances rapid plasticity of N1 and P2 source activation for unattended sounds. *Front. Hum. Neurosci.* **6,** 43 (2012).
20. Pantev, C. & Herholz, S. C. Plasticity of the human auditory cortex related to musical training. *Neurosci. Biobehav. R.* **25(10),** 2140–2154 (2011).
21. Tervaniemi, M. Musicians – same or different? *Ann. NY Acad. Sci.* **1169(1),** 151–156 (2009).
22. Alluri, V. *et al.* From Vivaldi to Beatles and back: Predicting lateralized brain responses to music. *NeuroImage* **83(12),** 627–636 (2013).
23. Burunat, I. *et al.* The reliability of continuous brain responses during naturalistic listening to music. *NeuroImage* **24,** 224–231 (2016).
24. Salimpoor, V. N. *et al.* Interactions between nucleus accumbens and auditory cortices predict music reward value. *Science* **340,** 216–219 (2013).
25. Wilkins, R. W., Hodges, D. A., Laurienti, P. J., Steen, M. & Burdette, J. H. Network science and the effects of music preference on functional brain connectivity: From Beethoven to Eminem. *Sci. Rep.* **4,** 6130 (2014).
26. Guo, S. & Koelsch, S. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Sci. Rep.* **6,** 19064 (2016).
27. van Wassenhove, V., Grant, K. W. & Poeppel, D. Visual speech speeds up the neural processing of auditory speech. *PNAS* **102,** 1181–1186 (2005).
28. Calvert, G. A. *et al.* Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport* **10(12),** 2619–2623 (1999).
29. Oray, S., Lu, Z. L. & Dawson, M. E. Modification of sudden onset auditory ERP by involuntary attention to visual stimuli. *Int. J. Psychophysiol.* **43(3),** 213–224 (2002).
30. Fuerst, D. R., Gallinat, J. & Boutros, N. N. Range of sensory gating values and test–retest reliability in normal subjects. *Psychophysiology* **44,** 620–626 (2007).
31. Rentzsch, J., Jockers-Scherübl, M. C., Boutros, N. N. & Gallinat, J. Test–retest reliability of P50, N100 and P200 auditory sensory gating in healthy subjects. *Int. J. Psychophysiol.* **67,** 81–90 (2008).
32. Kizkin, S., Karlidag, R., Ozcan, C. & Ozisik, H. I. Reduced P50 auditory sensory gating response in professional musicians. *Brain Cogn.* **61,** 249–254 (2006).
33. Olshansky, M. P., Bar, R. J., Fogarty, M. & DeSouza, J. F. X. Supplementary motor area and primary auditory cortex activation in an expert break-dancer during the kinesthetic motor imagery of dance to music. *Neurocase* **21(5),** 607–617 (2015).
34. Bar, J. & DeSouza, J. F. X. Tracking plasticity: Effects of long-term rehearsal in expert dancers encoding music to movement. *PLoS ONE* **11,** e0147731 (2016).
35. Brown, R. M. *et al.* Repetition suppression in auditory–motor regions to pitch and temporal structure in music. *J. Cogn. Neurosci.* **25(2),** 313–328 (2013).
36. Li, G. *et al.* Identifying enhanced cortico-basal ganglia loops associated with prolonged dance training. *Sci. Rep.* **5,** 10271 (2015).
37. Middleton, F. A. & Strick, P. L. Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain Cogn.* **42,** 183–200 (2000).

38. Seger, C. A. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neurosci. Biobehav. Rev.* **32(2),** 265–278 (2008).
39. Geiser, E., Notter, M. & Gabrieli, J. D. E. A corticostriatal neural system enhances auditory perception through temporal context processing. *J. Neurosci.* **2(18),** 6177–6182 (2012).
40. Takakusaki, K., Saitoh, K., Harada, H. & Kashiwayanagi, M. Role of basal ganglia–brainstem pathways in the control of motor behaviors. *Neurosci. Res.* **50,** 137–151 (2004).
41. Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B. & Taub, E. Increased cortical representation of the fingers of the left hand in string players. *Science* **270,** 305–307 (1995).
42. Li, J. *et al.* Probabilistic diffusion tractography reveals improvement of structural network in musicians. *PLoS One* **9(8),** e105508 (2014).
43. Zhang, L., Peng, W., Chen, J. & Hu, L. Electrophysiological evidences demonstrating differences in brain functions between nonmusicians and musicians. *Sci. Rep.* **5,** 13796 (2015).
44. Woods, E. A., Hernandez, A. E., Wagner, V. E. & Beilock, S. L. Expert athletes activate somatosensory and motor planning regions of the brain when passively listening to familiar sports sounds. *Brain Cogn.* **87,** 122–133 (2014).
45. Jacobsen, T., Schröger, E., Winkler, I. & Horvath, J. Familiarity affects the processing of task-irrelevant auditory deviance. *J. Cogn. Neurosci.* **17,** 1704–1713 (2005).
46. Vuust, P., Brattico, E., Seppänen, M., Näätänen, R. & Tervaniemi, M. The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia* **50(7),** 1432–1443 (2012).
47. Tervaniemi, M., Janhunen, L., Kruck, S., Putkinen, V. & Huotilainen, M. Auditory profiles of classical, jazz, and rock musicians: Genre-specific sensitivity to musical sound features. *Front. Psychol.* **6,** 1900 (2016).
48. Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B. & Wong, P. C. M. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* **30,** 267–275 (2009).
49. Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. & Haggard, P. Action observation and acquired motor skills: An FMRI study with expert dancers. *Cereb. Cortex* **15(8),** 1243–1249 (2005).
50. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303,** 1634–1640 (2004).
51. Tervaniemi, M. *et al.* Harmonic partials facilitate pitch discrimination in humans: electrophysiological and behavioral evidence. *Neurosci. Lett.* **279,** 29–32 (2000).
52. Tervaniemi, M., Schröder, E., Saher, M. & Näätänen, R. Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans – a mismatch negativity study. *Neurosci. Lett.* **290,** 66–70 (2000).
53. Meyer, M., Baumann, S. & Jäncke, L. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *NeuroImage* **32,** 1510–1523 (2006).
54. Fischer, C., Dailler, F. & Morlet, D. Novelty P3 elicited by the subject's own name in comatose patients. *Clin. Neurophysiol.* **119,** 2224–2230 (2008).
55. O'Kelly, J. *et al.* Neurophysiological and behavioral responses to music therapy in vegetative and minimally conscious states. *Front. Hum. Neurosci.* **12(7),** 884 (2013).
56. Adler, L. E. *et al.* Varied effects of atypical neuroleptics on P50 auditory gating in schizophrenia patients. *Am. J. Psychiatry* **161,** 1822–1828 (2004).
57. Lartillot, O. & Toiviainen, P. A Matlab toolbox for musical feature extraction from audio. International Conference on Digital Audio Effects, Bordeaux (2007).
58. Tzanetakis, G. & Cook, P. Music genre classification of audio signals. *Proc. IEEE T. Acoustic Speech* **10,** 293–302 (2002).
59. Delorme, A. & Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Meth.* **134(1),** 9–21 (2004).
60. Teder, W., Alho, K., Reinikainen, K. & Näätänen, R. Interstimulus interval and the selective-attention effect on auditory ERPs: "N1 enhancement" versus processing negativity. *Psychophysiology* **30(1),** 71–81 (1993).
61. Sambeth, A., Ruohio, K., Alku, P., Fellman, V. & Huotilainen, M. Sleeping newborns extract prosody from continuous speech. *Clin. Neurophysiol.* **119(2),** 332–341 (2008).
62. Polich, J., Aung, M. & Dalessio, D. J. Long-latency auditory evoked potentials: Intensity, inter-stimulus interval and habituation. *Pavlovian J. Biol. Sci.* **23(1),** 35–40 (1987).
63. Picton, T. W., Woods, D. L., Baribeau-Braun, J. & Healey, T. M. G. Evoked potential audiometry. *J. Otolaryngol.* **6(2),** 90–116 (1977).
64. Polich, J., Ellerson, P. C. & Cohen, J. P300, stimulus intensity, modality and probability. *Int. J. Psychophysiol.* **23(1),** 55–62 (1996).

Acknowledgements

Author Contributions

H.P., P.T. and M.T. conceived and conducted the experiment, H.P. analyzed the results. H.P., P.T. and M.T. wrote the main manuscript text and H.P. prepared Figures 1–4 and Tables 1 and 2. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Poikonen, H. *et al.* Early auditory processing in musicians and dancers during a contemporary dance piece. *Sci. Rep.* **6**, 33056; doi: 10.1038/srep33056 (2016).