Methods in Neuromusicology: Principles, Trends, Examples and the Pros and Cons

Christiane Neuhaus

Abstract Neuromusicology, also known as the Cognitive Neuroscience of Music, is a modern discipline devoted to the measurement of real-time processes in the human brain while perceiving and producing sound. Research topics range from acoustic feature processing and listening to melodies to composition and music performance. Before designing an experiment, researchers might find it helpful to be informed about the efficiency of methods and their pros and cons. The chapter at hand gives an overview of several methods used in the neurosciences with a special emphasis on their principles, constraints and fields of application. The focus is on transcranial magnetic stimulation (TMS), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG) and on event-related potentials (ERP). The reader will also become acquainted with trends and recent developments towards whole-brain analyses and real life studies based on the idea to improve ecological validity.

1 Introduction

Neuromusicology, also termed the 'Cognitive Neuroscience of Music', is a modern discipline that came into existence through its methods. It is still not clear whether it belongs to cognitive neuroscience as a sort of 'parent discipline', to empirical musicology or whether it has a status of its own. Any neuroscience method enables researchers to measure the brain's physiological processes in real-time, thus giving insight into the task-related or spontaneous functionings of the human brain without requiring any verbal or behavioral type of response.

The chapter gives an overview about the most frequent methods used in neuromusicology. I will particularly focus on transcranial magnetic stimulation (TMS), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG) and on event-related potentials (ERP). These

C. Neuhaus (✉)
Institute of Systematic Musicology, University of Hamburg, Hamburg, Germany
e-mail: christiane.neuhaus@uni-hamburg.de

types of research methods are all non-invasive, i.e. neurosurgical interventions are circumvented. For each method I will weigh up the pros and cons and give examples of application.

Since results obtained with these techniques essentially differ in nature, it is necessary to first classify methods according to type: Data achieved with EEG and ERP belong to the class of 'bioelectric potentials'. Those achieved with fMRI and PET belong to the group of 'neuroimaging data'. Results obtained with EEG and ERP have shifts of intracellular currents as their common source (and starting point), and bioelectric potentials on the head's surface are the final output (for details regarding electrogenesis, see Sect. 3). Thus, by using EEG and ERP, neural activity can *directly* be measured. Scans obtained with fMRI and PET, by contrast, reveal changes in energy consumption (mostly oxygen) in well-defined regions of the brain. These types of neuroimaging techniques therefore point to neural activity in a merely *indirect* way.

Jäncke [13, 14] addressed this issue of method classification in a slightly different manner by using the exact reference to anatomical structures as a specifying criterion. According to that, neuroimaging methods like fMRI and PET allow a precise assignment of physiological processes to neuroanatomical structures, whereas bioelectrical methods like EEG or ERP do not.

1.1 Transcranial Magnetic Stimulation: How Does It Work?

Let me start with some essential remarks on transcranial magnetic stimulation (TMS). In comparison to the methods mentioned above TMS is a sort of exception: It enables researchers to draw certain *causality-based* conclusions, i.e. to precisely relate 'cause' and 'effect'. In contrast to that, fMRI, ERP and other conventional neuroscience methods, although quite popular, allow only mutual relationships, or *correlative conjunctions,* which means that coincidences cannot be ruled out.

How does TMS work? Transcranial magnetic stimulation modifies the excitability of nerve tissue so that cortical processes may either be accelerated or severely be inhibited up to virtual lesions between 10 and 30 min length. The underlying principle is electromagnetic induction: A magnetic field is temporarily built up in orthogonal direction to the plane of a stimulation coil that is placed 2 cm above the head. This way, an electric current is induced in the small cortical regions underneath while tissue resistances of skin, skull and dura can be disregarded [13]. It depends on the coil's stimulation frequency whether processes speed up or slow down: A pulse series with repetition rates of 5 Hz or higher (repetitive TMS; rTMS) may cause facilitation through lowering the excitation threshold, whereas a pulse series with repetition rates of 1 Hz maximum (same for 10-ms intense single shots) may have the opposite effect and provoke inhibition by suppressing the intracellular current flows [13].

Note that, in principle, TMS does not belong to the class of 'neuroimaging or visualization methods'. Instead, TMS modulates pure neurophysiological (excitatory and inhibitory) activity, observable either through a decelerated or an accelerated overt behavior [13]. However, to complete the data, researchers often decide for a combined approach using TMS as well as fMRI (or ERP) for measuring purposes (e.g. [2]).

Things are more complicated regarding the auditory cortex: During stimulation, abrupt electromagnetic forces make the TMS coil produce noise, causing severe, disruptive effects, in particular when early auditory processes are investigated. Jäncke [13] reported that old apparatus produce sharp coil noise between 120 and 130 dB intensity (which is near the auditory pain threshold), making the use of earplugs necessary. In a combined high-resolution ERP-TMS study Tiitinen et al. [57] tested TMS noise of three volume intensities (80, 90 and 100 dB SPL) for its distorting effects. The study shows that coil clicks alone evoked auditory brain responses with a maximum amplitude at 170 ms post stimulus-onset. Furthermore, clicks seem to interact with a series of simultaneously presented pure tones serving as target stimuli (1000 Hz, 85 dB SPL, 50 ms duration, inattentively processed while reading a book), leading to attenuated brain responses in terms of the target tones. Due to these possible effects of contamination, TMS is less often considered as the method of choice whenever pure listening has to be investigated. However, by using elegant paradigms some difficulties in auditory rTMS research can be avoided: Andoh and Zatorre [2], for example, approached this issue by disentangling the time span between coil stimulation and subsequent testing: Two types of rTMS sequences were initially presented off-line in that the auditory cortex was first stimulated with trains of one pulse per second during the first 10 min, and then with a volley of ten pulses per second (followed by an ITI of 10 s) over the next 26 min, totaling up to 1200 (2 × 600) pulses (Fig. 1). Immediately after stimulation, a melody discrimination task had to be solved, showing interesting sex- and time-related results depending on the type of stimulation: Female participants (8 males, 8 females) significantly accelerated their recognition performance after stimulation with 10 Hz-rTMS sequences (RT became shorter), whereas male participants showed the opposite (RT became longer). In the second half of testing, the situation went into reverse, and for the female group, processing slowed down again.

2 Functional Magnetic Resonance Imaging: Basic Principles and Image Acquisition

Let me move on to the conventional types of neuroscience methods. I will first take a closer look at the principles of functional magnetic resonance imaging (fMRI). Since the underlying mechanisms are quite complex, I will omit the main parts of MR physics here and restrict myself to those aspects I consider relevant for
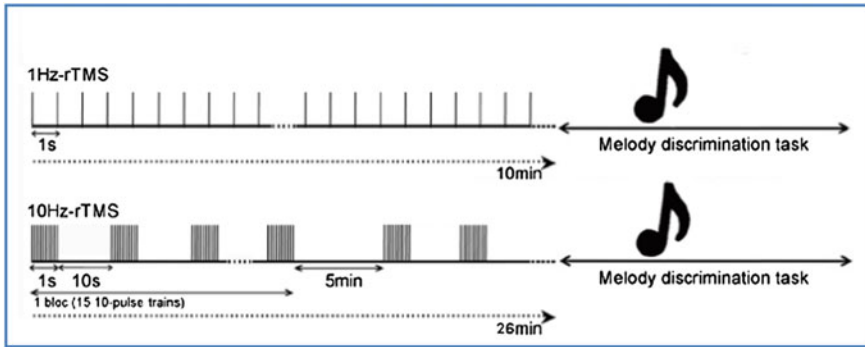
Fig. 1 Repetitive TMS (rTMS): for stimulating brain tissue different types of pulse sequences are in use [2]

discussing the methods' advantages and disadvantages. (For further information on MR physics please refer to the detailed, excellent explanations in [13, 19].

To investigate brain activity with fMRI, measurement has to be performed in two steps: In the initial phase, called Magnetic Resonance Imaging (MRI), pure neuroanatomical data are recorded to precisely reconstruct size, shape and structure of the individual brain. After that, changes of regional blood flow are registered to determine the amount of oxygen consumption. (This is the method's functional component, the 'f' in **f**MRI). Regarding the equipment, an expensive magnetic resonance scanner (tomograph) is the indispensable part; it should be able to produce a high-intense, static magnetic field B_0 with a field strength of 1 up to 7 Tesla (T).

MR physics uses the spin, i.e. the self-rotating property of hydrogen atoms (H). Inside the scanner, the hydrogen atoms of the human body (4.7×10^{27} in number) react like tiny compass needles. They orient along the external magnetic field B_0 and rotate with a special frequency termed 'Larmor precession frequency'. Whenever B_0 is modulated by introducing a sharp HF-impulse hydrogen atoms fold down from the Z- into the XY-plane, and magnetization changes from longitudinal to diagonal. To re-reach the starting position two types of 'backward forces', called 'spin-grid interaction' and 'spin-spin interaction' are effective, widely known as T1- and T2-relaxation [13].

In physical terms, T1-relaxation is defined as the point in time when longitudinal magnetization has regained 63 % of its original strength, whereas T2-relaxation is defined as the point in time when diagonal magnetization has decreased to 37 % of its initial value [13]. Gray and white matter, fat and cerebrospinal liquid differ in their relaxation times (T1 and T2), which enables researchers to use these parameters for adjustment, in particular for regulating brightness and image contrast of the MR scans. Neuromusicology uses almost exclusively MR scans of the T1-weighed type. This type of brain image makes gray and white matter (as well as other forms

of tissue with short T1-properties) look bright, whereas spaces filled with cerebrospinal liquid look dark, producing almost no signal.

In addition, several sub-processes, known as space-coding, are necessary to obtain precise 2- and 3-dimensional *spatial information* from the spin signal. Procedures for space-coding require a stimulation of selective layers as well as a modulation of the magnetic field and for this, paired gradient coils of X-, Y-, and Z-orientation are used [13]. These coils produce gradient fields superimposed on B_0, resulting in different strengths of the magnetic field at each point of the measuring volume. Regarding spatial information the MR signal can now be analyzed voxel by voxel [49].

Functional MR scans are the product of a further step of image acquisition, and for this aim, the so-called BOLD response is a necessary precondition (see next chapter).

2.1 BOLD Response and Its Underlying Principle

Let me describe the *physiological mechanism* responsible for the 'f' in fMRI: Blood itself, more precisely, the oxygen content of hemoglobin serves as a body-own (intravascular) indicator: Whenever a task has to be performed, either motor or cognitive, energy demand is high, and the regional cerebral blood flow (rCBF) increases. In the venous parts of the fine capillaries next to activated neural populations [49] an exchange of oxygen-deficient for oxygen-rich blood takes place, and this is the principle fMRI is based on. Note that the oxygenated type of hemoglobin differs from the de(s)oxygenated one in its magnetic susceptibility, the latter being slightly more susceptible than the former (also known as the para- vs. diamagnetic properties of blood; [13]). The principle of magnetic susceptibility was originally discovered by the American chemist Linus Pauling in the 1920s and was transferred to fMRI research in 1990, since then known as BOLD response (**B**lood **O**xygen **L**evel **D**ependent response, [36]). Jäncke [13] explains that deoxygenated (paramagnetic) hemoglobin has an inhibiting effect on the BOLD signal due to increasing magnetic field inhomogeneities yielding artifacts, but he concedes that the underlying mechanisms of signal production are still more complex.

The BOLD signal usually reaches its peak plateau between 4 and 8 s after task-onset. Thus, in comparison to bioelectrical methods, time resolution is poor. However, the main advantage of fMRI lies in its excellent spatial resolution with values between 1 and 4 mm^3 (or even below) which is a necessary precondition to precisely localize functionally activated areas.

Note that Talairach and Tournoux [53], two neurosurgeons of Swiss origin, developed a stereotaxic atlas, i.e. a sort of spatial coordinate system supporting researchers in their effort to localize specific brain areas. By using a color range from red to blue to indicate either activation or deactivation, clusters of functional activity can be marked voxel- or pixelwise on these brain scans. This stereotaxic atlas also provides the possibility to adjust the morphological structure of individual

brains or to transform data onto a template, the standard brain, used for data transfer between laboratories and for the comparability of results.

2.2 Techniques of Image Acquisition

Constructing an appropriate paradigm in fMRI research is no easy matter:

First, several options of pulse sequences must prove their suitability to stimulate the brain tissue in an adequate manner. Three of them, called *spin-echo-sequence*, *gradient-echo-sequence* and *echoplanar imaging (EPI)*, will be briefly introduced here (Fig. 2 shows similar considerations regarding TMS). In principle, stimulation starts with a HF-pulse, then a special combination of gradient coils is applied [13]. *Spin-echo-sequences* simply work with two (or more) HF-pulses. The initial 90° HF-pulse makes hydrogen atoms fold down into the XY-plane, whereas a second 180° HF-pulse forces the atoms to reverse their direction of precession; resulting in maximal strength of the to-be-analyzed signal. *Gradient-echo-sequences* and spin-echo-sequences differ from each other in a crucial point: Gradient coils change polarization on their own, which makes the process of signal generation less time-consuming while image quality of T1-weighed scans remains excellent.

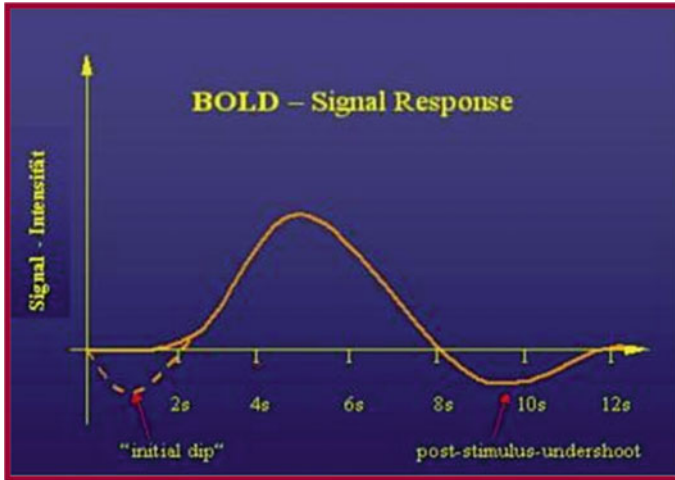**Fig. 2** Example of a T1-weighed scan (Universidad Autónoma de Zacatecas, Mexico)

Fig. 3 Time course of a BOLD response [49]

Echoplanar imaging (EPI) is the fast, high-resolution version of the gradient-echo technique. EPI enables researchers to record the whole brain in less than 2 s. A disadvantage of this stimulation type is, again, coil noise: The fast and permanent switching of gradient coils transmits small amounts of vibration onto the cylindrical scanner tube which automatically starts to resonate by producing volume intensities between 60 and 100 dB. Thus, it seems advisable to wear protective headphones or earplugs.

A second constraint is related to the BOLD signal itself: The fact that it needs at least 4 s to reach its plateau and, after decay, needs 12 to 15 s to recreate, imposes severe restrictions on the experimental design, meaning that signal presentation and inter-trial intervals (ITIs) have to be adapted to these limitations (Fig. 3).

2.3 The Auditory Cortex—A Challenge to fMRI Research

About 87 % of all fMRI studies use echoplanar imaging (EPI) as the method of choice, valued for rapid data collection [29]. Investigating the auditory cortex with fMRI, however, is a special problem. As already mentioned above, a disadvantage is that noise is produced by a fast switching of gradient coils during space-coding, ranging from 60 to 100 dB intensity. This side effect, similar to that observed for transcranial magnetic stimulation (TMS), makes the study of auditory processing complicated for, at least, three reasons: First, target sounds could be masked by the scanner noise and may hardly be detected. Second, the focus of attention may shift towards the noise and will eventually impair task performance. Third, emotional reactions to melodies as well as their aesthetic evaluation might extremely be

hindered, resulting in reduced brain activation of the limbic system (amygdala and hippocampus).

How can researchers deal with these side effects caused by scanner disturbance? Several effective solutions are suggested by Jäncke [13]: The first compensating technique he introduces is known as *sparse temporal sampling (STS)*. It is a variation of the 'echo-planar imaging' technique. STS is characterized by inserting pauses of up to 10 s length into periods of continuous scanning. These inter-stimulus intervals can make the coil noise fade out while at the same time, i.e. preceding the measurement, target melodies may fade in. Another possibility is termed *clustered acquisition*. This time, the entire set of target stimuli will be presented immediately before scanning, providing a time-frame between 4 and 6 s for image acquisition (data recording). Note that whenever *clustered acquisition* is used as the method of choice, the first two (of let's say, ten) fMRI scans have to be excluded from analysis: Owing to the fact that longitudinal magnetization has not been completely built up, signal strength is less than in the remaining scans [13]. Several other simple measures may also effectively reduce scanner noise. First, scanner-compatible headphones may attenuate the coil noise while at the same time the sound quality of target melodies can be enhanced. Another possibility is to line the inside of the scanner tube with sound-absorbing material (e.g. insulating mats) with attenuation effects up to 30 dB.

Faced with the challenge to increase the effectiveness of auditory fMRI designs Mueller et al. [31] tested the effects of three types of EPI-sequences on image acquisition: *continuous scanning*, *sparse temporal sampling (STS)* as well as a method called *interleaved silent steady state (ISSS)* that differs from STS in terms of scanner behavior in the silent gaps between recordings.

Mueller et al. [31] tested all three types of EPI sequences (continuous scanning, STS and ISSS) for possible differences in brain activation caused by the measuring technique itself. In each session of 12.5 min length the same 10 s excerpts of classical dances and their electronically distorted counterparts were taken as example pieces to examine the excitability of brain tissue in 20 volunteers (7 females). The study obtained two results: First, activations in left and right auditory cortices were significantly stronger for the original than for the distorted dance pieces. More interesting, however, is the observation of additional activations in the limbic system (left and right hippocampal structures) that could be made visible with ISSS but not with sparse temporal sampling. So, unexpectedly, the interleaved silent steady state method emerged as the most sensitive acquisition technique; thus, it might be the method of choice whenever subtle activities in sub-cortical structures or in deeper layers of the cerebrum have to be detected.

2.4 Positron Emission Tomography: Some Notes on the Signal and on Image Acquisition

It is the right time to take a closer look at positron emission tomography (PET), the older type of neuroimaging methods used for visualizing brain processes. Once again, energy consumption (oxygen and glucose) serves as an indicator to precisely localize brain activity. Spatial resolution obtained with this method is between 2 and 8 mm^3 depending on the type of PET scanner, i.e. it is slightly less accurate than that obtained with fMRI. When thinking about auditory paradigms neuroscientists often choose PET instead of fMRI. The reason why PET is preferred lies in the avoidance of scanner noise due to a different technique of data recording. In other words, PET in contrast to fMRI does not produce disturbing scanner noise at all. Thus, while lying in the tube of a PET scanner, participants can easily focus on the target sound, react emotionally and may also appreciate the target's aesthetic value. However, the major disadvantage of this method is that a radioactive tracer substance, mostly 15-Oxygen (^{15}O) or 2-Fluoro-2-Deoxyglucose, has to be injected intravenously. Radioactive isotopes emit positrons (particles of positive electric charge) that interact (or collide) with electrons (particles of negative electric charge), resulting in the emission of photons which can be measured with an array of scintillation detectors, i.e. a circle of sensors placed around the head [13].

Note that the half-life of each radioisotope imposes restrictions on the experimental paradigm in that time length available for task-related measurement is strictly determined by the rate of decay. The half-life of 15-Oxygen, for instance, is about 2 min, placing severe restrictions on the choice of appropriate stimuli: Single sound events such as pure tones or intervals have been proven suitable, whereas harmonic sequences or melodies should not exceed a time length of, on average, 10 s.

Jäncke [13] points out that due to risks of health only 10 injections of radioisotopes per participant seem acceptable, resulting in 20 min (10 shots à 2 min each) recording time, also restricting the number of the to-be-tested conditions. In compensation, and also to increase the statistical power, group-wise averaging of PET scans is advisable, and for this, PET data first have to be transferred to the template, the standard brain. Thus, in terms of *individual* PET data it seems almost impossible to achieve convincing statistical results.

2.5 Research with FMRI and PET: Example Fields of Music-Related Application

Scans depict functional activity in specific brain regions, using energy consumption as an indicator. Most results refer to the cerebrum, in particular the cortex, the basal ganglia and the limbic system, but, increasingly, the cerebellum and parts of the brain stem down to the pons (in particular the cochlear nucleus and the superior

olive of the *auditory* brain stem) are also investigated using neuroimaging methods (e.g. [50]). To give you a glimpse of an idea about the range of results obtained with fMRI and PET I will pick some out, choosing 'auditory processing', 'music theory' and 'creativity' as example fields of application.

Note, however, that from a neurophilosophical perspective fMRI results do not allow to draw a conclusion about the functioning of the *mind* or the type of knowledge representation (be it analogous or propositional; see e.g. [41, 45] for a further discussion). In other words: neuroimaging methods still cannot be used to distinguish between mind and brain, the old, 'hard' philosophical problem. Even so, first attempts have been made to reconstruct the mental content belonging to different semantic categories from brain scans showing cortical activation and deactivation (see e.g. [12]).

2.5.1 Studying the Human Auditory Cortex with PET and FMRI

Neuroimaging methods enable researchers to examine the specific functioning of the auditory cortex in detail. This way, many fundamental insights that were previously found by introspection, now can be verified via brain scans which may help disciplines like *Psychophysics* and *Tonpsychologie* strengthen their impact.

One example in this respect is a study performed by Zatorre and Belin [60]. Using PET, they were able to identify two functionally different parts in the auditory cortex, a core and a belt region. The core region is specialized in processing temporal features as typical for speech, whereas the belt region is specialized in processing spectral features as typical for tonal patterns. Furthermore, they observed a certain asymmetric shift (or functional lateralization) in that speech-like signals caused stronger activation in the left compared to the right auditory cortex whereas for signals with rich spectral content as in music the reverse was true. Zatorre and Belin [60] suggest that neuroanatomical structures on the micro-level, in particular different types of fiber myelination and cortical column width, may be the reason why rapid changes of signals are processed in the left auditory cortex, whereas spectral richness is more accurately processed in the right counterpart. (Note: It is the idea of starting with a common signal which was split up in two directions that makes the study trustworthy: Two pure tones in octave distance served as the standard signal and were randomly presented in alternating order. To either synthesize speech signals or simulate music, they were then speeded up in three steps (first condition) and enriched with additional spectral components (second condition)).

In addition, Warren et al. [59] could show by using fMRI that two regions of the secondary auditory cortex, termed planum polare (PP) and planum temporale (PT), react independently to pitch chroma and pitch height, the cyclical and linear dimensions of pitch as postulated by Drobisch [6] and by Révész [42]; (see Fig. 4). These findings confirm that the so-called 'Two-component theory of pitch' (*Zweikomponententheorie der Tonhöhe*), which is originally based on subjective assessments, indeed has a neuroanatomical counterpart. (To obtain these results,

individual sounds were manipulated in pitch height by attenuating the amplitude of the odd harmonics while keeping pitch chroma constant).

2.5.2 Tonality-Sensitive Areas—An Approach with fMRI

Another example field of application is music theory, although publications about the general laws in music and their neural base are scarce. Yet, Janata et al. [16] discovered a sort of *tonality-sensitive center* located in the rostromedial part of the prefrontal cortex (rmPFC). In this fMRI study, a melody was systematically modulated through all 24 major and minor keys, and eight musically experienced participants listened attentively to each transposed version.

While comparing the congruency of cortical activation patterns between transpositions, Janata et al. observed a certain "dynamic topography" in rmPFC (p. 2167) in that some activated voxels specifically responded to a certain group of keys but not to another. However, repetitions of the scanning procedure revealed that inter- as well as intraindividual variances in terms of these *key-sensitive* voxel-based activations were high (Fig. 5). Note that this type of key-sensitive, dynamic reaction in rmPFC was independent of the given task, (consisting in detecting tones of different timbre as well as in detecting those ones being out-of-key).

2.5.3 Musical Improvisation—An Example of Whole-Brain Image Analysis

In recent years, a tendency towards naturalness and authenticity has been observed showing that 'high ecological validity' becomes a core criterion in cognitive
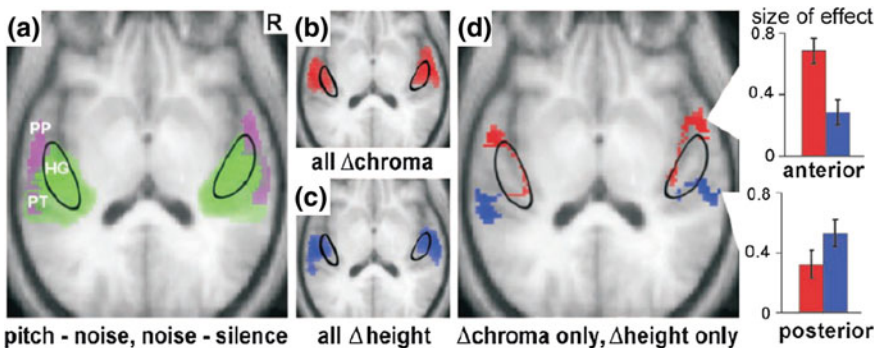

Fig. 4 Planum polare (PP) and planum temporale (PT) are differently activated when pitch height and pitch chroma are processed independently, confirming the 'two-component theory of pitch' [59]
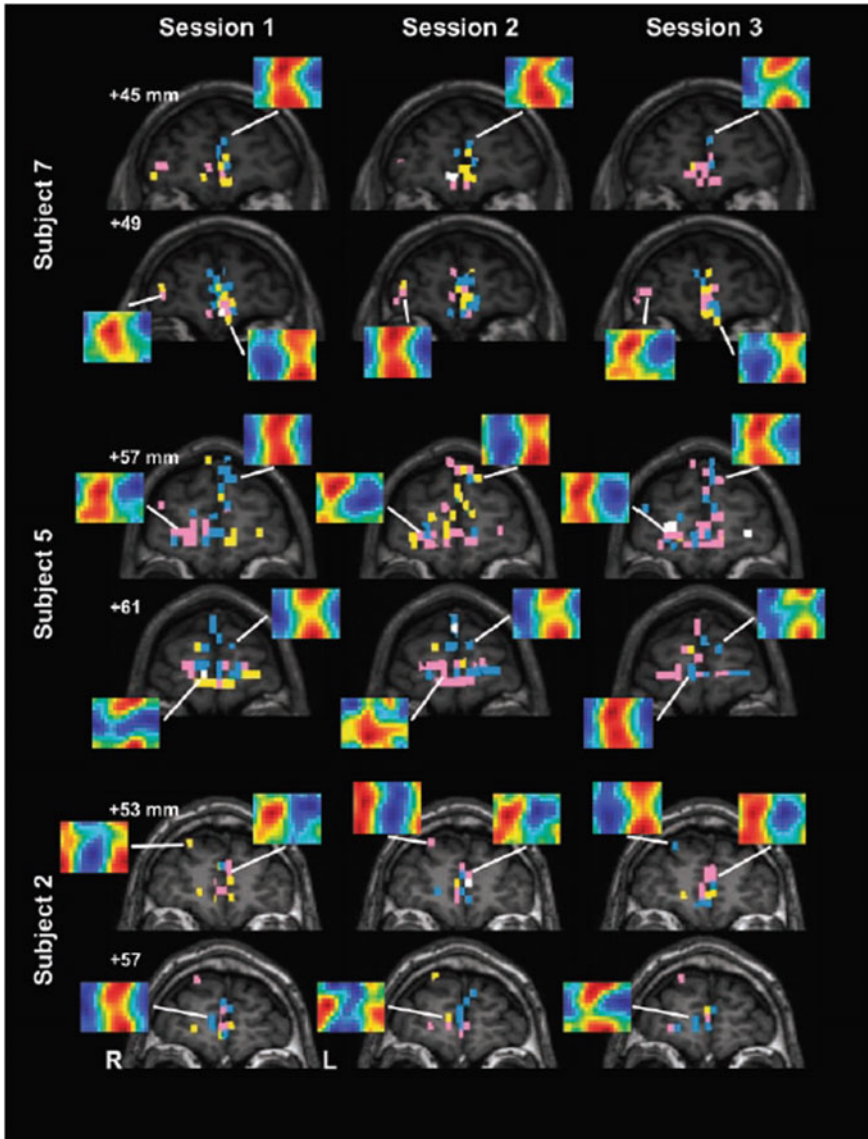
Fig. 5 Brain scans of subjects 2, 5 and 7 across three scanning sessions: the rostromedial PFC is sensitive to tonality. Within this area certain activated voxels reveal a key-specific behavior [16]

neuroscience. In this context some *neuromusicological field studies* (using EEG) occasionally have been put into practice: Fritz and colleagues, for instance, used a portable EEG equipment to record brain activity from native village inhabitants in Cameroon who listened to Western classical music for the first time (unpublished

work). A mobile EEG equipment has also been found useful to examine the effects of Cannabis on consuming rock music while sitting in the living room, smoking a couple of joints (see [9]).

Neuroimaging methods, by contrast, cannot fully meet the criterion of context-related authenticity as the scanning procedure should always be performed in a laboratory environment to obtain trustworthy results. Despite these constraints, a new tendency in fMRI research can be observed in that complex, natural musical pieces of several minutes length are used to investigate free natural listening and/or some types of spontaneous creativity.

In this context, Alluri et al. [1] demonstrated with fMRI that free listening to a modern tango of 8′32″ length activated cognitive, motor and emotion-related circuits on cortical and subcortical levels simultaneously while, at the same time, deactivations, mostly in pre- and postcentral cortical areas, were found (Fig. 6). In order to reliably relate the respective brain activity to the musical parameters of this tango, 21 music students judged 270 tango segments of 6 s length on a scale from 1 to 9 beforehand, according to the following parameters: 'brightness', 'fullness', 'activity' (i.e. 'change of timbre'), 'timbral complexity', 'rhythmic complexity', 'pulse clarity', 'key clarity', 'event synchronicity' and 'dissonance'. Interestingly, timbre features were not only processed in the auditory cortices, but also in some parts of the left and right cerebellum, while, at the same time, deactivations in the precentral and parietal regions could be found. 'Pulse clarity' was processed in the auditory cortices too, but showed deactivations in the insula and the limbic system (amygdala, hippocampus; see Fig. 6).

Complexity can even be increased when creative and performance aspects are added. One of the first neuroimaging studies addressing this issue was an fMRI experiment by Limb and Braun [28], testing the neural basis of improvisation. In regions of the prefrontal cortex, a dissociated pattern of activity was observed, showing deactivation in the dorsolateral and activation in the ventromedial parts simultaneously, which seems typical for any type of artistic, creative process (Fig. 7; see also [40] and text below): Dorsolateral deactivation stands for defocused attention combined with a lack of self-monitoring and volitional control, whereas the ventromedial activation may be interpreted as indicating basic attitudes and the preservation of rules. In addition, activity in the premotor and primary motor cortices indicated motor control as well as aspects of motor preparation (for playing a keyboard with the right hand in the scanner tube), whereas fine-grained adjustments of finger movements were regulated via cerebellar activity. Furthermore, emotionally inhibiting factors such as anxiety and fear seem to be suppressed during improvisation as the respective limbic areas, especially the amygdala, showed deactivation. Limb and Braun [28] conducted their study with six professional male pianists, each highly proficient in improvisation. While lying in the scanner, these pianists had to spontaneously modify a simple overlearned jazz melody on a small keyboard placed on their lap while listening to prerecorded combo chords via earphones.
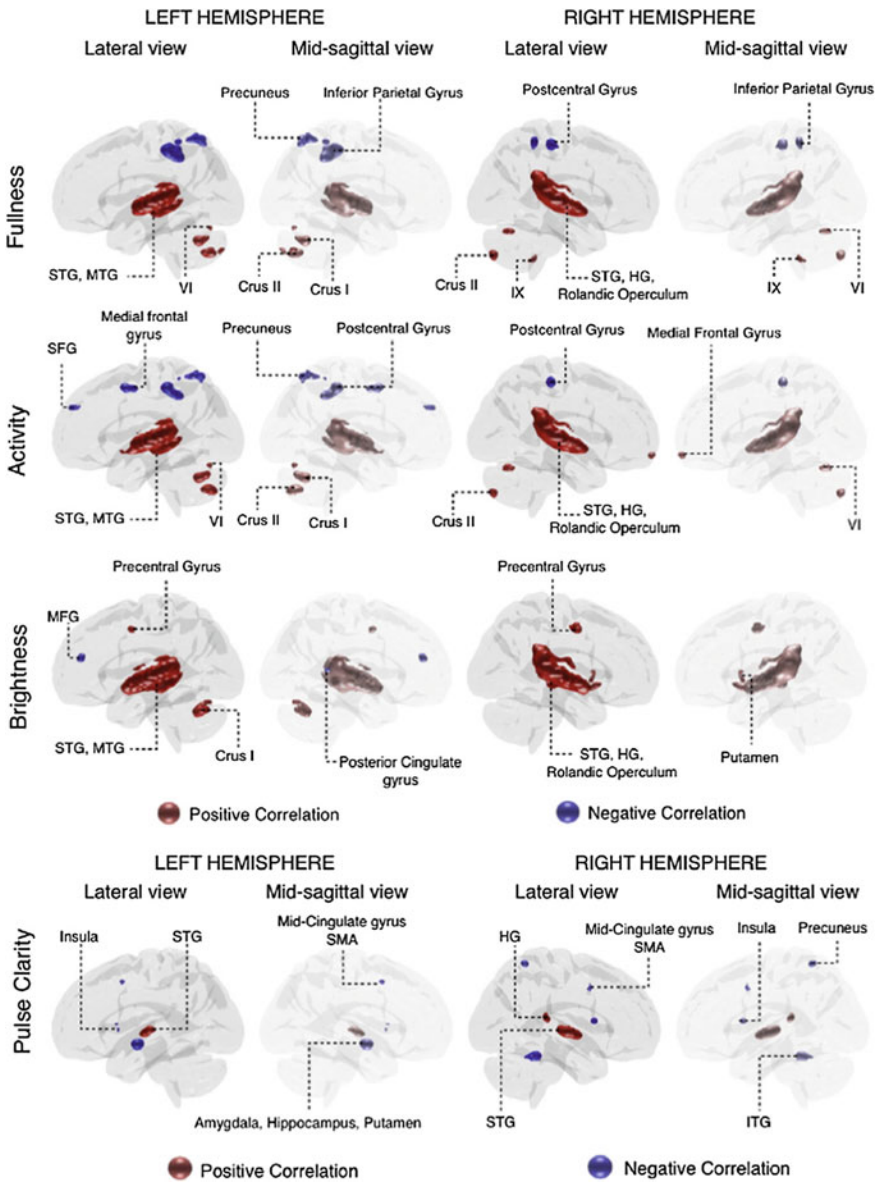
Fig. 6 Brain activity while listening to a complex piece of music—a modern tango. The auditory cortices are activated by high levels of pulse clarity and when processing several timbre features. Deactivations could be observed in the insula and the limbic system (for pulse clarity) as well as in parietal regions (for timbre attributes) [1]
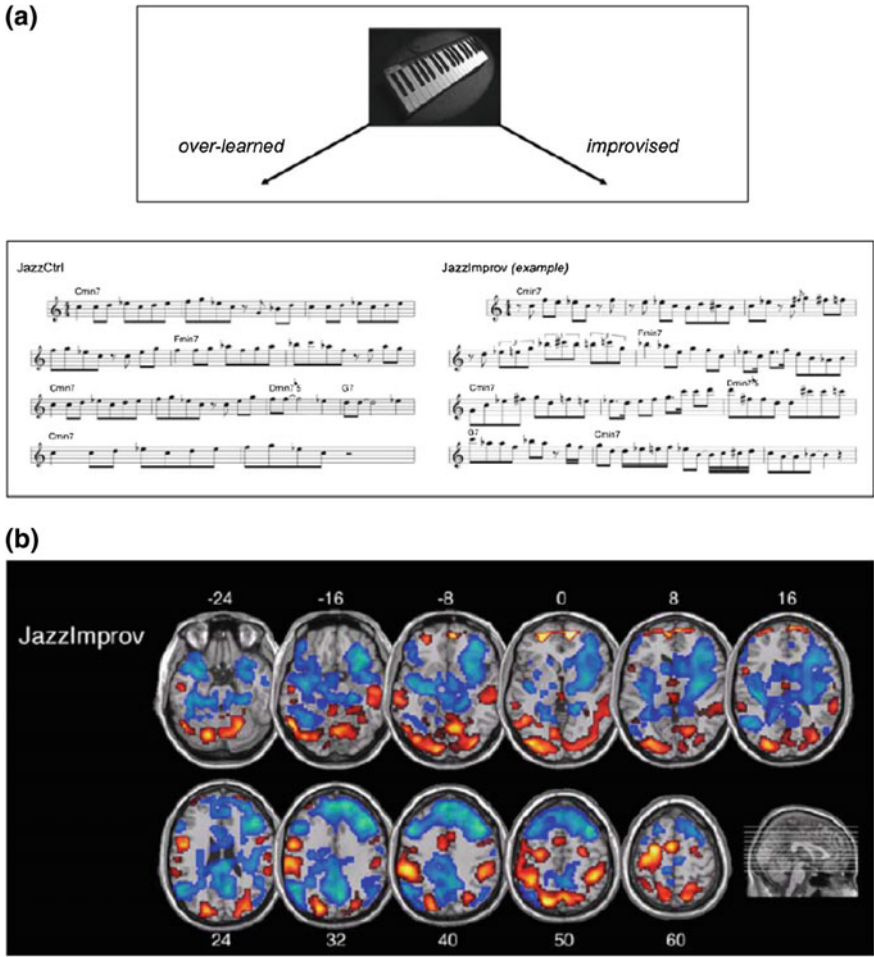
Fig. 7 a, **b** Activity in fMRI scans averaged over six male professional pianists. For any type of artistic creativity (here: jazz improvisation) a simultaneous division of the prefrontal cortex into deactivated (dorsolateral) and activated (ventromedial) parts can be observed [28]

2.6 Neuroplasticity in Musicians—Structural and Functional Types

Note that two decades ago, K. Anders Ericsson, an American psychologist of Swedish descent, developed a concept named 'deliberate practice', saying that high levels of proficiency (or expert performance) need years of intensive training especially in young adulthood (see [8]). Interestingly, deliberate practice leaves

'traces in the brain' through re-shaping its areas, a process widely known as *neuroplasticity*.

Neuroscientists distinguish between two forms, a functional and a structural type of experience-driven neuroplasticity. In the first case cortical activation strength, i.e. the susceptibility of brain tissue, is modified (see below), whereas in the latter significant enlargements, caused by an increase of dendritic branching as well as by intensification of synaptic strength, can be observed on the macro-level. A special method called 'voxel-based morphometry' enables researchers to precisely assess the extent of experience-induced anatomical changes while analyzing scans of the T1-weighed type (see [13] for details).

Münte et al. [32] considered the brains of professional musicians as the best fitting type to investigate these plastic changes, but also brains of sportspersons or chess players could serve as ideal models. Note that the brain's structural (and functional) changes seem to correlate with the age of learning to play a musical instrument in that effects are stronger the earlier piano or violin lessons start (i.e. typically before the age of 10) (see e.g. [7, 37, 38]).

A first result became visible in experienced string players, revealing an asymmetric (structural and functional) enhancement of the primary somatosensory cortex. The effect could clearly be detected for the fingering hand, i.e. for left hand fingers 2–5, but neither for the bow hand (right hand) nor for the left hand of a nonmusician control group [7].

Pantev et al. [37, 38] confirmed the fact of pure *functional neuroplasticity* in musical contexts: While listening inattentively to either piano versus pure tones (or to tones of familiar versus unfamiliar instrumental timbre) functional activity (or cortical dipole strength) was significantly enhanced in the auditory cortex of musicians. No similar effect could be observed for the nonmusician control group. Note that in these studies Pantev et al. [37, 38] decided for MEG (Magnetoencephalography) which is another non-invasive method of data recording enabling researchers to measure the brain's weak *electromagnetic* fields by using highly sensitive SQUID detectors. The distinguishing features of MEG are its high temporal as well as spatial resolutions, however, special software for source localization is required.

Another impressive result of neuroplasticity is beyond neuromusicology and refers to taxi drivers in London: Maguire et al. [30] could demonstrate that extensive experience in navigation, as typical for taxi drivers, causes a structural enlargement of the hippocampus, more specifically, a significant increase of gray matter density in the posterior hippocampal parts (Fig. 8).

3 Electroencephalography: The Basics

Let me move on to *electroencephalography (EEG)*, the oldest and most established type of neuroscience methods. By placing electrodes onto the head's surface, EEG research remains almost exclusively non-invasive. Note, however, that special
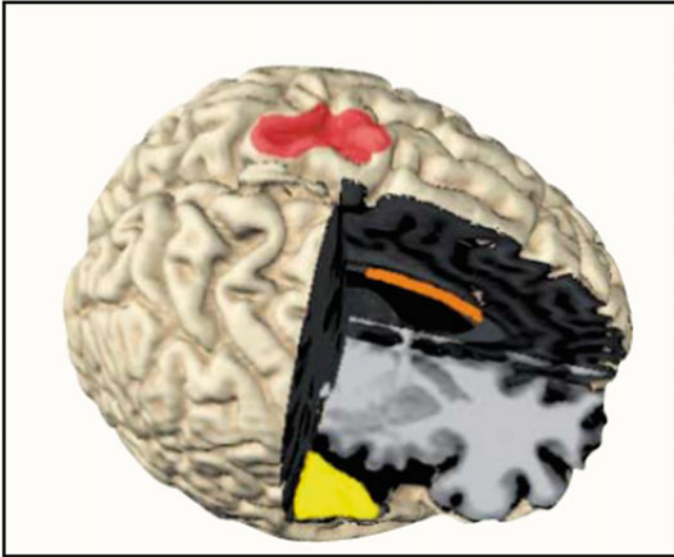
Fig. 8 The musician's brain serves as a model for neuroplasticity: structural enlargements of specific parts result from extensive training [32]

problems, for instance, localizing the source of musicogenic epileptic seizures during pre-surgical preparations, occasionally make it necessary, to either directly place electrodes onto the 'naked cortex', i.e. onto the gyri after opening the skull, or to implant them intracranially into the brain's tissue (e.g. [54]). The first variation is termed *electrocorticography (ECoG)*, the latter is called *depth EEG recording*.

The original, conventional EEG method was developed by Hans Berger, a German psychiatrist, at the beginning of the 1920s; he also coined the term 'electroencephalography' (Greek: enképhalon = brain; grapheîn = to write).

EEG potentials recorded from the scalp are raw signals, and each potential appears as the sum of extracellular field potentials stemming from different cortical layers. Extracellular field potentials, for their part, compensate intracellular voltage shifts when thousands of nerve cells 'fire' synchronously in order to solve a task. The electrogenesis of EEG scalp potentials can be described as follows (Tervaniemi and van Zuijen [55], p. 201):

> The EEG is a by-product of brain cells' information transfer in which intra- and extra cellular current flows are modulated with specific membrane mechanisms. When these current flows synchronize, potential differences summate, and become strong enough to be recorded with EEG. The post-synaptic activity of pyramidal dendrites (rather than action potentials) in the cortex particularly possess these characteristics and is therefore regarded as the main source of the EEG … Thus, in EEG, coherent activity of numerous cortical neurons (approximated by 10000) is recorded.

In most psychological contexts, EEG curves are registered by using *unipolar recordings* to measure potential differences between electrically active scalp

electrodes on the one hand and an electrically inactive reference point on the other (e.g., from an electrode placed on the earlobe). Furthermore, a standardized topographic schema named 'Ten-Twenty System', developed by a Canadian neuroscientist named Herbert Henri Jasper, helps to correctly place the scalp electrodes onto the head's surface. The 'Ten-Twenty System' creates a sort of individual anatomical coordinate system by taking two preauricular points as well as the nasion and the inion as reference points (see [17], for further details). This coordinate system enables researchers to precisely describe, compare or exchange sets with EEG data between research institutes (on the assumption that paradigms are equivalent) (Fig. 9).

Psychologically oriented EEG research often uses the Fast Fourier Transform algorithm (FFT) to separate four main frequency bands from each other (δ, θ, α, β; Fig. 10). Each band reliably shows a specific state of consciousness and/or level of arousal, ranging from 'coma', 'trance' and 'deep sleep' (indicated by a predominance of delta activity with oscillations between 0.5 and 4 Hz), to 'meditation' and 'drowsiness' (predominance of theta activity with oscillations between 4 and 8 Hz), followed by the state of 'being awake and relaxed' (predominance of alpha; 8–13 Hz) and the state of 'being mentally active and attentive' (predominance of beta, 13–30 Hz).

Note that a fifth type, the gamma band with frequencies between 30 and 80 Hz, is omitted in this context here: Gamma activity indicates 'feature binding', a specific process necessary to experience coherent percepts. It has mainly been found in the visual domain for binding spatially separate, 'static' properties together, like 'color', 'shape' and '(surface) texture'. Regarding the auditory domain, similar processes of temporal feature binding have been found less frequently. (Despite that [4], observed stronger gamma-band-synchronizations in musicians than in non-musicians when listening to musical excerpts, suggesting that musicians are more experienced in anticipating melodies of a certain style as well as in combining musical parameters such as contour, rhythm and timbre to a melodic entity.)
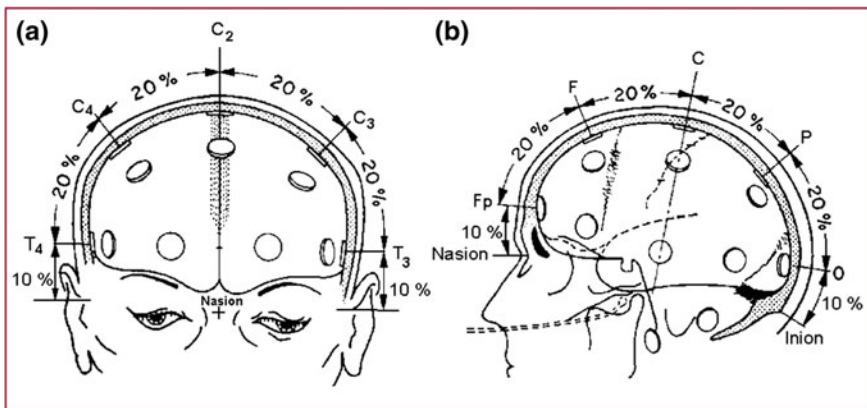

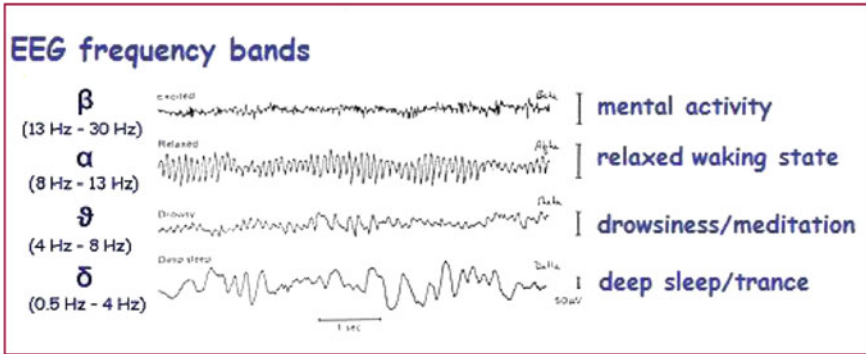Fig. 9 Electrode placements according to the 10–20 system [17]

Fig. 10 Fourier analysis: the EEG raw signal is divided into four frequency bands. Each indicates a specific state of consciousness and arousal (gamma activity is missing)

3.1 Research with EEG: Two Example Studies

It is common knowledge among ethnologists that states of Shamanic trance can be reached by taking drugs and/or by repetitive, monotonous drumming (e.g. [11, 44]). The brain reacts to this mind-expanding experience with a change in the spectral content of the standard EEG, in particular by an increase of theta and delta activity.

To elicit a form of sound-induced trance in a Western context, Kohlmetz et al. [23] chose a special piece of piano music named 'Vexations' (1893), written by the 27-year-old French composer Erik Satie (Fig. 11). Satie suggested playing this, per se, short (atonal and contrapuntal) composition 840 times without interruption, resulting in a total performance length of approximately 28 h. Here, trance is induced by an unusual playing instruction. Armin Fuchs, a German pianist, succeeded in getting through this 28-h 'endurance test' while the EEG was simultaneously recorded from parietal electrode sites. Despite mental stress and physical strains Fuchs kept tempo and motor performance relatively stable. He reported having experienced a 5-h-state of trance, in addition to a feeling of slowing down and that of lengthened time [23]. During trance, brain activity decreased bilaterally towards the delta-band and shifted slightly towards the left parietal electrode (P3). Note that only two electrodes were placed in the upper back of the pianist's head to avoid motion restrictions. However, from the viewpoint of EEG recording, this is below the minimum required by guidelines to fulfill the criterion of reliability.

Let me briefly describe a second example of music-related EEG research: During the 1980s, the Austrian neuroscientist Hellmuth Petsche came up with the idea of 'EEG coherence analysis', a methodological approach performed for each frequency band (δ, θ, α, β) separately to extend the conventional type of analyzing EEG raw signals via FFT. EEG coherence analysis has proven highly effective to investigate the interplay between cortical network structures during creative thinking and other mental processes of higher order. It describes the degree of similarity (or functional coupling) between EEG signals at adjacent electrodes of

Fig. 11 'Vexation', a piano piece by Erik Satie: It has to be played 840 times without interruption, resulting in a performance duration of 28 h and the experience of trance

the same hemisphere (the 'intrahemispheric type') or at homologous electrode sites on the opposite halves of the brain (the 'interhemispheric type').

In an EEG coherence study on composition, Petsche [40] put this approach into practice: Seven male professional composers had the task to spontaneously invent a tonal passacaglia and an atonal fugue, each of 5 min length, and write these pieces down immediately afterwards. Figure 12 shows the coherence patterns for one 56-year-old male composer while listening to a piece by Schönberg (the control condition) and while mentally composing in both styles. The most striking result appears as a shift of activation from left inferior-frontal regions (Fig. 12, left side), reflecting syntax analysis, towards the posterior parietal cortex (Fig. 12, middle), probably indicating some thoughts about the formal shape, or 'musical architecture' of the piece in progress.

Besides that, the study revealed two more results: First, any type of creative process in art (be it verbal, visual or musical) is indicated by a functional decoupling (or decrease of coherence) in the dorsolateral prefrontal cortex (dlPFC) so that bizarre, uncontrolled thoughts can enter (see also [5] and [28]). Second, long-distance coherences, for instance between fronto- and parietal electrode sites, increase during the mental act of composing, and interindividual differences are high (Fig. 13).
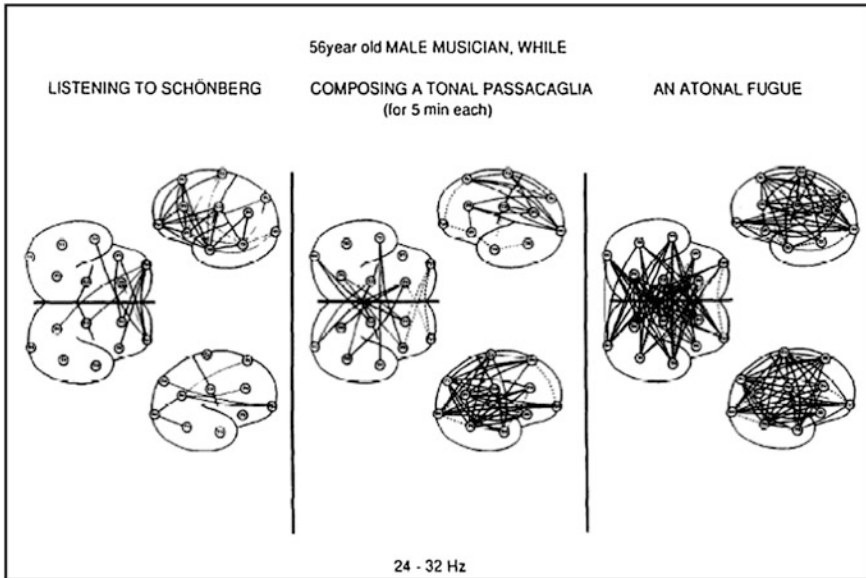
Fig. 12 Result of an EEG coherence analysis (beta band) for a 56-year old male composer: *Left* listening to Schönberg, *middle* composing a tonal passacaglia, *right* composing an atonal fugue (three views of the brain: from *above*, *left hemisphere*, *right hemisphere*) [40]

3.2 EEG Sports: A Promising Trend Using Mobile Devices

Investigating brain activity of humans *in action*, while playing golf, riding a bicycle or performing in a chamber music ensemble has been an unsolved problem in EEG research for many years. The most challenging aspect is not the real life situation per se but rather body movement as such: Any subject in motion produces many artifacts of extracerebral origin, arising from skin changes and muscle tension. In addition, sweating and loose electrodes may also contaminate the measurement and make EEG data not utilizable for further analysis (e.g. [56]). Furthermore, the recording equipment is unwieldy and heavy, including amplifiers and batteries, making it impossible to carry the device in a rucksack on the back. On the other hand, portable solutions would offer a wealth of opportunities in the field of human movement and sports science, while ecological validity would be high.

Until now, most studies in this context use EEG for neurofeedback-training in the lab, i.e. for investigating brain-based self-regulatory techniques that may help to modify the mental attitudes during several phases of practicing and performance. This way, certain EEG frequency bands, in particular alpha, theta and delta, can reliably be strengthened via monitor and other feedback devices, obviously increasing self-awareness, feelings of well-being and the supply of mental and physical energy necessary to succeed in any training session or sports competition outside [39, 56].
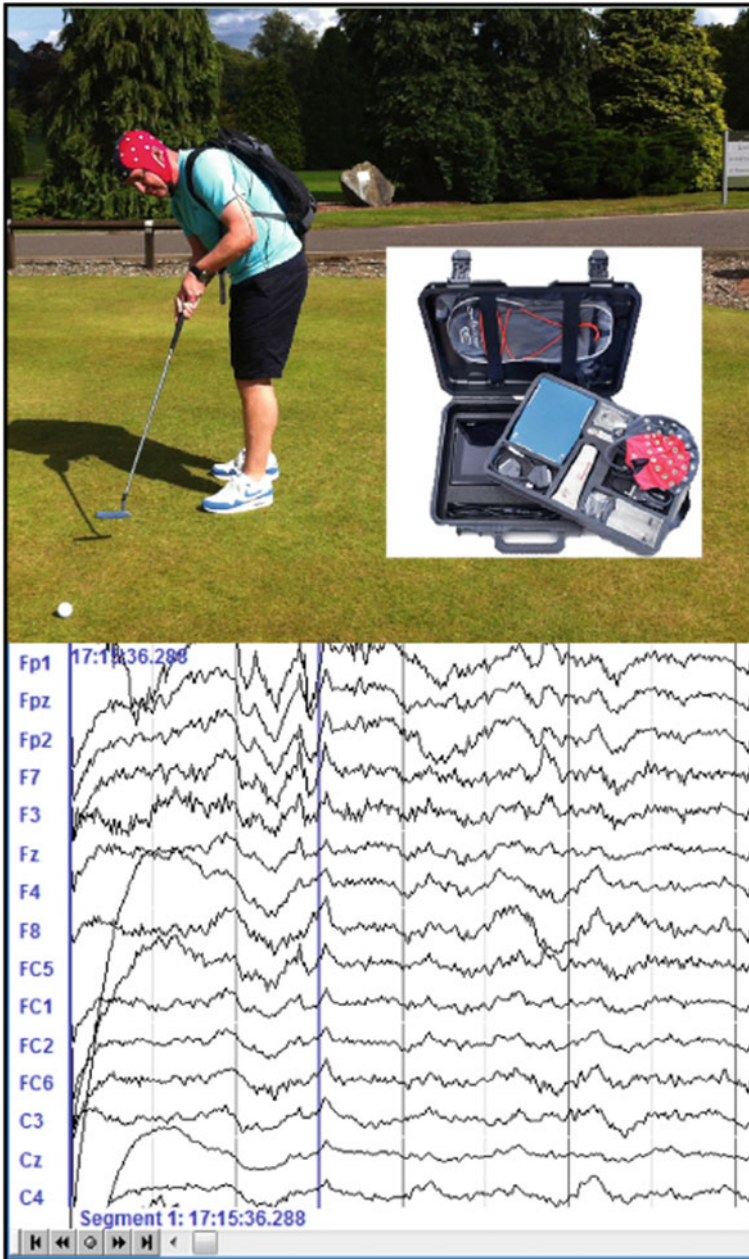
Fig. 13 A newly developed portable EEG solution, termed **eego**sports^TM, enables researchers to investigate brain activity in real life situations [39]

Recently, a new product series, termed **eego**sports ^{TM} has been developed by ANT Neuro, Enschede, a neuroscience company specialized in developing EEG hard- and software. Since 2013, they offer a portable, light-weight 64 channel EEG solution of less than 1000 g that enables researchers to freely investigate different types of movement as well as effects of training and physical exercise in a natural environment. Presumably, this mobile solution will be used in the context of 'music and motion' in the near future.

4 Event-Related Potentials (ERPs)—A Derivative of the EEG

Finally, let me describe the second type of bioelectric methods, known as the measuring of *'event-related potentials' (ERPs)*. ERP works on the precondition that, during recording, the same type of stimulus will be repeated at least 20 times which is not necessarily required for measuring the EEG.

Both methods, EEG and ERP, differ completely in their basic idea: EEG, on the one hand, allows to make individual recordings of several minutes length while disregarding transient brain activity, i.e. the components lasting some ms within ultra-short time frames. This way, the EEG informs about the brain's overall physiological state, i.e. the levels of consciousness and arousal while listening to music of different style and tempo.

ERP, on the other hand, is completely devoted to the basic idea of drawing an analogy between the computer and the human mind, meaning that both systems, the electronic and the human, should be considered similar in their strategies to select, transform, store or activate the respective information (see [34]). The ERP, therefore, directly points to a, mainly, serial form of processing input and comprises several independent processing steps in sequence [13]. (Note that according to this shift in thinking the word 'cognition', derived from the Latin word 'cognoscere', has lost its former philosophical connotations like 'becoming aware of', 'apperceive' or 'comprehend intuitively' and is now used in a simple, pragmatic way).

How are ERP results obtained? First of all, ERP uses the same initial recording procedure as EEG, so that raw traces as such cannot be classified as either belonging to the first or second type of method. Because of this, a special form of data analysis, termed *signal averaging*, is required to extract the amount of event-related brain activity from the raw data: Signal averaging enables researchers to split the raw signal into a spontaneous and a stimulus-related part by taking advantage of the fact that each stimulus repetition evokes a small but invariant brain response that can be summed up and divided by the number of presented stimuli, whereas spontaneous, stimulus-*uncorrelated* fluctuations converge against zero (Fig. 14). To further increase the signal-to-noise ratio, individual ERP curves are
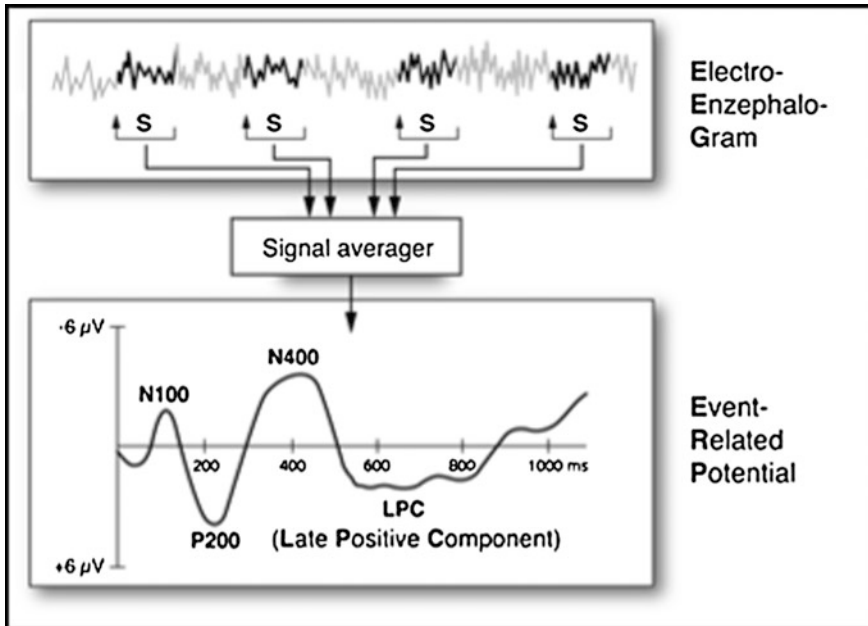
Fig. 14 Principle of signal averaging: marked EEG sections contain brain responses assigned to a specific stimulus S. They are summed up and divided by the number of stimulus repetitions, yielding the event-related potential. It informs about various steps of processing incoming information (MPI Cognitive and Brain Science Leipzig)

finally summed up to group-wise potentials, the so-called *grand average ERPs*, and this is the starting point for the analysis of ERP components (see below).

Rösler [43] points out that signal averaging and its product, the grand average ERP, are flawed with some weak points: First, brain responses stemming from several trial repetitions are summed up automatically which is considered inappropriate from a psychological point of view as it cannot be ruled out that participants might have changed attentiveness during recording. Second, brain responses are prone to habituation, i.e. amplitudes will be reduced the more familiar, or predictable, the often-repeated stimuli are. Third, grand average ERPs are produced at the expense of individual brain responses, meaning that conclusions regarding individual processing strategies cannot be drawn from the final product.

Note that it is not the ERP curve as a whole that serves as a unit for interpretation. Instead, each half wave, or ERP component, will be analyzed separately on the assumption that it responds independently, i.e. without any cohesive forces operating between adjacent components.

Regarding nomenclature, two details are needed to describe each ERP component properly: details about its 'polarity' and its 'latency'. The term 'polarity' describes the component's deflection, i.e. change in voltage direction either into the positive ('P') or the negative ('N'). 'Latency', by contrast, refers to the timespan

between stimulus onset and the peak amplitude and can either be described as a rounded value (in ms; e.g. P200) or as an ordinal number (e.g. P2). These sparse but essential details may be completed by some more information about the component's 'maximum amplitude value' [μV], its 'brain topography' and its 'waveshape'.

The curve example in Fig. 14 shows five components, termed P50, N100, P200, N400 and LPC (late positive component). The first ones up to 300 ms indicate exogenous processes that, in principle, are determined by stimulus parameters such as frequency, intensity or presentation rate. N400 and LPC, by contrast, indicate endogenous processes, reflecting some task-related cognitive processing steps for which attention is required. However, since recent results could show that exogenous components can be modulated by top-down processes too (e.g. [33]), contemporary ERP research directly focuses on the characteristics of the particular component itself, i.e. it omits this additional exogenous-endogenous classification.

To illustrate which aspects of cognitive processing a component may indicate, I will first pick out the so-called Mismatch Negativity (MMN).

4.1 The 'Mismatch Negativity' (MMN)—An Example Component of the ERP

Imagine you hear the following example: AAAA AAAA AA**B**A AAAA A**B**AA AAAA, i.e. a basic type of sequence in which two elements, a frequently repeated pure tone A (the standard) and an occasionally inserted pure tone B (the deviant) form an auditory chain which can be partitioned into bunches of four (the ratio of A to B may be 0.9:0.1 or 0.75:0.25). Interestingly, the brain will not only process the acoustic parameters of A and B separately (by developing a P50 and an N100), but will also react to this irregular and unpredictable alternation of items within the sequence. The detection of deviant 'B' will be indicated by a specific ERP component, called Mismatch Negativity (MMN) on the implicit assumption that a (temporary) memory trace has already been built for the regular sequence AAAA AAAA (Fig. 15). The MMN has its origins in the primary and secondary auditory cortices and the right frontal lobe, thus, indicating detection processes mainly for the auditory domain. Its maximum amplitude is between 130 and 250 ms measured from deviant onset.

From a functional point of view, the MMN indicates some trace-building processes within sensory (echoic) memory which gives rise to the assumption that a pattern-based process is the underlying driving force [20]. The second attribute is its independency from attention, enabling researchers to investigate both, attentive (controllable) as well as preattentive (automatic) processes of sound detection. Regarding the latter, attention is caught by instructing subjects to watch a silent video or read a book which prevents them from taking particular notice of the sounds themselves.
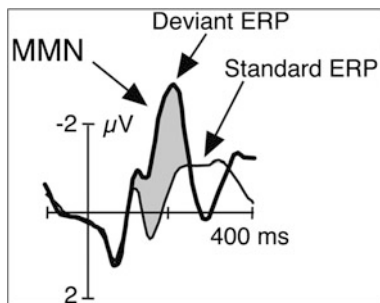
Fig. 15 The MMN is the first component that does not react to the properties of a single tone, but rather to some irregularities within the auditory sequence. Subjects may either focus on the tone series as such (attentive condition) or be distracted by watching a movie (preattentive condition). Adapted from Kujala and Näätänen [24]

Interestingly, these preattentive mechanisms of sound detection are modifiable by longstanding experience in that amplitudes are higher the more musical training participants have, in other words, the more accurate sound-templates in long-term memory are stored: Violinists, who are well-experienced in shading intervals and chords according to good intonation, automatically detect a slightly impure major chord (with frequencies of 396-491.25-596 Hz instead of 396-495-596 Hz), and this discrepancy between the actual input and the stored template will be indicated by a clear MMN. Musically inexperienced participants do not show a similar result [21].

4.2 Syntactic and Semantic Incongruities in Language and Music: ELAN/ERAN, P600 and N400

As already seen in the previous paragraph, ERP works best when tone rows are investigated, that is, when structure unfolds along the time axis. This way, sequence structures of any type will match the method's distinguishing feature of registering transient brain activity in high resolution on a ms-time scale.

By using a different paradigm, three specific ERP components have been found in the language domain, named ELAN (early left anterior negativity), N400 and P600, that are connected with a *rule-based type of sequence structure*: the ELAN and P600 indicate error detection in terms of syntax structure, whereas the N400 indicates deviation regarding semantic content.

In more detail, omissions with regard to word category ('the blouse was on [ ] ironed') elicit an ELAN between 150 and 200 ms (measured from onset of the word 'on'), whereas the late P600 indicates some sort of conscious re-analysis of the entire sentence structure (Fig. 16). Thus, ELAN and P600 are two ERP components

reflecting both, the preconscious (automatic) as well as the controlled aspects during the initial and later phases of syntax analysis [10].

Interesting parallels can be drawn between syntax processing in language and music: That is, melodies ending with a non-diatonic, incongruous final tone [3] evoke early and late components (named ERAN [early right anterior negativity] and P600) of similar shape and latency as those ones that were previously found for processing spoken sentences, allowing the conclusion that underlying processes are domain-general (Fig. 16).

(Note that similar comparative results were achieved when processing prosodic information: A specific ERP component termed Closure Positive Shift (CPS) was found for processing intonational phrase boundaries in spoken language as well as for processing phrase boundaries in music (while listening to binary-structured melodies); cf. [18, 35, 51]).

Besides that, a prominent component, termed N400 reacts to challenges of the semantic type: The N400 is visible whenever absurd or meaningless words in otherwise grammatically correct sentences are identified ("He carries his daughter in his *nostrils*"). The N400 is therefore interpreted as indicating violations of semantic expectancy [25].

It is worth investigating how the N400 'behaves' in the context of music, since music is commonly regarded as being more ambiguous than language. Koelsch et al. [22] approached this question by creating a priming situation, i.e. by presenting either a spoken sentence or a musical excerpt *preceding* a word, serving as a target stimulus. An N400 was evoked after both types of priming, making Koelsch suggest that a common associative network might be activated (Fig. 17).
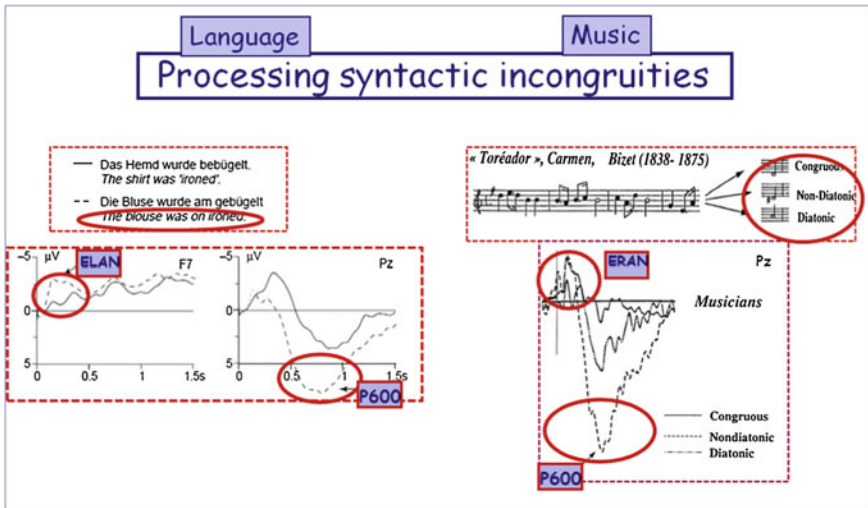

Fig. 16 Syntactic incongruities in spoken sentences (*left*) and in melodies (*right*) elicit similar components, termed ELAN/ERAN and P600 [3, 10]
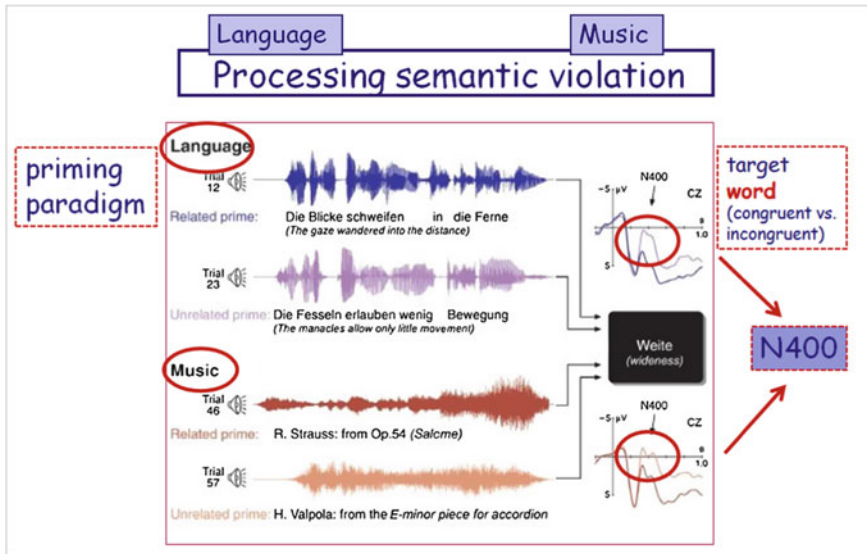
Fig. 17 An incongruous target word ("Weite") elicits an N400. It is unrelated to the previous context taken from music or the language domain [22]

To my knowledge, no study exists at present where the N400 is evoked by a semantic mismatch between a *musical context* on the one hand and a *musical target* (e.g. a chord) on the other, supporting the widespread view that chords, intervals and musical excerpts are less clear in meaning than words.

5 Do Advantages Outweigh the Disadvantages?—A Final Assessment of the Methods' Pros and Cons

Let me sum up the latest developments in neuromusicology. In my view, the following three tendencies become apparent: First, there is the endeavor to precisely relate *cause and effect*, i.e. to prefer causal relationships to correlative ones. This means that transcranial magnetic stimulation (TMS) is increasingly applied to music-related questions [26, 52], allowing researchers to assign an either slowed down or accelerated overt musical behavior to differently stimulated brain tissue.

The second trend is towards investigating brain activity in *real life situations*, i.e. to increasingly fulfill the criterion of *ecological validity*. Recently developed EEG mobile solutions (**eego**sports^TM) match well with this concept: They enable researchers to investigate the 'brain in action' while sledging, jogging, preparing a solo recital or playing in a jazz combo, in short: during every type of sport and performance activity. This trend also includes the endeavor to record brain activity

in natural environments, for instance, while performing cross-cultural field studies in non-Western countries.

Third, several labs advocate a holistic approach in that *whole-brain activity* is explored with fMRI while listening to complex musical pieces or while spontaneously improvising on the piano [1, 28]. Regarding this holistic approach, EEG coherence analysis, developed by Petsche as early as 1996, might be considered as a forerunner, since functional coupling of cortical network structures (while composing or listening to short pieces of music) can also be investigated by using this type of analysis.

Starting the assessment of methods with *EEG*—What are the main advantages and disadvantages of this oldest and most established type of neuroscience methods?

EEG allows the recording of *unspecific brain activity* over a time span of several minutes length while disregarding the transient components. Accordingly, the EEG is an appropriate method whenever *experiencing music*, be it repetitive drumming or a Mozart symphony. So, EEG is the method of choice whenever the focus of research is on the level of consciousness, on attention or arousal. Furthermore, Fourier analysis enables researchers to precisely observe changes in the spectral content of EEG signals over the entire time of recording. A disadvantage might be that interpretation is limited to statements about the brain's physiological state in general. However, a further plus point is coherence analysis, an option showing the functional coupling of brain activity at near and distant electrode placements, yielding information about the interplay between cortical network structures. So, whenever coherence analysis is included as an additional tool, the method's power is considerably increasing.

A second advantage is the possibility to analyze EEG raw traces for each subject separately. This is without alternative whenever mental states during creative processes are investigated, making the EEG method indispensable for creativity research (Schaffenspsychologie). However, there is also an option for a group-wise EEG analysis by comparing frequency bands of, e.g. musicians vs. non-musicians in relation to a specific task or a particular piece of music.

What are the pros and cons of measuring *'event-related potentials' (ERPs)*, the second type of bioelectric methods?

In principle, ERP components indicate information processing in a step-by-step manner. This way, ERP supports the basic idea of cognitive psychology which says that the human mind and the computer work on analogous principles [34]. In addition, excellent time resolution allows the discovery of new components beyond the established ones, for instance the face-sensitive N170, indicating the earliest stage of face recognition (e.g. [46]).

Regarding contents, ERP is the appropriate method to investigate three specific aspects: (a) to examine brain responses to frequency, intensity and other sound parameters in the context of psychoacoustics (eliciting a P50, an N100 and a P200, respectively), (b) to investigate brain responses to (simple) auditory sequence structures (eliciting an MMN) and (c) to examinethe processing of rule-based

syntax structures (evoking components such as ELAN/ERAN and the P600) so that comparisons between language and music can be drawn.

However, a lot of methodological constraints are imposed on each ERP design with restrictive effects on the interpretation of results: First, to increase the signal-to-noise ratio—between the event-related potential and spontaneous, unrelated EEG activity—the same type of stimulus has to be repeated between 20 and 1000 times (depending on the respective paradigm). Second, grand average ERPs refer to the entire number of trials and subjects, making it impossible to measure inter-individual differences or to reconstruct individual responses in retrospect. Third, brain responses are prone to habituation which means that, for each subject, trials are subsumed to a single common curve while disregarding minor or even major shifts in attention which is considered inappropriate from a psychological point of view.

Fourth, empirically working musicologists should know that ERP does not allow any conclusion about processing musical pieces of a particular epoch, a specific genre or the personal (idiosyncratic) style of composer (e.g. Händel vs. Bach). Fifth, the ERP responds to structure-violation in an overall-sense, for instance to any deviant chord within the standard scheme, be it a Neapolitan Sixth (N^6 or s^n) or a double dominant (D^D) in a musical cadence. These types of deviant chords will evoke the same ERP component (ERAN), no further specification in terms of harmonic progression will be possible.

Finally, let me weigh up the main advantages and disadvantages of *fMRI* and *PET*, the most popular types of neuroimaging methods:

Due to an excellent resolution in the spatial domain (approximately 1–4 mm^3 for fMRI and 2–8 mm^3 for PET depending on the scanner type) both neuroimaging methods provide the possibility to localize even the smallest functionally activated brain areas, based on a voxel-wise analysis. This way, the complex interplay between cortical and subcortical network structures, including the basal ganglia, the cerebellum and parts of the brain stem, can be made visible.

However, the precise 'Where' in the brain is at the expense of the 'When': The BOLD signal reaches its peak plateau between 4 and 8 s after task-onset, thus, in comparison to EEG and ERP, time resolution is poor.

Among the various options of stimulating brain tissue with pulse trains and sharp HF-impulses to obtain high-quality imaging data, echoplanar imaging is the fastest, enabling researchers to record the whole brain in less than 2 s. However, a disadvantage of this stimulation type is the technical noise in the scanner with volume intensities between 60 and 100 dB due to a fast switching of gradient coils during space-coding, a necessary step for image acquisition. Note that the interleaved silent steady state method is the most sensitive of all echo-planar imaging techniques, suitable for detecting subtle activities in subcortical structures as well as in deeper layers of the cerebrum [31].

However, in terms of musically-related contexts neuroscientists often choose PET, the older one of both imaging techniques. PET in contrast to fMRI does not produce any disturbing scanner noise at all. This enables participants to deepen their emotional experience and perceive a musical piece in an aesthetic sense while lying

in the scanner. However, the major disadvantage of the PET method is that a radioactive tracer substance has to be intravenously injected. This procedure imposes several restrictions on the recording procedure and the choice of stimuli, both aspects are strictly determined by the rate of decay.

To sum up: Do advantages outweigh the disadvantages? The question should be answered in the positive: Neuroscience methods offer elegant solutions to measure cognitive processes in real-time, yielding results of either high-temporal or high-spatial resolution. This fits nicely with a proposal by Leman [27]: To solve research problems more successfully he recommends a "joint correlative approach between different research methodologies; in particular musicology, computer modeling, experimental psychology and […] neuromusicology" (p. 194f), in short, a "convergence paradigm" . This is in accordance with a truly systematic approach as advocated by Schneider [47]: "The ultimate goal of systematization is to establish coherent systems of knowledge that should be free from any contradictions, and should be as complete in descriptions and explanations of the principles and phenomena of a certain field as is possible." (p. 20).

However, to dampen euphoria and overoptimism regarding the available neuroscience methods and their capacities, take notice of the following:

"The goal of neural science is to understand the mind—how we perceive, move, think, and remember." Despite all efforts, this statement by Eric Kandel (cited in [58], p. 1) still cannot be put into practice. (Some experiments on mental rotation make an exception, e.g. [15, 48]). Until now, many impressive methods inform about the *physiological state* and the *functional activity of the brain*. But how the mind works, is a different matter. Scientists still do not know for sure how thoughts are generated and how mental knowledge representations precisely look like. Nevertheless, attempts have recently been made to reconstruct the mental content belonging to different semantic categories from fMRI scans showing cortical activation and deactivation (e.g. [12]). However, the main disadvantage of this approach is ambiguity in that, until now, no clear assignment between both types of substance, the material and the immaterial world, can be made. Even so, innovations in the field of neuroscience are growing rapidly, so there are grounds to believe that the dualism between mind and brain, the so-called 'hard problem', may be solved in the near future.

References

1. Alluri, V., Toiviainen, P., Jääskeläinen, I.P., Glerean, E., Sams, M., Brattico, E.: Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. NeuroImage **59**, 3677–3689 (2012)
2. Andoh, J., Zatorre, R.J.: Interhemispheric connectivity influences the degree of modulation of TMS-induced effects during auditory processing. Front. Psychol. **2**, Article 161, 13 pages (2011). doi:10.3389/fpsyg.2011.00161

3. Besson, M., Faïta, F.: An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. J. Exp. Psychol.: Hum. Percept. Perf. **21**(6), 1278–1296 (1995)
4. Bhattacharya, J., Petsche, H., Pereda, E.: Long-range synchrony in the γ-band: role in music perception. J. Neurosci. **21**(6), 6329–6337 (2001)
5. Dietrich, A.: The cognitive neuroscience of creativity. Psychon. Bull. Rev. **11**, 1011–1026 (2004)
6. Drobisch, M. W.: Über musikalische Tonbestimmung und Temperatur [On musical pitch estimation and temperature]. In: Abhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften **2**, 1–120. Hirzel, Leizpig (1855).
7. 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**(5234), 305–307 (1995)
8. Ericsson, K.A.: The influence of experience and deliberate practice on the development of superior expert performance. In: Ericsson, K.A., et al. (eds.) The Cambridge Handbook of Expertise and Expert Performance (Chapter 38, pp. 685–706. Cambridge University Press, New York (2006)
9. Fachner, J.: Topographic EEG changes accompanying Cannabis-induced alteration of music perception—Cannabis as a hearing aid? J. Cannabis Ther. **2**(2), 3–36 (2002)
10. Friederici, A.D.: Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. **6**(2), 78–84 (2002)
11. Gingras, B., Pohler, G., Fitch, W.T.: Exploring Shamanic journeying: Repetitive drumming with Shamanic instructions induces specific subjective experiences but no larger Cortisol decrease than instrumental meditation music. PLOS One **9**(7), 9 pages (2014)
12. Haynes, J.-D., Rees, G.: Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. **7**, 523–534 (2006)
13. Jäncke, L.: Methoden der Bildgebung in der Psychologie und den kognitiven Neurowissenschaften. W. Kohlhammer, Stuttgart (2005)
14. Jäncke, L.: Lehrbuch Kognitive Neurowissenschaften. Huber, Bern (2013)
15. Jäncke, L., Jordan, K.: Functional neuroanatomy of mental rotation performance. In: Mast, F.W., Jäncke, L. (eds.) Spatial Processing in Navigation, Imagery and Perception, pp. 183–207. Springer, New York (2007)
16. Janata, P., Birk, J.L., van Horn, J.D., Leman, M., Tillmann, B., Bharucha, J.J.: The cortical topography of tonal structures underlying Western music. Science **298**, 2167–2170 (2002)
17. Jasper, H.H.: The ten-twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. **10**(2), 370–375 (1958)
18. Knösche, T.R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O.W., Friederici, A.D.: Perception of phrase structure in music. Hum. Brain Mapp. **24**(4), 259–273 (2005)
19. Köchli, V.D., Marincek, B.: Wie funktioniert MRI?. Springer, Berlin (1998)
20. Koelsch, S.: Music-syntactic processing and auditory memory: similarities and differences between ERAN and MMN. Psychophysiology **46**, 179–190 (2009)
21. Koelsch, S., Schröger, E., Tervaniemi, M.: Superior pre-attentive auditory processing in musicians. NeuroReport **10**, 1309–1313 (1999)
22. Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., Friederici, A.D.: Music, language and meaning: brain signatures of semantic processing. Nat. Neurosci. **7**, 302–307 (2004)
23. Kohlmetz, C., Kopiez, R., Altenmüller, E.: Stability of motor programs during a state of meditation: Electrocortical activity in a pianist playing 'Vexations' by Erik Satie continuously for 28 hours. Psychol. Music **31**(2), 173–186 (2003)
24. Kujala, T., Näätänen, R.: The mismatch negativity in evaluating cental auditory dysfunction in dyslexia. Neurosci. Biobehav. Rev. **25**(6), 535–543 (2001)
25. Kutas, M., Hillyard, S.A.: Reading senseless sentences: brain potentials reflect semantic incongruity. Science **207**, 203–208 (1980)

26. Launay, J., Dean, R.T., Bailes, F.: Rapid learning of associations between sound and action through observed movement. A TMS study. Psychomusicology **26**(1), 35–42 (2016)
27. Leman, M.: Relevance of neuromusicology for music research. J. New Music Res. **28**(3), 186–199 (1999)
28. Limb, C.J., Braun, A.R.: Neural substrates of spontaneous musical performance: an fMRI study of Jazz improvisation. PLoS One **3**(2), e1679 (11 pages) (2008)
29. Logothetis, N.K.: What we can do and what we cannot do with fMRI. Nature **453**, 869–878 (2008)
30. Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J., Frith, C.D.: Navigation-related structural change in the hippocampi of taxi drivers. PNAS **98** (8), 4398–4403 (2000)
31. Mueller, K., Mildner, T., Fritz, T., Lepsien, J., Schwarzbauer, C., Schroeter, M.L., Möller, H. E.: Investigating brain response to music: a comparison of different fMRI acquisition schemes. NeuroImage **54**, 337–343 (2011)
32. Münte, T.F., Altenmüller, E., Jäncke, L.: The musician's brain as a model of neuroplasticity. Nat. Rev. Neurosci. **3**, 473–478 (2002)
33. Musacchia, G., Sams, M., Skoe, E., Kraus, N.: Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. PNAS **104**(40), 15894–15898 (2007)
34. Neisser, U.: Cognitive Psychology. Meredith, New York (1967)
35. Neuhaus, C., Knösche, T.R., Friederici, A.D.: Effects of musical expertise and boundary markers on phrase perception in music. J. Cogn. Neurosci. **18**(3), 472–493 (2006)
36. Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W.: Brain magnetic resonance imaging with contrast dependent on blood oxygenation. PNAS **87**, 9868–9872 (1990)
37. Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M.: Increased auditory cortical representation in musicians. Nature **392**, 811–814 (1998)
38. Pantev, C., Roberts, L.E., Schulz, M., Engelien, A., Ross, B.: Timbre-specific enhancement of auditory cortical representations in musicians. NeuroReport **12**(1), 169–174 (2001)
39. Park, J.L., Fairweather, M.M., Donaldson, D.I.: Making the case for mobile cognition: EEG and sports performance. Neurosci. Biobehav. Rev. **52**, 117–130 (2015)
40. Petsche, H.: Approaches to verbal, visual and musical creativity by EEG coherence analysis. Int. J. Psychophysiol. **24**, 145–159 (1996)
41. Pylyshyn, Z.: Return of the mental image: are there really pictures in the brain? Trends Cogn. Sci. **7**(3), 113–118 (2003)
42. Révész, G.: Tonpsychologie. Voss, Leipzig (1913)
43. Rösler, F.: Statistische Verarbeitung von Biosignalen: Die Quantifizierung hirnelektrischer Signale. In: Baumann, U., et al. (eds.) Klinische Psychologie: Trends in Forschung und Praxis 3, pp. 112–156. Huber, Bern (1980)
44. Rouget, G.: Music and trance. A theory of the relations between music and possession. Chicago University Press, Chicago (1985)
45. Rumelhart, D.E., Norman, D.A.: Representation in memory. Stevens Handbook of Experimental Psychology 2, 2nd edn, pp. 511–587. Wiley, New York (1988)
46. Sagiv, N., Bentin, S.: Structural encoding of human and schematic faces: holistic and part-based processes. J. Cogn. Neurosci. **13**(7), 937–951 (2001)
47. Schneider, A.: Foundations of systematic musicology: a study in history and theory. In: Schneider, A. (ed.) Systematic and Comparative Musicology: Concepts, Methods, Findings, pp. 11–61. Peter Lang, Frankfurt am Main (2008)
48. Shepard, R.N., Metzler, J.: Mental rotation of three-dimensional objects. Science **171**, 701–703 (1971)
49. Siedentopf, C.M.: (Internet source) University of Innsbruck, Austria (2013). www.fMRI-easy.de
50. Sigalovsky, I.S., Melcher, J.R.: Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers. Hear. Res. **215**(1–2), 67–76 (2006)
51. Steinhauer, K., Alter, K., Friederici, A.D.: Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nat. Neurosci. **2**(2), 191–196 (1999)

52. Stupacher, J., Hove, M.J., Novembre, G., Schütz-Bosbach, S., Keller, P.E.: Musical groove modulates motor cortex excitability: a TMS investigation. Brain Cogn. **82**, 127–136 (2013)
53. Talairach, J., Tournoux, P.: Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging. Thieme Medical Publishers, New York (1988)
54. Tayah, T.F., Abou-Khalil, B., Gilliam, F.G., Knowlton, R.C., Wushensky, C.A., Gallagher, M.J.: Musicogenic seizures can arise from multiple temporal lobe foci: intracranial EEG analyses of three patients. Epilepsia **47**, 1402–1406 (2006)
55. Tervaniemi, M., van Zuijen, T.L.: Methodologies of brain research in cognitive musicology. J. New Music Res. **28**(3), 200–208 (1999)
56. Thompson, T., Steffert, T., Ros, T., Leach, J., Gruzelier, J.: EEG applications for sport and performance. Methods **45**, 279–288 (2008)
57. Tiitinen, H., Virtanen, J., Ilmoniemi, R.J., Kamppuri, J., Ollikainen, M., Ruohonen, J., Näätänen, R.: Separation of contamination caused by coil clicks from responses elicited by transcranial magnetic stimulation. Clin. Neurophysiol. **110**, 982–985 (1999)
58. Wagemans, J., Vertraten, F.A.J., He, S.: Editorial—beyond the decade of the brain: towards a functional neuroanatomy of the mind. Acta Psychol. **107**, 1–7 (2001)
59. Warren, J.D., Uppenkamp, S., Patterson, R.D., Griffiths, T.D.: Separating pitch chroma and pitch height in the human brain. PNAS **100**(17), 10038–10042 (2003)
60. Zatorre, R.J., Belin, P.: Spectral and temporal processing in human auditory cortex. Cereb. Cortex **11**, 946–953 (2001)

Author Biography

Christiane Neuhaus studied Systematic Musicology at the University of Hamburg. She has strong research interests in neuromusicology and empirical musicology and was a postdoctoral fellow at the Max Planck Institute for Human Cognitive and Brain Sciences Leipzig for seven years. She wrote her Habilitation thesis about 'Structural hearing' and was recently invited as a guest professor at the Eberhard Karls University Tübingen. She presently works as a Privatdozentin at the University of Hamburg.