Chapter 8

Brain, music and musicality: Inferences from neuroimaging

Robert Turner and Andreas A. Ioannides

> The man that hath no music in himself
> Nor is not moved with concord of sweet sounds
> Is fit for treasons, stratagems and spoils;
> The motions of his spirit are dull as night
> And his affections dark as Erebus:
> Let no such man be trusted.
>
> Shakespeare, *The Merchant of Venice*

8.1 **Introduction**

We approach the question of human musicality by considering the spatial and temporal patterns of brain activity that may be specific to the human experience of music. Commonalities in any activity or behaviour across people and cultures point towards biologically rather than culturally defined competencies of mind and brain. As brain scientists, we wish to discriminate such innate cerebral competencies for music from conventional, culturally learned skills, and to make sense of the functional localizations for both types of activity in particular parts of the brain that are amenable to study using neuroimaging techniques.

There is an important parallel here with studies of language in which, over the past century, the question of innateness has been much debated, especially with reference to the theory of generative grammars and the possibility of a brain-based 'language instinct'. Since instrumental music and some songs lack explicit reference to things experienced in the world, yet excite powerful cognitive and emotional effects, it is likely that the study of music, and especially its foundations in the motives for musicality, can provide insights into the innate competencies of the brain that enable the production and interpretation of the ordered sound sequences that constitute spoken language.

In any given culture, the classification and terminology used by educated people in characterizing the structured, ordered and intentional sequences of sounds understood as 'music' take on a unique authority. In Western studies of music, terms such as 'pitch', 'rhythm', 'melody', and 'timbre' tend to be regarded as applicable to music found in any human group, whether or not that group itself employs corresponding concepts in their music. Studies seeking universal principles of brain organization or brain activation for these Western concepts could potentially be taking an answer for granted. However, it is outside the scope of this chapter to review the ethnomusicological literature on musical terminology worldwide (and see Cross and Morley, Chapter 5, this volume). For the purposes of describing the corpus of research exploring brain representations of musical competencies, we will accept the conventional terminology widely used to identify features of music in Western cultures—melody, harmony, rhythm, chord, attack, timbre, pitch, tempo and tonality—because this research corpus has done just that. Even where the terms of analysis do not correspond to human universals, much can still be learned about how our brains internalize the collective representations that make up a musical culture, by studying the brain processes by which these musical concepts come to derive their unquestionable character.

Music, as a structured and intentional succession of movement-produced sounds, is clearly comparable to other complex products of human activity, including both speech and non-spoken language (Dissanayake Chapter 2, Brandt Chapter 3, this volume). These areas are now the subject of intensive research using the neuroimaging techniques that we will discuss below, such as magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI. However, until recently, these methods have been little used to investigate music and musicality—partly because music and musicality have been seen by influential language-focused scholars as less important to human life, and partly because they (especially the motivational processes of musicality) are more difficult to fit into easy categories of object or stimulus.

For some time, imaging techniques have been established that can tell us where and when the brain has increased activity. More recently, new techniques have been proposed that can also quantifiably explore patterns in regional electrical brain activity; while the topic of the neural correlates of musical expression and performance is at least as interesting as those of musical perception, we will, given the limitations of current research methods, address mainly the latter in this chapter.

Combining these powerful imaging techniques can be used to address the following questions:

- Are the experienced effects of stimulation by music and language related in the brain, and if so, how?
- What aspects of music correspond to our basic ‘hard-wired’ (i.e., early developed and regulatory) brain capabilities?
- Are specific parts of the adult brain specialized for perceiving components of musical awareness, such as those we define as pitch, harmony and rhythm?
- How does brain activity at different timescales relate to different aspects of musical activity or stimulation, and how are these dynamic events integrated into a unified and powerful percept of a musical mood, theme or narrative?
- What changes occur in the brain when we learn to read music or to play an instrument?
- How does the system of sound awareness relate to the functions of the emotion system, to cause us to be profoundly moved by certain qualities of music?

There is a plethora of techniques for mapping brain activity. Here we will review non-invasive methods that allow the localization and timing of human brain activity over spatial and temporal scales that are relevant to the study of music. These methods fall into two categories: those which track changes in cerebral blood flow or blood oxygenation correlated with changes in neuronal activity, and those which directly detect changes in electrical or magnetic signals caused by neuronal activity. We begin with a summary of the main physical principles involved, and outline advantages and limitations for research on psychological events (see Table 8.1).

8.2 Blood flow response methods: PET and fMRI

The first type of technique encompasses two modalities: positron emission tomography (PET) (tomography is the process of forming cross-sectional images or maps of activity corresponding

Table 8.1 Techniques for functional brain imaging

Detection method	PET	fMRI	EPI	EEG	MEG
Source of measurements of neural activity	Local increase of radioactive oxygen carried by blood flow	Disturbance of imposed magnetic field by changes in deoxygenated blood in active brain, observed by detecting MRI signal from protons in tissue water	Very fast version of fMRI	Direct measurement of electrical activity of synchronized neuron groups, with electrode arrays attached to the scalp	Detection of generators of neural activity by magnetic coils in an array round the subject's head
Temporal resolution	2 minutes	4–8 seconds, detecting events over 0.2 sec	As for fMRI	Fractions of milliseconds, but averaging reduces this	Fractions of milliseconds, but averaging reduces this
Spatial Resolution	5 mm	1–3 mm	As for fMRI	≥ 10 mm or more, depending on number of sources	A few millimeters, with careful computation
Restrictions of participant's activity and environment	Participant in prone position on bed, head firmly immobilized	Participant in prone position on bed, head firmly immobilized	As for fMRI	Immobile in comfortable chair	Immobile in comfortable chair. Enclosed in chamber of scanner
Stimuli or response conditions	≤ 12 tracer injections, one condition each	Several stimuli or response conditions applied while subject is scanned for 15–20 minutes	As for fMRI	Nil	Nil
Advantages	Silent	Comparatively good spatial and temporal resolution	As for fMRI	Silent and non-invasive. Excellent temporal resolution allowing discrimination of rhythmic components over a wide range. Relatively cheap and available	Silent and non-invasive. Most direct view of brain activity. Expensive and not widely available
Disadvantages	Potentially toxic tracer substance. Poor spatial and temporal resolution	Loud noise of magnetic coils. Overcome by intermittent 'sparse scanning' with coils periodically switched off	As for fMRI	Problems of computation to locate neural activity. Low conductivity of the skull makes the extraction of detailed information about the generators difficult. Needs MRI to correct for individual brain anatomy. Complex processing to interpret data	Problems of computation to locate neural activity. Needs MRI to correct for individual brain anatomy. Complex processing to interpret data

to slices of the brain) and fMRI (Turner and Jones 2003). These two modalities involve very different physical principles. PET tracks radioactive decays of rapidly decaying positron-emitting nuclei in radioactive compounds injected into the bloodstream that make their way into the brain. Most PET studies of brain function have used water molecules labelled with the isotope oxygen-15, which decays into the stable nucleus nitrogen-14, emitting a positron in the process. The positron quickly comes to rest and mutually annihilates with an electron. Before annihilation, the positron travels an average distance of about 50 mm, which sets the intrinsic limit for the spatial accuracy of PET. The annihilation event produces two high-energy photons that travel (at the speed of light) away from the emission site in opposite directions. Rings of detectors register these photons, recording the time and position of detection. The nearly simultaneous detection of the two photons identifies them as originating from the same positron annihilation event. A mathematical algorithm converts these data into a map of average cerebral blood flow (CBF) during the oxygen-15 half-life of 122 seconds, the time by which half of the nuclei has decayed. In brain-mapping experiments, volunteer participants are given up to 12 injections of the relatively harmless oxygen-15 labelled water, and their brains are scanned while they are presented with a stimulus or task (the experimental 'condition') to test such factors as their awareness, thoughts or memories. Local changes of CBF due to variations in neuronal activity can be inferred using precise statistical techniques that provide parametric maps of the location and intensity of activity. These maps have a spatial resolution of about 5 mm at best, and due to demanding radiation dose restrictions, it is usually necessary to average results from several volunteers to obtain significant results in brain-mapping studies. This limits the sensitivity of the method for immediate, labile and idiosyncratic experiences.

PET has the great advantage for studies of music and human brain that it is entirely silent in operation, by contrast with fMRI (see below). Its main disadvantages are the requirement of injections of a radioactive tracer (limiting any given volunteer to only one set of 12 scans in a lifetime), the relatively poor spatial resolution, and the poor temporal resolution of about two minutes, which is the time it takes to hear six verses of an average song, or to read this paragraph at least six times.

Functional magnetic resonance imaging (fMRI) is based on magnetic resonance imaging (MRI). This widely available medical imaging technique provides images of the tissue distribution of hydrogen nuclei (protons) in water molecules in any organ. In the presence of a very uniform, large, steady magnetic field, the protons can be resonantly excited by a radio-frequency magnetic field of appropriate frequency. The frequency of precession of these spins can be controlled by imposing gradients of intensity in the magnetic field, generated by 'gradient coils' driven by large rapidly switched currents. These magnetic field gradients allow the distribution of the water molecules to be mapped. The tiny voltage induced in a receiver coil by the precessing protons is detected with sensitive radiofrequency electronics, and converted into a high-resolution image of a chosen slice of tissue. The MR image intensity depends mainly on the density of water protons, and is modulated by intrinsic local properties of the tissue, such as biochemical composition and magnetic heterogeneity.

The different properties of different human tissues is of particular importance for imaging effects in brain tissue. The deoxyhaemoglobin, found in venous blood, has a greater magnetic susceptibility than tissue and oxyhaemoglobin; that is, if a person's head is inside a large magnet, the net magnetic field in the capillaries and small veins of the brain will be very slightly larger than anywhere else in the head. This disturbs the otherwise uniform magnetic field, causing a decrease in the MR image intensity around the smallest veins. The venous deoxyhaemoglobin thus acts as an endogenous paramagnetic MRI contrast agent. Any intervention that changes the oxygenation of venous blood can thus be observed as a change in MR image intensity.

Using a very fast MRI technique known as echo-planar imaging (EPI), these changes can be followed second by second, and thus the technique can be used for tracking changes in brain activity. Focal increases in neuronal activity give rise to increases in blood flow that surpass changes in oxygen consumption, so that the net oxygenation of blood leaving an area where neurons are more active is increased. Thus, this venous blood becomes more similar in its magnetic properties to the surrounding tissue, the magnetic field becomes more uniform, and there is a local rise in MR image intensity, the so-called BOLD (Blood Oxygenation Level-Dependent) response. At typical MRI static field strengths of 1.5 Tesla (a measure of magnetic field intensity), this increase is about 4 per cent of the total image intensity, and is thus quite easily observable. Because the level of blood flow accurately reflects locally increased neuronal activity, and changes in oxygen demand correlate well to these electrical changes, fMRI can provide quite accurate localization of neuronal activity. However, the change due to blood flow (the haemodynamic response function) is slow compared with that of the neuronal activity that causes it. Typically, it takes 4–8 s to build up and decay, compared with neuronal time constants of a few tens of milliseconds. Nevertheless, it is possible to sample the entire brain with a spatial resolution of 3 mm in about 3 s. Studies are usually conducted by presenting participants with several successive tasks or conditions while they are continuously scanned for periods of 15–20 minutes, and estimating the effects statistically.

The great advantages of fMRI for studies of music and brain are the relatively high spatial resolution (3 mm or better), the reasonable temporal resolution (a few seconds), the high sensitivity over wide areas of the brain (it takes only a few minutes of scanner time to record activity in many brain areas), the complete non-invasiveness (allowing for an indefinite number of repeat scans), and the wide availability of MRI scanners. The main disadvantage is the quite loud noise produced by the vibration of the gradient coils as their current is varied. Though the sound is repetitive, it can still be distracting and can mask specific frequencies in musical stimuli. Methods for shielding this noise from participants' ears during auditory presentation are far from perfect, although active noise cancellation methods have shown promise. One method of reducing this problem is known as 'sparse scanning'. Because of the slowness of the haemodynamic response, it is feasible to increase the time delay between scans by up to 8 s, presenting the auditory stimuli during the periods of silence. When the brain is scanned, the blood flow and BOLD response to the auditory stimulus are only just reaching their maxima and are thus easily detectable. Sparse scanning has been the mainstay of much of the fundamental imaging research on brain auditory response.

A further disadvantage of fMRI, compared with MEG and EEG (see below), is its low temporal resolution. Brain events can be located accurately in time only to about half a second, a very long time compared with typical cortical processing and transit times between brain areas that are usually in the range of a few to tens of milliseconds, and in comparison with the timing of significant events in music.

8.3 Electrophysiological methods: EEG and MEG

The most important electrophysiological methods are electroencephalography (EEG) and magnetoencephalography (MEG). MEG and EEG use detectors with high temporal resolution, capable of recording signals with a resolution of a fraction of a millisecond. MEG and EEG signals are direct measures of brain activity, but relating these measures to their underlying neuronal generators requires elaborate computation. Before discussing the difficult aspects of interpreting MEG and EEG signals, we describe three properties that ease the burden and provide useful data about how these signals should be interpreted.

First, the signals propagate from the source in the brain to the detector at the speed of light, so the data recorded are, for all practical purposes, the consequences of instantaneous changes in the brain's electrical activity. Second, the detected signal is normally well above instrumental noise, generated, for example, by thermal fluctuations. Since the activity of a single or a few neurons is too small to produce a detectable signal, it follows that the observed changes in electrical activity are produced by sources with compact spatial organization and near-synchronous timing. Third, raw signals can be naturally analysed into many components. These include slow components with time constants of many tens to hundreds of milliseconds, and fast components lasting just a few milliseconds (Ioannides *et al.* 2005).

Typically, MEG and EEG studies have used high- and low-pass filtering of the signal, and the averaging of many single trials, which removes the signal components arising from fast, intermittent activity. Thus, the results reflect only the slow components in the signal that are closely time-locked to external stimuli. Better analysis of non-averaged single-trial MEG data has demonstrated that both slow and fast intermittent activity corresponds to generators in brain areas known by other methods to be activated by the tasks and stimuli used in experiments (Ioannides 2001). In short, both the properties of the signal and evidence from localization studies from single-trial MEG data suggest that the generators at each point of time are both focal and sparse. They must be focal, because only nearby generators can be sufficiently spatially (and possibly temporally) organized to produce coherent electrical activity. They must be sparse, because the brief activation of these focal generators is consistently seen across single trials. They occur in highly variable sequence from trial to trial, yet they lead to an overall signal with well-defined temporal structure. Recent publications based on possibly related observations on raw EEG data (Freeman and Holmes 2005) and high density arrays of electrodes placed on primary sensory cortices of animals (Freeman 2005) provide intriguing insights that may be particularly important for advancing our understanding of how music excites the brain.

To relate the measured signal to what is happening in the brain, one must solve two problems, each with its own difficulties. The first is the 'forward problem', from source to signal, i.e., the computation to identify the signal generated by a given source configuration in the brain activity. The relevant physics is well-understood, involving only the classical laws of electromagnetism and the details of the coil or electrode arrangement making up each MEG and EEG detector. However, the details for the conductivity of the intervening space between generators and detectors must be taken into account: these can be messy, as they usually are when biological structures—such as the head, with its bones and cavities—are involved. In this respect, MEG has a distinct advantage over EEG. The MEG signal detected by a given sensor can be computed remarkably accurately with simple models of the generators, e.g., a sphere model with the centre extracted from the curvature of the inner part of the skull close to the sensor. For EEG, however, a very accurate definition of the conductivity is required throughout the head.

The second problem is the 'inverse problem', from signal to source, i.e., the task of reconstructing the generators, given the signal. It has been known for a long time that, in its purest mathematical expression, the bioelectromagnetic inverse problem has no unique solution (von Helmholtz 1853). However, in practice, given the sparse nature of the generators, a reliable estimate of where they are can be obtained using very broad and reasonable constraints on the assumed nature of the generators, specifically limiting the maximum strength and maximizing the information extracted from the signal (Ioannides *et al.* 1990; Taylor *et al.* 1999). It is, in fact, possible to recover tomographic estimates or maps of activity from individual timeslices of single-trial MEG data (Ioannides 2001; Ioannides *et al.* 2005). These single-trial tomographic solutions have revealed that the brain activity excited by the perception of stimuli is very dynamic. Responses to repeated presentations of identical stimuli differ from trial to trial, even for very basic sensory stimuli, as we have repeatedly demonstrated for the auditory (Liu *et al.* 1998), somatosensory (Ioannides *et al.* 2002a) and visual (Laskaris *et al.* 2003) modalities.

The results demonstrate that, even for simple sensory stimuli, the route from the peripheral nerves to the cortex involves competition within systems of different neuronal networks, some of which may be beyond the reach of available imaging techniques. However, the range of responses for a given stimulus is often sufficiently limited to allow the classification of responses in one area, and to separate out the nodes and interactions of at least part of the underlying system (Ioannides *et al.* 2002a). The averaged data represent a mixture of different processes, even for individual subjects. In contrast, detailed single-trial analysis reveals not only common loci of activity across subjects, but common sequences of interactions in individuals, both for normal processing (Ioannides *et al.* 2002a) and its changes in pathology (Ioannides *et al.* 2004).

In terms of localization, recent studies have demonstrated that tomographic reconstructions of MEG data and post-reconstruction statistical analysis can recover the generators close to the cortical surface with an accuracy of a few millimetres (Moradi *et al.* 2003). Even for deep sources, as long as they are distant from the centre of the head, the generators can be recovered well enough to distinguish, for example, the gaze centres on either side of the brainstem (Ioannides *et al.* 2004). Tomographic reconstructions of single-trial MEG data and post-reconstruction analysis produce an extraordinary wealth of information about the brain. That said, the resources required for the computation and storage of the data are huge, and this limits most investigations to the study of just a few subjects.

8.4 **State of the art, and problems of implementation**

Given the uncertainties outlined above, we believe that there is a need for improved standards of research planning and the more circumspect interpretation of data. We need better models and better methods of analysis. The analysis of MEG and EEG data continues to rely on oversimplified assumptions, partly because of the computational demands of single-trial tomographic analysis and partly for historical rather than scientific reasons (Ioannides 1995). Even today, most MEG studies use equivalent current dipoles (ECD)—point source models—to describe brain activity (a current dipole is completely described by six parameters: three to establish its position within the head, two to define its orientation and one to define its strength). Usually one or more ECDs are fitted to MEG signals that have been heavily filtered and averaged. Following the successful use of baseline conditions in PET and fMRI, many electrophysiological studies have localized ECDs from difference signals between the averages of different conditions. Admittedly, some interesting results have been reported using this approach, as we will describe below. However, describing brain activations elicited by complex linguistic and music stimuli by established methods becomes increasingly dubious and difficult to justify conceptually, as long as the methodology relies on averaging the signals, grand-averaging across subjects, obtaining differences of averages between conditions, and modelling with point-like sources.

As we have mentioned, EEG and MEG measurements are completely silent and non-invasive. The measurements can be repeated as often as a participant is willing to put up with the inconvenience of relative immobility in a comfortable chair. EEG is widely available and cheap to run, but the low conductivity of the skull makes the extraction of detailed information about the generators difficult. MEG provides a more direct view of the brain, but it is expensive and not widely available. MEG and EEG on their own access only functional information. Additional information must be provided from a separate MRI scan for each individual's brain anatomy, to show the generators in the context of the background anatomy at whatever spatial accuracy EEG or MEG can provide. The direct relationship to neuronal activity is obviously a desirable property, but it also has drawbacks.

In the case of music, where so much of the brain is excited, many different neuronal networks are in play in response to the music at the same time, so separating out distinct contributions requires complex processing. The non-uniqueness of the inverse problem turns out to be less of a problem in practice than in theory, at least for MEG. However, the use of simplified models to define the generators poses a major problem in interpreting many of the results. These models are founded on assumptions that the time course of evoked events will be repeated, and that the generators of detected effects are fixed and focal in nature, conditions that clearly are not realized in the brain. Until recently, the computational demands of more robust and powerful methods limited their use to a few laboratories. Today, the computational disadvantages are reduced, but they are still serious enough, and reliance on simpler and less reliable methods of analysis is still prevalent.

8.5 **Gathering the threads thus far**

Much effort is currently devoted to combining haemodynamic techniques (to study the forces involved in the circulation of blood) and electrographic techniques to interpret evoked changes in the brain. The most straightforward way of doing this is to use the MEG and/or EEG signal to recover the time course of regional sources after constraining them to the locations where foci of activity have been identified in the same or similar experiments by fMRI. However, the different techniques rely on very different mechanisms, and provide mappings of events spanning very different timescales. It is all too easy to combine different techniques without caution, and to end up with a result that is determined primarily by the limitations rather than advantages of each method. There is much to be said for advancing each technique as far as possible to reap its fullest benefits before combining their data, as has recently been demonstrated (Moradi *et al*. 2003).

8.6 **Talk and music**

Before we look at the correlates of components in musical stimuli with detected brain activity, it is worth considering the relationship of music and spoken language. Recent results from electrophysiology suggest common processes for syntactic (rhythmic and dynamic) motivating features of music and speech (Besson and Schon 2001) and for aspects of pitch perception. In contrast, some imaging studies suggest that certain features of musical perception are strongly lateralized in a different way from those identified for language perception, as has long been claimed by neuropsychologists (Wallin 1991). We will return to this controversy, since it may provide valuable information on the innateness or developmental origin of human musicality and its relationship to acquired musical skills.

Both music and language are generated in time and contain basic units or elements of intentional action that are planned in time. Cognitive theory supposes that these elements must first be defined or recognized, organized according to rules, and integrated into a ‘phrase’ that is correct both in terms of grammatical and semantic rules. However, all of these features of expressive organization in communication by music and speech (or text) are effects of the intrinsic prospective control of movements by the brain, and their perception is part of the experience of acting in motivated ways (Lee and Schögler, Chapter 6, this volume).

For the purposes of the current chapter, we consider words (or syllables) and music notes (or chords) as the perceived elements of language and music, respectively. Much has been written on the similarities and differences between language and music, as controlled processes of human expression for communication. As we will see, much of the evidence about the similarity of the recognition of organized form in music and speech suggests a partially shared ‘syntactic production/analysis’ apparatus. Evidence presented in Part 2 of this volume shows that use of musical forms of production and reception in communication precedes language in development, and lesion studies suggest at least some degree of dissociation between the motivations for music and language in the human brain (Peretz 2002). However, much of the neural apparatus is likely to be shared, as it relates to the regulation of all body actions and their monitoring in all sensory modalities: it has been strongly argued by Besson and Schon (2001) and others that a prerequisite for understanding how our brains deal with language is first to understand how they deal with the dynamic intentions and the narrative power of music.

In relation to the question of innate musicality in humans, this area of discussion presents another twist. If, as is claimed by Chomsky and his followers (e.g., Pinker 2000), language learning were quasi-instinctive for humans, and if it were true that music and language occupy much the same brain areas, one might argue a fortiori that musical production and perception, or the impulses for their development, are also quasi-instinctive, perhaps representing a critical step in the process of leading to language acquisition after infancy, and indeed in human evolution (Dissanayake Chapters 2 and 24, Cross and Morley Chapter 5, this volume).

8.6.1 PET and fMRI studies

In one of the few direct comparisons between language and music-like stimuli using functional brain imaging, Binder and colleagues used fMRI to contrast passive and active listening to words and tone sequences (Binder *et al.* 1996). In the passive condition, participants heard in alternating periods only the background scanner noise or the background scanner noise with English words and random tone sequences. In the active condition, participants performed either a semantically based word decision or a tone-pattern analysis. The study identified several left hemisphere cerebral areas that respond more strongly to word conditions, including the superior temporal sulcus, middle temporal gyrus, angular gyrus and lateral frontal lobe. In contrast, the planum temporale responded equally well to tones and words in the passive listening condition, and more strongly to tones during active listening. These authors concluded that the planum temporale is likely to be involved in early auditory processing, while specifically linguistic functions are mediated by multimodal association areas distributed elsewhere in the left hemisphere. Parenthetically, Tzourio-Mazoyer *et al.* (2002) provided evidence that the traditional language areas—Broca's and Wernicke's areas—are not uniquely specialized for mature speech and language. These two areas were activated when 2-month old babies looked at pictures of a woman's face. The proficiency of 2-month-olds in 'protoconversation', which employs the same parts of the body as are used in adult speech and hand sign language, also demonstrates that these parts of the brain are active in regulation and perception of forms of vocal expression, even the prelinguistic ones that infants use.

The best-known proponent at present of a music-specific lateralization of brain function based on neuroimaging studies is Robert Zatorre. Most of his work has been done using PET. The musical tasks performed during scanning include pitch judgements within melodies (Zatorre *et al.* 1994), imagined imagery for tunes (Halpern and Zatorre 1999), and the reproduction of tonal rhythm patterns. From these results, he argues that tonal processing, an essential ingredient of music, but not necessarily of the textuality of highly practised and/or literate language, is predominantly a right-hemispheric task. However, since laterality in these studies was inferred by the questionable technique of visually examining arbitrarily thresholded z-maps, rather than by making an explicit statistical comparison between homologous regions in each hemisphere, these findings are not definitive.

By contrast, a recent fMRI study showed greater activation in left-hemispheric primary auditory cortex (Heschl's gyrus) for monaurally presented pure tones that are sinusoidally modulated at 5 Hz (Devlin *et al.* 2003). This study used a robust method for assessing laterality, in which a lateralization index was repeatedly calculated for several different significance thresholds. In another fMRI study, Levitin and Menon (2003) compared listening to normal musical excerpts with listening to scrambled versions of the same pieces, which differed mainly in their temporal coherence. They found activation in areas generally accepted to be related to language perception, and also in their right-hemisphere homologues, with no clear lateralization. A further fMRI study by Koelsch *et al.* studied the cortical response to unexpected musical events that disobey normal syntax, such as abrupt changes in key. These authors found widespread areas activated, with good overlap with traditional language areas, including Broca's and Wernicke's areas, but with no distinct lateralization (Koelsch *et al.* 2002).

Summarizing recent data, it appears that homologous areas in both hemispheres can carry out some of the processing required for musical perception, while left-hemisphere areas have become specialized for acquiring control of articulations of language during development. Interestingly, there are many accounts (Abo *et al.* 2004; Yonemoto 2004; Blank *et al.* 2003; Calvert *et al.* 2000; Woods *et al.* 1988) of right-hemisphere structures—anatomically symmetric to normal left-hemisphere language areas—that become active during recovery of language capabilities following left-hemisphere stroke. This suggests that not much further reorganization of these areas is required for them to support language production and reception.

8.6.2 EEG and MEG studies of language and music

Studies comparing music and language with MEG and EEG have demonstrated that the evolution of effects occurs through distinct time-shared stages. We consider three temporal stages of processing (Table 8.2). We acknowledge from the outset that these stages are convenient ways of summarizing the probable cascades of activity, some of which will straddle the boundaries of the stages as we define them; we also recognize that many subdivisions of timing are possible within each stage.

The first stage is completed soon after 100 ms; it corresponds with what is conceived as the initial processing of the physical properties of the sound that make up the linguistic or music element. The second stage extends from the first stage to around 200 ms; it deals with (semi-) automatic processing of rules that link the element just analysed to an online, tentative framework for the grammatical structure of the yet-to-be completed phrase. The third stage reflects processes that integrate the element with the surrounding context or require a previous sentence parcellation to be re-evaluated. The third stage covers latencies after 250 ms; it must be analysed further into three parts—one for latencies between 250 and 400 ms, one for latencies around 400 ms, and one for latencies around 600 ms. The separation of processing into temporal stages for the different processes of music and language is not surprising when one sees that all human motor activity shares the same fundamental hierarchy of rhythms (Trevarthen 1999). Thus, beyond the analysis of the physical properties of a stimulus, the timing of the successive stages of 'processing' matches nicely with different 'step rates' in walking or thinking, from presto to andante, or alternatively to the duration of a single act and its prospective perceptual monitoring. It has been argued that the thinking process itself—the 'mindness' created by brain function—is a by-product of the evolutionary embedding of motor activities and their associated rhythms (Llinas 2001).

We will next discuss briefly the main conclusions regarding the similarities and differences in language and music processing for each of these stages, focusing, in chronological order, on the work of groups that have studied extensively both linguistic and music responses with MEG and EEG. The first significant contributions were from EEG in the linguistic domain, beginning with the discovery of the N400 component by Kutas and Hillyard (1980; 1984). This is a centroparietal

Table 8.2 Temporal stages of processing for music and language

Processing times	**Early components, < 250 ms**		**Late components, > 250 ms**		
	< 100 ms	**100–200 ms**	**250–400 ms**	**~ 400 ms**	**~ 600 ms**
Speech	Initial processing 40–100 mset al.(P1m); in auditory cortex	Early left anterior negativities (ELAN) following syntactic incongruity or 'phrase structure violation' Mismatch negativity (MMN) for a change in a repeated auditory signal .	Left anterior temporal negativity (LAN) for incongruities in syntax/grammar	N400 centro-parietal negative event-related potential to unexpected words and semantic errors. Associated with semantic integration (knowing what the thing is) N400 priming effect for language, in posterior middle temporal gyrus. N400 for incongruent words sung in tune	P600 a positive centro-parietal deflection after the onset of an unexpected word. Associated with syntactic rather than semantic processing (knowing where it and oneself is). Re-evaluation of earlier processing, checking meaning in a sentence
Music sounds	Initial processing 40–100 ms (P1m); in auditory cortex	Early right anterior negativities (ERAN) following unexpected chords, and right MMN for a change in a repeated chord	Right anterior-temporal negativity (RATN) for incongruous musical sound	N400 priming effect for music, in posterior middle temporal gyrus	P600 for musical incongruities and for correct words sung out of tune
Movements of musical performance	Trills	Vibrato, arpeggios	Presto pulse/beat		Andante pulse or beat

negative ERP (event-related potential) component; it is elicited by unexpected words, especially semantic errors, beginning around 250 ms and peaking around 400 ms, hence the name. The N400 corresponds to the second part of our third stage, and is associated with semantic integration (knowing what the thing is). The P600 is a positive deflection that, like the N400, is over centroparietal electrodes and starts fairly early after the onset of an unexpected word, but that peaks later, around 600 ms (Osterhout and Holcomb 1992). The P600 corresponds to the third part of the third stage and is associated with syntactic rather than semantic processing (knowing where it and oneself is). The P600 is strong when a re-evaluation of earlier processing is required, such as the following of 'garden-path' sentence structure.

Except for some early work by Besson and Macar (1987), until recently there were relatively few studies of evoked effects of music incongruity (Besson *et al.* 1994; Janata 1995). The first study that directly compared music and language claimed that linguistic and music incongruities elicited positivities around 600 ms that were statistically indistinguishable for both language and music (Patel *et al.* 1998). In the same paper, an earlier music-specific ERP component was observed that showed anterotemporal right-hemisphere lateralization, termed right anterior-temporal negativity (RATN). This component resembled the left anterior negativities (LAN) reported earlier for language-related syntactic processing (Friederici 1995). The LAN and RATN correspond to the first part of our third stage.

The similarity between late evoked components elicited by music and language has been reported in a series of recent studies. In Besson and Schon (2001), 200 excerpts from French operas were used with either the last word sung in or out of tune, or the last word replaced by a semantically incongruent word sung in or out of tune. The results demonstrated a N400 for incongruent words sung in tune and a P600 for correct words sung out of tune. The double incongruity (incorrect words sung out of tune) elicited both an N400 and a P600, not significantly different from the sum of the effects observed for each incongruity separately. The near additivity of these effects suggests that, despite the similarity in language- and music-induced N400 and P600 effects, relatively independent neural processing underpins the semantic aspects of language and the harmonic aspects of music.

In Koelsch *et al.* (2004), behavioural measures and the N400 ERP component elicited by a visually presented target word were used to quantify the priming effect of either semantically related or unrelated preceding sentences or music. The semantic relationship between words and music was chosen either on the basis of self-reports of the composers or was based on musicological terminology—for example, the musical prime for the word 'narrowness', used intervals set in closed position (covering a narrow pitch range, and being dissonant), and the prime for the word 'wideness', used intervals set in open position (covering a wide pitch range). The N400 priming effect was similar for language and music in terms of time course and strength. ECD (equivalent current dipole) localization of the difference signal ('semantically related' subtracted from 'semantically unrelated' evoked potentials) produced very similar results for both language and music, localizing the source in the posterior portion of the middle temporal gyrus.

Early anterior negativities (< 250 ms) following syntactic incongruity, in the front half of the brain, have been labelled ELAN and ERAN, depending on whether they are recorded primarily over the left or right hemisphere, respectively. The ELAN was first identified in linguistic syntactic violations, e.g., following phrase structure violation (Friederici 1995). The localization of the magnetic correlate of these decreases in EEG signal, using tomographic analysis of MEG data with limited head coverage (Gross *et al.* 1998) and with ECD analysis from full-head coverage (Friederici et al. 2000), showed early activity in both left and right auditory cortex and in left and right inferior frontal cortex, corresponding to Broca's area and its right-hemisphere homologue. These studies were followed by experiments using a sequence of five chords, introducing unexpected variation (Neapolitan chords) in the third and fifth chord to study the music analogue of the linguistic syntactic violation. The unexpected chords elicited anterior *right* negativities, ERAN, at the same latency (around 200 ms) as the linguistic violation effect on the *left* (ELAN), followed by a *bilateral* negative signal, greatest around 500–550 ms. The ERAN was identified with both EEG and MEG, but the later signal was evident with only EEG. ECDs were localized in the middle of the Heschl's gyrus bilaterally from the average MEG signal of the in-key chords around 200 ms, a location within or near the primary auditory cortex (Maess *et al.* 2001). In the same early time period, the ECD were localized in more anterior regions around Broca's area and its right-hemisphere homologue, using the same stimulus as for the ERAN signal, the difference between the Neapolitan chords and the in-key chords (Maess *et al*. 2001).

To fully appreciate the ELAN and ERAN results, we must contrast them with the well-studied 'mismatch negativity' (MMN) (Naatanen 1992). MMN is elicited by a change ('deviant') in a repeating auditory stimulus. It is an 'automatic' response, since it can be elicited even when participants' attention is engaged in a different task. MMN is seen as a negative ERP wave between 100 and 200ms in the difference signal between the deviant and standard ERP waveforms. A direct comparison between encoding of elementary musical and phonetic sounds was made using MEG recording in an MMN paradigm (Tervaniemi *et al*. 1999). It was found, only within the right hemisphere, that the MEG counterpart of MMN (MMNm) elicited by an infrequent *chord* change was stronger than the MMNm elicited by a *phoneme* change. The ECD model was used separately for phonemes and chords to describe the early MEG signal (the increase in signal occurring between 40 and 100 ms, termed P1m) and the MMNm (the difference signal between frequently repeated and deviant items). The MMNm ECD localizations differed slightly for phonemes and chords, and more noticeably for binaural stimulation, but for P1m did not differ significantly for phonemes and chords. However, a PET study using very similar stimuli showed an even more lateralized response, suggesting a left-hemisphere specialization for the processing of phonetic stimuli and a right-hemisphere specialization for the processing of chords.

The traditional interpretations of ERAN and MMN results have been rather different. MMN is thought to be an index for the neural traces of short-term auditory memory, which must be related to ongoing awareness of the environment, while ERAN is proposed as an index of early syntactic processing, which will be more concerned with prospective control of participants' actions. However, a case can be made for the equivalence of ERAN and MMN. Both ERAN and MMN can be elicited automatically by deviant auditory stimuli. Evidence for anticipatory auditory cortex activity within a sequence of tones with variable ISI (inter-stimulus interval) was recently presented (Ioannides *et al*. 2003). If the auditory system cannot easily distinguish anticipatory neural activity from the memory trace set up by a repetitive stimulus, then much of the argument for the distinction between MMN and ERAN, and between 'memory trace' and 'prospective perceptual control', evaporates. For now, we note that the shared properties for ERAN and MMN include an overlap in latency and similarity in the spatial distribution of the corresponding ERP components. The evidence thus far cannot eliminate the possibility that additional processes, specific to syntactic 'cognitive' processing, give rise to ERP components in ERAN that are not present in MMN (Koelsch *et al*. 2001).

The few studies that looked at the processing of sound stimuli before the ERAN and MMN range have reported no differences between the processing of language and music. For example, the ECD localization for the P1m response was not significantly different for phonemes and chords (Tervaniemi *et al*. 1999).

8.6.3 Innateness and plasticity at different temporal scales: results from electrophysiology

The first year of human postnatal life is characterized by enormously actively engaged brain plasticity (Dawson and Fischer 1994; Trevarthen 2004a). During this first year, infants, while intending to communicate, pick up the statistical and prosodic patterns in a culturally formed language input, thus discovering phonemes and words (Trevarthen 2001, 2004b). Through their highly intelligent social interaction with other humans, their speech learning accelerates in a way that has been compared to communicative learning in songbirds (Kuhl 2004), but which differs fundamentally in the shared intentions it serves. Changes in capacity to learn new speech sounds occur more slowly after early childhood, but 'plasticity' for auditory learning is still present in adults. Although the evidence suggests that 'hard-wired' changes leading to new neural 'maps' are slow to occur (requiring activity-dependent changes at the synaptic level), and changes in grey matter (that can be detected by imaging the local anatomy at very high resolution) are slower still, adequately motivated learning can still be fast, notably in 2-year-olds 'one trial' learning of new words (Carey and Bartlett 1978; Halberda 2003. A catchy tune can make an indelible impression the first time it is encountered, and this is true for infants as it is for adults (Trainor 1996; Trevarthen 2002).

In many studies of long-term changes in responsiveness to stimuli, a single ECD is fitted to the averaged MEG or EEG data, and the derived dipole strength is regarded as a quantitative measure of the plasticity effect. However, it is not always clear whether the increase in ECD strength arises because the real-time activation is stronger in each single trial, or because the single-trial activity is better 'organized' (e.g., better time-locked to the external stimuli). As we have seen, some studies that involve the exposure of naive participants to sequences of tones (and presumably music) for hours or days have tended to use real-time measures. These studies point to the reorganization of background rhythms in distributed cortical representations, especially in the *gamma* band of frequencies (~ 40 Hz), as the most likely change that correlates with performance (Bosnyak *et al.* 2004).

Some indication of the specificity of innate neural competence for musical perception, as compared with language perception, is provided by the intrasurgical studies of Chauvel and colleagues. During open-skull procedures with conscious patients, they intracerebrally recorded the auditory evoked potentials from the primary and secondary auditory cortices of both hemispheres of the human brain, in response to syllables or tones. In one study, they showed that the response to the acoustic elements of syllables is lateralized. The effect of voiced and voiceless syllables was distinguishable in the left, but not in the right Heschl's gyrus (HG) and planum temporale (PT); only the evoked potentials in the left HG, and to a lesser extent in PT, reflected a sequence of responses to the different components of the syllables. This acoustic temporal effect in the brain was not limited to speech sounds, but applied also to non-verbal sounds mimicking the temporal structure of the syllable (Liégeois-Chauvel *et al.* 1999). In a second study, clear spectrally organized tonotopic maps were observed (that is, spatially organized according to gradations of sound frequency), with distinct separations between different frequency-processing regions in the right hemisphere. In contrast, tonotopic organization was less evident in the left hemisphere, where different regions were involved in response to a range of frequencies (Liégeois-Chauvel *et al.* 2001).

Ioannides and colleagues (2003) used MEG to compute the so-called echoic memory trace (EMT)—the short-term retention of stimulus-related information. They used trains of auditory stimuli designed specifically for computing the EMT lifetime and its contextual sensitivity, and observed time-dependent EMTs with different lifetimes at different latencies, suggesting the existence of multiple neural delay lines. The longer EMT lifetimes showed handedness and gender-dependent interhemispheric asymmetry. Specifically, all participants except left-handed males showed longer EMT lifetimes in the left hemisphere. These EMT results, together with the results of Liégeois-Chauvel *et al.* (2001), suggest that very fundamental properties of the auditory system, likely to determine how language and music are first perceived in the brain, are differently established in the left and right hemispheres of the adult human brain.

8.6.4 **Conclusions**

The differences between electrophysiological and imaging studies are not easy to reconcile. Patel (2003) makes a valiant effort in this regard by proposing a common neural substrate for music and linguistic 'syntactic processing', possibly in the frontal lobe and related to the production of speech and musical sounds. He proposes that this neural substrate operates on distinct linguistic and musical representations of experience, very likely localized in posterior parietal areas. Clearly, from Chauvel's work, innate or early-maturing networks in primary auditory areas are also involved in bringing about these distinct representations. Our discussion of the similarities between MMN and ERAN point to a similar conclusion to that of Patel, but with more emphasis on the role of anticipatory mechanisms.

This picture is consistent with the evidence that a newborn human is already music-competent, with specific areas in both hemispheres capable of interpreting structured sequences of sounds—especially those with distinct rhythms linked with the body rhythms of cardiac pulsation, respiration, gestural movement, and walking—and that this shared rhythmic sense, or communicative musicality, facilitates the later acquisition of language through the elaboration of expressive communication between babies and intimate companions. Lateral prefrontal areas, which include Broca's area, are innately configured to be involved in the anticipations and predictions that define the syntactic structures of both linguistic and musical expression. It is also clear that the processing requirements specific to language or music, such as the distinctions between very similar consonants needed for semantic precision, require access to early acoustic analysis in the auditory cortex with distinct hemispheric lateralization for different features.

8.7 **Cortical specializations for components of music perception**

8.7.1 **Pitch and melody**

Several careful brain imaging studies have addressed the following questions.

1. What brain areas are selectively involved in processing notes of identifiable pitch as compared with noise?
2. What areas of the cortex are 'tonotopic' (i.e., map tones in a systematic pattern)?
3. What brain areas are involved in the detection of the fundamental musical aspects of pitch—i.e., which octave (pitch height) and what note within the octave (pitch chroma)?
4. What brain areas are concerned with listening to melodies compared with simple repeated note sequences and noise?

There are strong reasons to believe that none of these questions is seriously ethnocentric, or even anthropocentric. From studies of animals, especially monkeys (Morel *et al.* 1993; Brugge 1985) it is clear that tonotopy is a universal feature of the primate brain, and that pitch discrimination is an important component of the recognition of sound sources, with obvious survival value. Pitch, like colour, is a conscious percept, constructed by the brain; this percept bears a relationship to the frequencies that are picked up by the ear, but this is not necessarily a simple relationship—for instance, there is the well-known example of the 'missing fundamental', in which a listener will confidently claim to have heard the fundamental note of a harmonic series of pitches played together, even when the fundamental is physically missing. The studies that address the questions posed above will now be presented.

1. The definitive fMRI experiment comparing pitched sound with noise was performed by Griffiths *et al.* (1998). To avoid the artificial character of a pure tone, and to provide emphasis for the percept of pitch rather than the response to auditory frequency, a stimulus consisting of iterated noise was used, which has a very broad frequency spectrum and can easily be adjusted to give varying degrees of perceived pitch strength. This allowed Griffiths and colleagues to perform a parametric experiment, in which the temporal regularity (equivalent

to the strength of the perceived pitch) was systematically increased, and brain areas were found that showed a corresponding increase. These brain areas were located in the primary auditory cortex (Heschl's gyrus) bilaterally, with no activation in brain areas that are earlier in the auditory pathway (inferior colliculus, medial geniculate body).

2 It is apparent that our auditory cortices provide spatial maps of sound frequency. While this was carefully explored in one of the earlier MEG studies (Pantev *et al.* 1988), the use of fMRI has provided much greater detail (Formisano 2003; Talavage 2004), showing several tonotopic gradients in early auditory areas, consistent with results in owl monkeys (Brugge 1985; Recanzone *et al.* 1993). The implication is that distinguishable sound frequencies are naturally important in human hearing, which enables us to perceive meaningful structures of sounds based on sequences of notes.

3 There is abundant evidence that the perception of octave equivalence (in which the sound frequency is precisely doubled) is a human universal—it is even found in rhesus monkeys (Wright *et al.* 2000), which can generalize tonal melodies (but only tonal) across octave transpositions. Tones within each octave are uniquely consonant with their corresponding tones in other octaves, while within each octave there is a subjectively similar cycle of notes, whatever scale is considered. In an fMRI study, Warren *et al.* (2003) investigated the brain areas involved in perceptual changes along these two dimensions of pitch variation, known as pitch height (which octave?) and pitch chroma (which note within the octave?). In their experiment testing musically untrained adults, the stimuli were harmonic complexes, in which chroma and height could be varied continuously, while the total energy and spectral region remained fixed. This controlled properly for these important auditory variables, leaving the perceptual aspect of pitch as the important experimental variable. They found that change in pitch chroma, but not in height, activates bilateral areas in front of Heschl's gyrus in the planum polare, while change in pitch height, but not in chroma, activates bilateral areas in the posterior planum temporale. While Heschl's gyrus itself is involved in all auditory experience, these regions, which bracket the primary auditory cortex on the superior surface of the temporal lobe, represent distinct brain substrates for processing the two musical dimensions of pitch (Figure 8.1).

Is this specialization of auditory cortex innate or acquired? The participants in this study were not musically trained, but as adults they would already have been exposed to a great deal of mainly Western music, and as we will see later in this chapter, our adaptable brains are capable of permanent change in the process of acquiring many musically related skills. Thus, without performing similar studies of very young infants, with almost no musical experience, this question is impossible to answer definitively. However, the cross-culturally universal perception of the octave interval strongly suggests that there are brain regions that are 'prewired' to be sensitive to this feature of pitch, and there is no reason to suppose that the regions identified as activated in this study differ across cultures.

4 The earliest imaging study to investigate whether there are particular brain areas concerned with melody was a PET by Zatorre *et al.* (1994), in which participants listened to simple melodies and acoustically matched noise sequences. The results showed that cerebral blood flow increased in the right superior temporal and right occipital cortices. The authors concluded that specialized neural systems in the right superior temporal cortex participate in the perceptual analysis of melodies, but as with several other early studies, appropriate statistical methods were not used to evaluate lateralization. These areas can be considered to form part of the pervasive mirror or 'sympathy' system of the brain for all sensory modalities (Decety and Chaminade 2003; Jeannerod 2004).

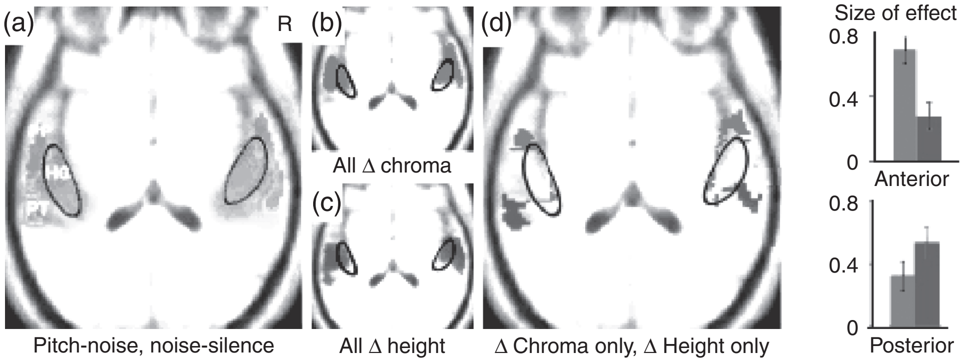

Fig. 8.1 Statistical parametric maps for Warren *et al*. (2003). The 90 per cent probability boundaries for primary auditory cortex are outlined (black). (*a*) Broadband noise contrasted with silence (noise–silence, green) activates extensive bilateral superior temporal areas including both medial and lateral Heschl's gyrus (HG). The pitch-producing stimuli contrasted with noise (pitch–noise, lilac) produce more restricted bilateral activation in lateral HG, planum polare (PP), and planum temporale (PT). (b) Pitch chroma change (Δ chroma) contrasted with fixed chroma (all Δ chroma, red) activates bilateral areas in lateral HG, PP, and anterolateral PT. (c) Pitch height change (Δ height) contrasted with fixed height (all Δ height, blue) activates bilateral areas in lateral HG and anterolateral PT. (d) Voxels (volume elements) in b and c activated both by pitch chroma change and pitch height change have been exclusively masked. Pitch chroma change but not height change (Δ chroma only, red) activates bilateral areas anterior to HG in PP; pitch height change but not chroma change (Δ height only, blue) activates bilateral areas in posterior PT. These areas represent distinct brain substrates for processing the two musical dimensions of pitch. The relative magnitude of the blood oxygen level-dependent (BOLD) signal change in anterior and posterior areas is shown for each of the contrasts of interest (right). The height of the histogram columns represents the mean size of effect (signal change) relative to global mean signal for the contrasts Δ chroma-only (red) and Δ height-only (blue) at the peak voxels for each contrast in the right hemisphere; vertical bars represent the standard error of the mean size of effect. The histograms demonstrate opposite patterns of pitch chroma and pitch height processing in the anterior and posterior auditory areas. (Adapted and reproduced with permission.) (See also colour plate.)

Another study by the Griffiths group (Patterson *et al.* 2002) addressed this question, this time with an fMRI experiment that allowed more detailed investigation to identify the main stages of melody processing in the auditory pathway. Spectrally matched sounds that produce no pitch, fixed pitch, or melody were all found to activate Heschl's gyrus and planum temporale. Within this region, sounds with pitch produced more activation than those without pitch only in the lateral half of Heschl's gyrus. When the pitch was varied to produce a melody, there was activation in regions beyond Heschl's gyrus and planum temporale, specifically in the superior temporal gyrus and planum polare, with significantly more activation in the right hemisphere, in partial agreement with Zatorre's findings. Unsurprisingly, regions in the planum polare are also specifically involved in pitch perception within the octave (chroma). The results support the view that there is a hierarchy of pitch processing in which the centre of activity moves forward and more laterally away from the primary auditory cortex as the processing of melodic sounds proceeds. An analogous hierarchical functional distribution has been identified in visual areas at the rear of the brain, where areas further forward are concerned with progressively more complex features of a visual scene (Zeki 1993).

Again, the question remains whether these brain areas have undergone a special adaptation as a result of experience or whether they are congenitally earmarked for the task they perform. Perhaps cross-cultural studies could be helpful here, if possible involving participants from cultures where far less emphasis is placed on melody in music. It is intriguing to hypothesize that the observed lateralization arises because language processing comes to dominate left hemisphere auditory areas during development, leaving the homologous right hemisphere areas to music by default.

8.7.2 **Timbre**

The term 'timbre' describes the harmonic content of sounds as they evolve through time during the playing of each note (McAdams *et al.* 1995). It is this feature of a sound that allows us to distinguish one resonant source from another when the other perceptual features—pitch, loudness and duration—are held constant. Thus, the perception of timbre is likely to be relatively species-invariant, since it may well be required for potentially survival-dependent auditory identification, and for the detection of expressive effort and level of excitement or emotional modulation in calls.

Changes in timbre are an inherent part of a musical performance, and they play an important role in the emotional effects of music, which we will discuss in more depth later. In a well-designed fMRI study, Menon *et al.* (2002) used selected sound stimuli to investigate the neural correlates of timbre perception. They presented non-musician volunteer subjects with the same melodies performed with contrasting timbres, and compared the resulting brain activations. The two timbres comprised tones with (a) a fast attack, low spectral centroid, and no spectral flux, and (b) a slow attack, higher spectral centroid, and greater spectral flux. Loudness and subjective pitch were carefully controlled. The participants' low-level task was a key-press response at the end of each short melody. The results indicated that both left and right hemispheres are involved in timbre processing, challenging the conventional notion that the elementary attributes of musical perception are predominantly lateralized to the right hemisphere.

Significant timbre-related brain activation was found in well-defined regions of posterior Heschl's gyrus and superior temporal sulcus, extending into the circular insular sulcus. Although the extent of activation was not significantly different between left and right hemispheres, temporal lobe activations were significantly posterior in the left hemisphere when compared with those of the right, suggesting a functional asymmetry in their respective contributions to timbre processing. Apart from Heschl's gyrus activation, which appears to be common to most auditory perceptions, the timbre-specific activation areas appear to be different from those of pitch and melody; they are generally deeper and lower in the temporal lobe, perhaps indicating greater involvement with the emotive effects of the sounds. (For an analysis of the physical information that defines the different emotional intensity of expressive movements, including those of musical performance, see Lee and Schögler, Chapter 6, this volume).

8.7.3 **Rhythm**

A considerable proportion of music in most cultures has definable rhythms. It is probable that the perception of rhythm is mediated by specific, innately predefined brain areas. However, as with melody, cultures have developed highly sophisticated elaborations of rhythmic motifs.

Rhythmic movement and sounds are intrinsically part of animal life—heartbeat and respiration must occur cyclically to give tissue continuous nourishment, and animal locomotion and chewing are repetitive. Indeed, all movements of animals, even those such as jellyfish or worms, without jointed body members, have a cyclic time control or rhythm (Llinas 2001). For all animal movement, increases in physiological arousal are accompanied by increases in frequency, providing music-makers with a uniquely natural means of modulating arousal in their listeners. Rhythmic sounds have a powerful entraining effect on listeners (Molinari *et al*. 2003), encouraging a response in the form of rhythmical movements or dance (Cross and Morley, Chapter 5, this volume).

Primate species share a brain system known as the 'mirror system' (Rizzolatti *et al*. 1996; Iacoboni *et al*. 2001). In a given participant's brain, neurons forming part of this system produce action potentials whether the individual is performing a specific movement, or whether another conspecific is viewed performing the movement. Rhythmical sounds easily evoke in the imagination the movements required to produce them, and thus it would not be surprising if they activate areas of cortex involved in movement generation, specifically the premotor cortex and supplementary motor areas. Perhaps even more pertinent for understanding the brain mechanisms of musical communication and art is the increasing evidence that mirror systems in the brain are conveying information between participants about dynamic emotional states, qualities of purpose expressed in movement, and the goals of actions (Adolphs 2003; Gallese 2001).

There have been few brain imaging studies that address directly the question of what brain regions are specifically responsible for discrimination of rhythms, either in music or in the wider context of regularly repeated actions. Parsons (2001) performed a PET study in which musicians and non-musicians discriminated pairs of rhythms with respect to pattern, tempo, meter, or duration. Among other areas, the cerebellum appears to be a crucial component of the rhythm perception network. Sakai *et al.* (1999) used fMRI to explore the brain areas involved in maintaining the short-term memory of a rhythmical sequence of sounds, while varying the complexity of the rhythm. They found that for simple rhythms, as commonly found in music, activation occurred in predominantly left-hemisphere premotor and inferior parietal areas in the cerebrum, together with right anterior cerebellum. These are areas important in 'mirroring' intentions of movements. Overy *et al*. (2004), in a ground-breaking fMRI study, investigated melody and rhythm perception in children of about 6 years old. By contrast with earlier studies on adults of rhythm discrimination, based on the effect of cortical lesions (Samson *et al*. 2001), this study found no strong cerebral lateralization to the left hemisphere. Overy argues that this lateralization is more likely to develop during maturation.

The production of regular movements, for instance finger tapping, has received much more attention from imaging neuroscientists (Rivkin *et al*. 2003; Dhamala *et al*. 2003, Ullen *et al*. 2003), but space does not permit a full discussion here. (See Lee and Schögler, Chapter 6, this volume, for information on brain systems that are involved in the regulation of the timing of movements and their perceptual control).

8.8 **Temporal aspects of music (and how to study them)**

Music and time are inseparable. Notes, pitch and chords are defined by the natural frequencies of their elements ('real' or 'virtual'; e.g., missing fundamental), and hence they are intrinsically based on the temporal notion of oscillation at a very fine temporal scale (Osborne, Chapter 25, this volume). Melody is defined by the temporal arrangement of musical elements over longer time periods, and melodies are linked together in patterns specific to musical themes, phrases, songs and composers, that span timescales from seconds to a composer's lifetime, i.e., the full range of conscious experience and its recollection. Rhythm, with elaborated narratives of rhythms, provides the defining organizational principle of temporal sequences. It is therefore clear that all elements of music are conncected with unfolding sequences in time.

Brain activity is also characterized by events and oscillations (chemical and electrical) that unfold in time at different timescales. Two ingredients are therefore highly desirable for the analysis of brain activity elicited by music. First, a method is needed to describe in a quantifiable way the temporal attributes of music in all its aspects. It is especially important to allow for structure over wide temporal scales and to include variations expressing the performer's interpretation that are known to have a profound effect on the way the music is perceived. Second, brain activity needs to be described in ways that produce regional time-courses of activity that can be analysed in exactly the same way as music, so that correlations can be found between the temporal organization of the piece of music, its specific expression in a given performance, and the brain activity it elicits. We will emphasize a new methodology that embodies these two ingredients, after summarizing the achievements and limitations of other recently proposed methods that also allow flexibility in the stimulus content.

8.8.1 Conventional analysis, and moves in the right direction

Methods relying on haemodynamic processes lack the temporal resolution to capture the high and moderate frequencies encountered in music. MEG and EEG do have the required temporal resolution, but their conventional use does not lend itself easily to such an investigation. Even studies relying on musical phrases containing a harmonically, melodically, or rhythmically incongruous note or chord fall short of the mark. Although such approaches rely on elaborated cognitive categories and concepts of music theory, they can only probe the natural processes of expectancy at a restricted level, identifying isolated responses to musical violations, rather than probing the responses evoked during the generation of experience of the local and global attributes of an actual piece of music.

The steady-state response (SSR) of individual MEG or EEG channels has been used to study the temporal neural correlates of the auditory sequences in a number of recent studies. This approach provides a measure of the real-time variation elicited by a continuous stimulus, albeit often an indirect one. The evidence from these studies suggests that phase reorganization plays a more important role than amplitude changes, as the two following examples suggest.

Patel and Balaban (2000) have studied the SSR elicited by the amplitude modulation of melody-like sequences. They found that energy changes in the MEG signal did not track the stimulus structure well. In contrast, the phase of the MEG signal at particular sensor locations showed marked resemblance to the contour of the pitch-time series. The distribution of sensors showing this phase tracking of the auditory sequences was bilateral, but with a 'statistical tendency' for higher density in the right hemisphere. A similar approach was followed in Bosnyak *et al.* (2004) to address the question of plasticity in the adult human brain. Adult non-musicians were trained to discriminate small changes in the carrier frequency of 40 Hz amplitude modulated pure tones. Although changes in the amplitude of the transient evoked response were identified, no change in the SSR amplitude was identified, and only changes in the phase of the SSR were observed after training.

Working at the level of the SSR of individual channels limits what can be deduced from the data. For MEG in particular, the signal recorded by any one sensor is sensitive to generators that can be very distant from each other, so changes at the level of sensors will be more apparent if large-scale synchrony is established. Conversely, a single focal generator produces a dipolar radial field pattern; therefore, a strong MEG signal will be recorded at two separate locations (the distance between the sensors of maximum sensitivity increases with the depth of the focal source). Temporal synchronization between areas cannot be reliably concluded on the basis of measures of synchronization measures between sensors at different locations, because each single focal generator will create strong correlation (with no time delay) in the output of sensors at different locations.

Two recent publications exploited the availability of full tomographic descriptions of brain activity, millisecond by millisecond, elicited by continuous music (Ioannides *et al*. 2002b; Popescu *et al*. 2004). For the tomographic analysis, the CURRY 4.5 (Compumedics Neuroscan) source localization software was used, with a minimum L2-norm constraint for the currents and the L-curve regularization. Five right-handed male participants were used, none of whom had formal music training. The music stimulus was a 2 minute 50 second solo piano piece played at a moderate tempo. It had a simple anapaestic metre, i.e., its basic metrical foot consisted of crotchet, crotchet, minim (i.e., ratio of 1:1:2). The piece was initially unfamiliar to the participants, and became familiar through a training procedure that each participant underwent before actually listening to the excerpts used for the analysis. The procedure ensured that the material was neither novel not overlearnt, but rather equally familiar to each participant. The piece was divided into motifs that could be used for memorization and recall. The MEG signal was then recorded while the participants listened to two motifs each lasting 10 seconds (motif I and motif II). The MEG signal was recorded for 20 repetitions of each of the two motifs.

The first study (Ioannides *et al*. 2002b) used the average and single trial MEG signal from the 20 repetitions of motifs I and II. The average MEG signal at the beginning of the motif showed a dipolar pattern on the lateral surface (Figure 8.2a), and the tomographic solutions showed loci of activity in and around the auditory cortices (Figure 8.2b), stronger on the right. The average MEG signal was computed separately for sensors around the maximum and minimum of the dipolar pattern for each hemisphere, using five sensors for each sum as marked by the pentagons in Figure 8.2a. The differences between the averages around the maximum and minimum of each dipolar pattern define a virtual sensor (VS) for each hemisphere. The VSs can be applied to the single-trial MEG signal to obtain a real-time estimate of activity in the left and right auditory and nearby cortices. Figure 8.3 shows that the frequency spectra for the time series representing the sound of music and the activity in the left and right auditory cortex (as described by the two VSs) shared similar peaks, especially on the right, as would be expected by the results of Patel and Balaban (2000).

The second study extended these results in two ways. First, the frequency spectra were computed for regional brain activations extracted from the tomographic solutions. The results showed that musical attributes at different temporal scales are processed in distributed and partially overlapping networks: low frequencies are encountered in networks distributed in the anterior part of the temporal lobes and frontal areas (Figure 7 in Popescu *et al*. 2004). These networks presumably deal with 'higher-order patterns' formed by slow features of the melody. The results showed further that high frequency features corresponding to individual notes are analysed in regions within and around the auditory cortex, as already demonstrated in the earlier study (Ioannides *et al*. 2002b), and in motor areas, specifically in primary sensorimotor area (SM1), premotor area and supplementary motor area (SMA) (Figure 8 in Popescu *et al*. 2004).

A novel analysis was introduced and applied to one of the two motifs to capture the temporal characteristics of the music score and brain activity. For the purpose of this analysis, the motif was divided into four melodic contour segments (A–D); the first three segments (A–C), each lasting a little over two seconds, were analysed in detail. The performance-rhythm of each segment was characterized by temporal deviations from the reference interval ratio (DRIR), introduced either by the performer's artistic expression or the limitation/restriction of his physical capabilities. The mean interval ratios were 1.0:1.2:2.0 in segment A, 1.0:1.3:2.4 in segment B and 1.0:1.1:2.1 in segment C—that is, segment C corresponds best to the anapeastic 1:1:2 interval ratio defined on the score (crotchet, crotchet, minim), segment A comes close to it, and segment B

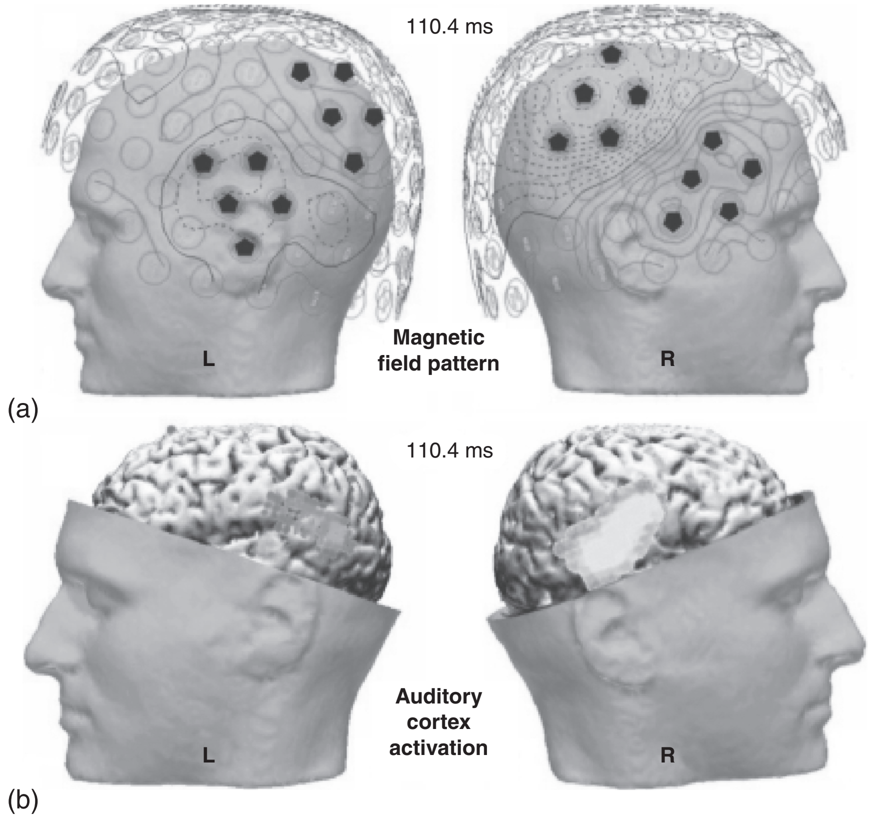

Fig. 8.2 (a) Contour plot of the magnetic field over the left (L) and right (R) side of the head. The signal topography is dipolar, i.e., the pattern separates clearly into positive and negative values corresponding to magnetic field in and out of the plane. The sensors around the maxima and minima of this nearly dipolar pattern are marked by regular pentagons pointing in opposite directions. This signal topography is consistent with a rather focal generator somewhere in the brain in the region between the two extrema, which corresponds to the primary auditory cortex and/or the surrounding auditory association areas. An estimate of the activity around each auditory cortex can be calculated directly from the signal values at each time slice by taking the difference between the averages of the signal recorded for each half of the brain by each set of sensors (each set marked on the figure with pentagons that point in the same direction). (**b**) Activity from tomographic reconstructions of the signal shown in (**a**) are superimposed on rendering of the left and right cortical surface. (See also colour plate.)

deviates the most. The partitioning of the 10-second motif therefore reveals two switches in note-duration ratios that occur during the musical motif: the first switch marks the transition from a close to metrical segment (segment A) to a segment characterized by higher DRIR (segment B), whereas the second switch marks the transition from segment B (with higher DRIR) to a segment with the smallest DRIR (segment C). Full details on the music score and its decomposition can be found in Popescu *et al.* (2004).

The new measure, introduced in Popescu *et al*. (2004), was to quantify the similarity in the temporal structures of music and brain activity through the correlation between

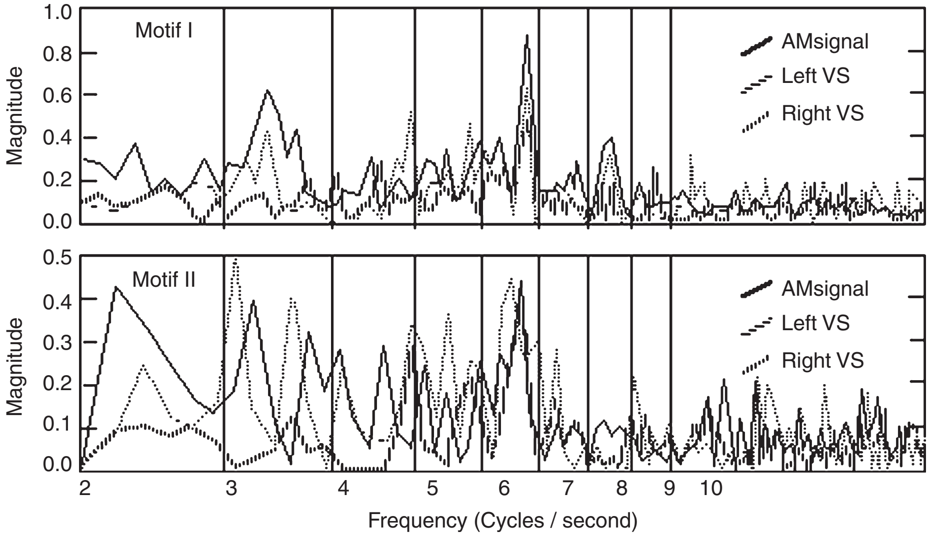

Fig. 8.3 Frequency spectra for the amplitude envelope of two musical motifs and for the auditory cortex activity (derived from the tomographic solutions) show a high correlation in the 7 Hz frequency band, which roughly corresponds to the repetition rate of the piano key sounds. The match at other frequencies is better for the right hemisphere (adapted from Ioannides *et al.* 2002b).

two beat-spectra (Popescu *et al.* 2004). The beat spectrum provides a robust characterization of the time-series across different timescales. When applied to a music score, the beat spectrum encapsulates the musical performance rhythm in accordance to its perceptual features. The beat spectrum is derived directly from the amplitude modulation of a time series (Todd 1994), namely from self-similarity measures of the multi-resolution wavelet decomposition of the amplitude envelope of the music signal (Foote and Uchihashi 2001). The same analysis is applied to the time courses of regional brain activity to obtain a corresponding beat spectrum. The beat spectrum captures the rhythmic properties of the amplitude modulation: rhythmically transient responses that preserve the stimulus periodicities will produce similar beat spectra. The beat spectrum is sufficiently sensitive to capture small changes in rhythm on the sound track of different musical excerpts produced by the expressive performance of the artist.

We made use of this sensitivity by studying how the new measure changed for different brain areas with changes in DRIR in the three segments of the motif. For each participant, the correlation was computed between the beat spectra for each music segment and each of six brain areas—three motor-related areas in each hemisphere. Figure 8.4 shows the mean correlation for each segment and the statistical significance for the two changes (between segments A–B and B–C). The results demonstrate that when external rhythm is close to metrical (i.e., for small DRIR), the coherent mode of oscillatory activity is encountered in the left hemisphere. The increased synchronization between the internal and external rhythms might result in a subsequent mental stability of the rhythm percept and the efficient self-generation or retrieval of these rhythms, and easier memorization and reproduction of these rhythmical patterns (Essens and Povel 1985; Sakai *et al.* 1999). The change to high DRIR reduces the rhythm-tracking in the left but not the right hemisphere. Hence, the right hemisphere dominates when high DRIR rhythms are encountered, which is consistent with better tracking of non-metrical rhythms

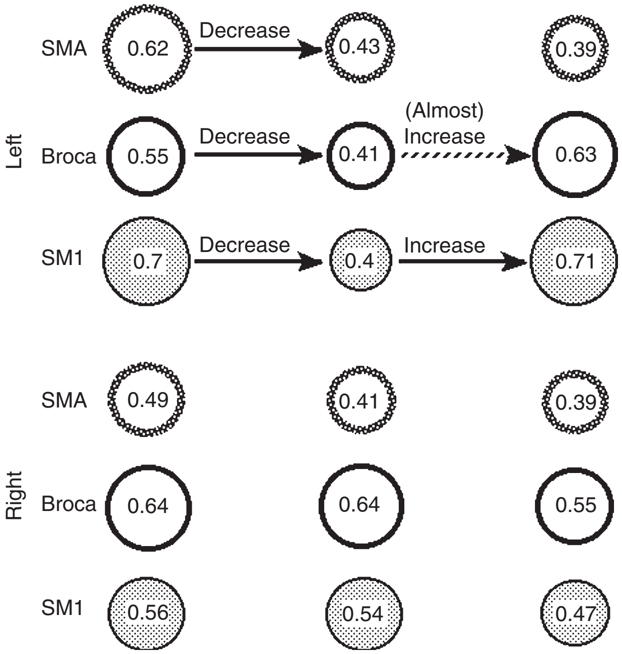

Fig. 8.4 Measures expressing the relationship between the temporal organization of musical sound and brain activity. Each measure corresponds to the correlation between two beat spectra. The first beat-spectrum is extracted from the time series for the regional activation of one brain area. Six brain areas were used: three on each side, namely left and right supplementary motor area (SMA), left and right Broca area and left and right primary sensorimotor area (SM1). The results for one area are shown in one row in the figure. The second beat-spectrum is extracted from the time series of a music segment. Three successive music segments (A–C) were used. The first (a) and third (c) music segments were more regular in their rhythmic properties relative to the middle segment (b). The values of the mean beat-spectrum correlation for each segment is shown in one column. A solid arrow between successive segments marks statistically significant change in the beat spectrum correlation, while a dash arrow corresponds to a trend that narrowly fails to reach significance.

(Roland *et al*. 1981). However, the overall conclusion from the beat spectrum analysis is that the processing of rhythm is not confined to only one hemisphere, in agreement with recent studies (Peretz 1990; Sakai *et al*. 1999).

The full dynamics of brain activity can only be appreciated in the real-time tomographic displays of brain activity as the music is played. Single-trial analysis of the MEG signal has been performed using magnetic field tomography (MFT) (Ioannides *et al*.1990; Ioannides 2001). Statistical parametric mapping (SPM) of these solutions were computed for 500 ms windows, comparing brain activity during music listening with the pre-stimulus period. The SPM analysis was made using the MFT solutions derived separately for each of 20 repetitions of the motif. The 500 ms window was run at the same rate as the music so the resulting audiovisual display shows changes in brain activity unfolding in real time (click on ‘MUSIC and animation’ at http://www.hbd.brain.riken.jp/auditorymusic.htm). The display shows the slow build-up of activity, beginning in motor-related areas, that eventually engulfs much of the brain and outlasts the music.

In summary, tomographic analysis of MEG responses to real music demonstrates that very large areas of the brain are activated when we listen to music. These activations differ in the left and right hemispheres; the left hemisphere is more engaged when regular rhythms are encountered. The activity in different brain areas reflects musical structure over different timescales; auditory and motor areas closely follow the low-level, high-frequency musical structure. In contrast, frontal areas contain a slower response, presumably playing a more integrative role. All of these results show that listening to music simultaneously engages distant brain areas in a cooperative way across time. This might be one reason why music has such a profound impact on humans.

8.9 **Learning musical techniques**

The difficulty of assigning innate status to a particular brain competence is underlined by a number of studies of brain changes that result from training. Clearly, our brains are made to take the imprint of experience, and especially human-created experience, or there would be much less point in having them. Thus, it can become very difficult to disentangle nature from nurture. It is understandable that in forming a neural imprint corresponding to a particular skill, our brains will engage in heightened activity in areas that are in some sense the most appropriate to the task. The outline of cortical functional segregation—especially the elaborate functional adaptations of subcortical systems, including emotional and communicative adaptations with which we are born—inevitably regulates which areas are brought into play in acquisition of a learned skill (Panksepp and Trevarthen, Chapter 7, this volume). Neurons that fire together, wire together, as Hebb (1949) pointed out, and thus areas that are exercised mature into more fully specialized regions that are fine-tuned for efficient and accurate performance.

Electrophysiological measures of changes can be traced at different temporal scales in the adult brain. Pantev and colleagues used MEG techniques in a number of studies to investigate the changes that occur in the human auditory cortex when a skill is acquired, such as when learning to play a musical instrument (Pantev *et al.* 2003). These studies showed that increased neuromagnetic response to musical stimuli was correlated with the age at which musicians began to practice, and that this response was preferentially enhanced for the timbre of the instrument on which the musician was trained (comparing violinists and trumpeters). In one of the studies, short-term laboratory training, involving learning to perceive virtual pitch (perceived frequency created through processing by the brain) instead of spectral pitch (frequency that is physically present), showed that the switch to perceiving virtual pitch was correlated with a significantly stronger gamma band frequency response, combined with a shift towards a more medial source location (Schulte *et al.* 2002).

Using fMRI and making very careful measurements of the amount of grey matter at any given part of the brain has enabled the tracking of brain changes resulting from practice, in particular musical practice. In the forerunner fMRI study of this work, volunteer participants were asked to learn the same sequence of finger-to-thumb tapping for 10 minutes each day, similar to a five-finger piano keyboard exercise (Karni *et al.* 1995). The participants were scanned each week while slowly performing the tapping sequence they were learning, and also when performing a very similar but not practised control sequence. When tested for speed and accuracy, over the first three weeks the performance level outside the scanner reached a plateau level. During this time, the brain areas in primary motor cortex (M1) that were activated during the practised sequence grew larger by comparison with those for the unpractised sequence. Several weeks after the end of the study, when tested again in the scanner, the participants showed that this increase in area, like the new skill associated with it, had remained unaltered, suggesting a permanent change in brain organization, with more neural tissue devoted to the trained task.

This study was followed by one that was more explicitly music-related—the investigation of brain changes associated with the learning of musical notation (Stewart *et al.* 2003); music notation-naive participants were trained for 12 weeks up to Grade One standard in sight-reading (the first grade of the Associated Board, UK). Before training, they were scanned while looking at a musical score and performing a simple keyboard task that did not require an understanding of the notation; after training, they were again scanned, this time performing an exactly analogous keyboard task, which now required a reading of the score appropriately. When the results were averaged across all 12 participants' brains, the only significant changes after training were a task-related increase in brain activity in the right dorsal parietal lobe, and a decrease in activity in the hippocampus. The parietal increase can be interpreted as increased specialization of this region of cortex to enable the automatic spatial decoding of the score from a vertical arrangement of notes to a horizontal placement of fingers. The hippocampal decrease can be seen as a familiarity-driven reduction of activity in this region, which is well-known to be important in the laying down of long-term memories.

It is now becoming apparent that part of the process of cortical specialization is a quantitative change in the amount of grey matter in task-specific areas. This has been demonstrated in a series of studies of increasing sophistication and convincingness, including Gaser and Schlaug (2003), who showed that professional musicians have increased grey matter density in several brain regions, dependent on the particular skills required. From an early age, musicians learn complex motor and auditory skills, which they practise extensively from childhood throughout their entire careers. Using a voxel-by-voxel morphometric technique and high-resolution structural MR images, Schlaug and his colleagues found that grey-matter volume increases in motor, auditory, and visual–spatial brain regions when comparing professional keyboard players with a matched group of amateur musicians and non-musicians. Schneider *et al.* (2002), made careful measurements using MR images of the Heschl's gyrus of 12 professional musicians, 12 amateur musicians, and 13 non-musicians, and used MEG to record the evoked response to pure tones in the same participants. The results showed that the evoked response, and the size of Heschl's gyrus, were both strongly correlated with the musical experience of the participants. In Sluming *et al.* (2002), increased grey matter density was identified in Broca's area of the left hemisphere of orchestral musicians, compared with non-musician controls. Since Broca's area and its right hemisphere homologue appear to be involved with musical syntax (Levitin and Menon 2003), and sight-reading unquestionably involves Broca's area (Parsons 2001), this is entirely consistent with the other studies in this area.

It has been argued that some of these regional differences could be attributable to innate predisposition, but a very recent study has shown highly significant morphometric changes in the visual motion areas V5 and parietal cortex in participants who learned to juggle over a period of three months (Draganski *et al.* 2004).

The studies we have reviewed support the notion that differences between brain areas in different participants related to musical skill, including volumetric differences, are largely acquired rather than innate. One may ask in this context, then, whether the question of human brain musicianship is not so much, 'Is the brain hard-wired for music production?' as 'Can the brain be trained to produce music?'. The answer, for musicianship, would appear to be 'Yes, the brain can be trained to produce music'. It is abundantly clear that what the hearer, at least in our culture, enjoys as music is well-tuned to what the musician can produce, which could then support the idea that part of musical enjoyment is bound up with the system of neurons that inherently equips us to move with the actions of others—that in hearing music we can vicariously participate in its production, which is often a fulfilling and satisfying experience (Cross and Morley, Chapter 5, this volume).

8.10 **Music and emotion**

This moves us on to a crucial question: what motivates us to experience music? Since musical structures provide no specific information, make no obvious short-term structural changes in our tissue, and are emphatically transient, why should we spend so much time listening to music, why should music accompany most forms of entertainment, and why should the music industry comprise such a large proportion of the economy? The answer to these questions can only be that music is a powerful regulator of our emotions. Music makes us feel, in a controllable and safe way. It allows us to experience feelings of great joy and sadness without the costs associated with the social and personal events that would otherwise be required to induce these emotions. It can calm and soothe us, stir us up, ravish us with its beauty.

It is well known that the human brain structures most involved with emotion are old, in evolutionary terms. While it is still quite uncertain (*pace* Orpheus!) whether any other species experience music in the way that we do, it seems likely that our emotional response to music is underlain by innate processes taking place in ancient brain structures, which receive complex modulation as we mature and are exposed to the characteristic musical forms of our own cultures (Panksepp and Trevarthen, Chapter 7, this volume).

Because musical forms often have clearly defined structures, one element of emotional response is surely related to anticipation and resolution (Meyer 1956). Many areas of our brains appear to be concerned with planning, whether in the feedforward circuitry entailed in motor control (Wolpert *et al.* 1995), in the phenomenon of 'priming', where perceptual expectations are set up in specific cortical areas, or in the prefrontal areas associated with selection for action (Rowe and Passingham 2001) and working memory (Cohen *et al.* 1997) (Lee and Schögler, Chapter 6, this volume).

Krumhansl (1990, 2002), following Lehrdahl and Jackendorff (1983), has carefully analysed the expectations set up by musical passages, showing that the degree of tension can be quantified. There may well be an autonomic response, with emotional force, when a high degree of tension built up in a musical passage is eventually resolved in a satisfying way. A further source of emotional response may occur when the music takes an unanticipated direction, which is only later seen as logically implicit in what has gone before. This can provoke a type of 'aha!' experience. Whether the perception of musical tension has innate foundations, or whether the expectations that are generated are entirely brought about by exposure to a culture's typical musical forms, is a question still without answer. Infants very quickly learn to recognize a 'favourite' song; this is facilitated by the way the song is encountered as part of a meaningful interaction between a baby and its adult companion. The structure of 'motherese' songs that are sung to babies are highly adapted to the babies' preferences (Trevarthen and Malloch 2002), and often take a dramatic form leading to a strong cadence or an ending with physical involvement (e.g., 'Round and Round the Garden'), suggesting that brain circuits which provide a sense of anticipation are active, even in early infancy.

Perhaps the most powerfully rewarding emotional experience that music can provide has been given the name of 'chills' (e.g., Panskepp 1995). This describes a response to specific musical passages which induce an intensely pleasurable, euphoric sensation that can bring tears to the eyes and shivers up the spine (and see Panksepp and Trevarthen, Chapter 7, this volume). Because such chills are clear, discrete events and are often highly reproducible for a specific piece of music in a given individual, they provide a good model for neuroimaging studies of emotional responses to music. A seminal study by Blood and Zatorre (2001) explored this phenomenon using PET. Ten musically trained students, equal numbers of males and females, were interviewed to discover which musical passages reliably elicited the chills response. One such passage was selected for each participant, making sure that it did not elicit the chills for any of the other participants, for whom it could act as a control stimulus. All musical passages were purely instrumental and came from the Western classical tradition.

The PET data revealed a network of brain areas that are closely involved in the chills experience. Some of these, including several that have previously been associated with reward, showed increased activity, and others with decreased activity, when compared with listening to the control musical passages that did not elicit this experience. Regional CBF (cerebral blood flow) increases were found in left ventral striatum, which includes the nucleus accumbens and dorsomedial midbrain, and decreases were found in right amygdala, left hippocampus/amygdala and ventromedial prefrontal cortex. Increases in chills intensity were also observed in paralimbic regions (bilateral insula, right orbitofrontal cortex) and in regions associated with arousal (thalamus and anterior cingulate) and motor processes (supplementary motor area and cerebellum). The pattern of activity correlating with music-induced chills is similar to that observed in other brain imaging studies of euphoria and/or pleasant emotion. Dopaminergic activity in the nucleus accumbens appears to be the common mechanism underlying the reward response to all naturally rewarding stimuli (e.g., food and sex) and to drugs with euphorogenic properties and/or abuse potential. Activity in the insula is associated with subjective feeling states involving representation of bodily responses elicited by emotional events (Critchley *et al*. 2004).

When such an extensive network of brain areas is seen to be involved with a particular experimental condition, interpretation becomes difficult, especially with the relatively poor spatial resolution of PET scanning. Thus, it is not yet possible to identify areas in this network that are specific for music-induced emotional response. However, it is encouraging to be able to observe a distinct brain-state correlate for the powerful feelings that are subjectively reported. Many of these areas are phylogenetically ancient. The human genre of music has somehow enlisted these areas, and it will be a pressing task of music and brain research to discover how this comes about.

8.11 **Importance and innateness of music: insights from imaging studies**

Brain imaging research related to music is rapidly growing in volume. Increasing numbers of neuroscientists have recognized that there is a dearth of empirical research in this area, and that many important questions can be raised relating to what seems to be a human need for music.

This brief survey has focused on a few of these questions, summarized here.

Language and music. It has become clear that while language and music share wide areas of neuronal substrate, there are differences, even at the level of primary auditory cortex. Generally these differences take the form of greater bilaterality for music than for language; langauge tends to be strongly left-lateralized, especially in right-handed males. This is consistent with a view that the capacity to be affected by music, when compared to language, is more likely to be innate, and is supported by specialized brain areas. Language, by this argument, is a highly specialized subset of musical cognition.

Features of music. Several experiments have shown that certain elements of music, such as pitch, have specific neural representations. There is good reason to believe that the neural apparatus that supports these features is predefined, even though it may become much more effective as a result of musical experience.

Acquisition of musical skill. The limited number of truly longitudinal studies using brain imaging techniques makes it hard to assert that there are human brain areas that are uniquely adapted for learning musical techniques. It is more likely that areas which have a general competence, for example for control of hand movements or vocal production and articulation, become specialized as a result of extensive practice, and there is increasing evidence that this specialization is accompanied by an increased bulk of neuronal tissue in some parts of the cerebral cortex.

Music and emotion. Music can induce powerful and primordial emotions in humans which are experienced as rewarding and pleasurable. Brain imaging techniques have allowed localization of activity associated with such emotions, and the areas involved are consistent with those involved with emotional responses to other types of stimuli. There remains the burning question: how and why can a structural sequence of non-referential sounds produce such a powerful response?

8.12 **Future work**

The imaging neuroscience of music is a research area that is experiencing rapid growth. The non-invasiveness of MEG and fMRI, and the continued existence of deep questions regarding the power and purpose of music, have encouraged increasing numbers of eager young researchers to enter this field. However, only a small fraction of published and planned studies are relevant to the study of the innateness of musicality. Most are concerned with music as an acquired cultural skill, which may or may not shed light on whether the brain structures involved are uniquely suited to the task.

When it comes to assessing electrophysiological methods, there is an urgent need to bring together two types of information. On the one hand, studies with large number of participants using average EEG and/or MEG data and simple models for the generators, show that the brain deals with language and music in remarkably similar ways. On the other hand, studies using very detailed single-trial tomographic analysis, but of only a few participants, show a much more dynamic picture of brain activity. These studies reveal distinct differences in how the left and right hemispheres share the processing of simple tone sequences, linguistic or music material, depending crucially on the fine details of the material (e.g., expectation and rhythmic properties). Between these extremes, the data from fMRI and PET show a degree of specialization for music and language material, but their low temporal resolution makes the establishment of contact between the haemodynamic and electrophysiologically based methods difficult without sacrificing the finer points of each. The synthesis of this wealth of neuroimaging data and the establishment of further associations with behaviour is one of the challenges that the field now faces.

Whatever technique is used, there is a particular need for longitudinal neuroimaging studies of the development of musical perceptions and skills, especially in ecologically valid contexts, such as mother–baby interactions. These could allow us to determine the way in which musical expression may precede language, and to relate music with rhythmical movement and expressive gestures that are adaptive, minimizing the influence of acculturation. MEG lends itself better than fMRI for such studies, because it is silent in operation and the scanning conditions can be adapted more easily to suit young babies and their mothers. However, much can be achieved using fMRI, provided that attention is given to the comfort of the participants. Because musical skills are often learned much later than language skills, there are rich opportunities for further longitudinal studies of music training, like that of Stewart *et al.* (2003), described above. Changes in brain organization relating to keyboard fluency and levels of rhythm complexity could be studied, both with fMRI and MEG.

Further work on the relationship of music and language perception and production is highly desirable. Existing studies do not often control the tasks or stimuli sufficiently well to pinpoint and fully characterize the differences in brain structures involved. By contrast, Callan *et al.* (2006) compare brain responses to hearing familiar songs with those to hearing spoken versions of the same songs by the same speakers. This experimental design involves approximate controls for semantic content, and for timbre and auditory source; results show remarkably small differences in brain activation.

Song itself needs further study. Neuroimaging studies of song and singing may approach the question of human innate musicality more directly than research involving such culturally relative auditory sources as musical instruments.

Still further studies could investigate the relationship between brain areas involved in the perception and production of rhythmical sounds, and those dealing with other controlled and repetitive movements. Research on the brain's mirror systems is still in its infancy, and further experiments explicitly addressing the integration of hearing and sight in a musical context would be of great interest, especially with regard for sympathetic brain responses to expressive forms of moving (Calvo-Merino *et al*. 2005). For example, an fMRI or MEG study where participants view dancers moving to music, in which in some conditions the music is incongruous with the movements, could identify brain areas responsible for this integration. Finding activations in areas associated with the autonomic system would provide evidence of the innateness of the association of auditory and motor rhythms.

Other research might include studies of musically induced emotion, comparing brain activations produced by passages selected for their emotional effects with emotionally laden natural sounds, such as the cry of a baby, the coo of a dove, the crash of breaking glass, and a cry of joy. This could help to discriminate the pathways by which musical perception enters our innate emotional networks, which have clearly evolved in close relationship with our requirements for survival.

References

Abo M, Senoo A, Watanabe S, Miyano S, Doseki K, Sasaki N, Kobayashi K, Kikuchi Y, Besson M and Macar F (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, **24(1)**, 14–25.

Abo M, Senoo A, Watanabe S, Miyano S *et al.* **(2004).** Language-related brain function during word repetition in post-stroke aphasics. *Neuroreport*, **15(2)**, 1891–1894.

Adolphs R (2003). Cognitive neuroscience of human social behaviour. *Nature reviews, Neuroscience*, **4(3)**, 165–178.

Besson M and Macar F (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, **24**, 14–25.

Besson M and Schon D (2001). Comparison between language and music. *New York Academy of Scences*, **930,** 232–258.

Besson M, Faita F and Requin J (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience Letters*, **168,** 101–105.

Binder JR, Frost JA, Hammeke TA, Rao SM and Cox RW (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, **119,** 1239–1247.

Blank SC, Bird H, Turkheimer F and Wise RJ (2003). Speech production after stroke: the role of the right pars opercularis. *Annals of Neuroogy*, **54(3)**, 310–320.

Blood AJ and Zatorre RJ (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences USA*, **98(20)**, 11818–11823.

Bosnyak DJ, Eaton RA and Roberts LE (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, **14(10)**, 1088–1099.

Brugge JF (1985). Patterns of organization in auditory cortex. *Journal of the Acoustical Society of America*, **78(1/2)**, 353–359.

Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H and Turner R (2006). Song and speech: Brain regions involved with perception and covert production. *Neuroimage*, **31,** 1327–1342.

Calvert GA, Brammer MJ, Morris RG, Williams SC, King N and Matthews PM (2000). Using fMRI to study recovery from acquired dysphasia. *Brain and Language*, **71(3)**, 391–399.

Calvo-Merino B, Glaser DE, Grezes J, Passingham RE and Haggard P (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex*, **15(8)**, 1243–1249.

Carey S and Bartlett E (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, **15,** 17–29.

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J and Smith EE (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, **386,** 604–608.

Critchley HD, Wiens S, Rotshtein P, Ohman A Dolan RJ (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, **7(2)**, 189–195.

Dawson G and Fischer KW (eds) (1994). *Human behavior and the developing brain*. The Guilford Press, New York.

Decety J and Chaminade T (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, **41,** 127–138.

Devlin JT, Raley J, Tunbridge E, Lanary K, Floyer-Lea A, Narain C, Cohen I, Behrens T, Jezzard P, Matthews PM and Moore DR (2003). Functional asymmetry for auditory processing in human primary auditory cortex. *Journal of Neuroscience*, **23(37)**, 11516–11522.

Dhamala M, Pagnoni G, Wiesenfeld K, Zink CF, Martin M and Berns GS (2003). Neural correlates of the complexity of rhythmic finger tapping. *Neuroimage*, **20(2)**, 918–926.

Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U and May A (2004). Neuroplasticity: changes in grey matter induced by training. *Nature*, **427(6972)**, 311–312.

Essens PJ and Povel DJ (1985) Metrical and nonmetrical representations of temporal patterns. *Perceptual Psychophysics*, **37,** 1–7.

Foote J and Uchihashi S (2001). The beat spectrum: A new approach to rhythm analysis. *Proceedings of IEEE International Conference on Multimedia and Expo*. Paper available from http://rotorbrain.com/foote/papers/allpapers.html

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K and Goebel R (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron*, **40(4)**, 859–869.

Freeman WJ (2005). Origin, structure and role of background EEG activity. Part 3. Neural frame classification. *Clinical Neurophysiology*, **116,** 1118–11129.

Freeman WJ and Holmes MD (2005). Metastability, instability, and state transition in neocortex. *Neural Networks*, **18,** 497–504.

Friederici AD (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, **50(3)**, 259–281.

Friederici AD, Wang Y, Herrmann CS, Maess B and Oertel U (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, **11(1)**, 1–11.

Gallese V (2001). The 'Shared Manifold' hypothesis: From mirror neurons to empathy. In E Thompson, ed., *Between ourselves: Second-person issues in the study of consciousness*, pp. 33–50. Imprint Academic, Charlottesville, VA/Thorverton, UK.

Gaser C and Schlaug G (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, **23(27)**, 9240–9245.

Griffiths TD, Buchel C, Frackowiak RS and Patterson RD (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, **1(5)**, 422–427.

Gross J, Ioannides AA, Dammers J, Maess B, Friederici AD and Muller-Gartner HW (1998). Magnetic field tomography analysis of continuous speech. *Brain Topography*, **10(4)**, 273–281.

Halberda J (2003). The development of a word-learning strategy. *Cognition*, **87,** B23–B34.

Halpern AR and Zatorre RJ (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebal Cortex*, **9(7)**, 697–704.

Hebb DO (1949). *The organization of behavior*. Wiley, New York.

Iacoboni M, Koski LM, Brass M, Bekkering H, Woods RP, Dubeau MC, Mazziotta JC and Rizzolatti G (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences USA*, **98(24)**, 13995–13999.

Ioannides AA (1995). Estimates of 3D brain activity ms by ms from biomagnetic signals: Method (MFT), results and their significance. In E Eiselt, U Zwiener and H Witte, eds, *Quantitative and topological EEG and MEG analysis*, pp. 59–68. Universitaetsverlag Druckhaus-Maayer GmbH, Jena.

Ioannides AA (2001). Real time human brain function: Observations and inferences from single trial analysis of magnetoencephalographic signals. *Clinical EEG*, **32,** 98–111.

Ioannides AA, Bolton JPR aqnd Clarke CJS (1990). Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problem*, **6,** 523–542.

Ioannides AA, Fenwick PBC and Liu LC (2005). Widely distributed magnetoencephalography spikes related to the planning and execution of human saccades. *Journal of Neuroscience*, **25,** 7950–767.

Ioannides AA, Kostopoulos GK, Laskaris NA, Liu LC, Shibata T, Schellens M, Poghosyan V and Khurshudyan A (2002a). Timing and connectivity in the human somatosensory cortex from single trial mass electrical activity. *Human Brain Mapping*, **15,** 231–246.

Ioannides AA, Poghosyan V, Dammers J and Streit M (2004). Real-time neural activity and connectivity in healthy individuals and schizophrenia patients. *NeuroImage*, **23,** 473–482.

Ioannides AA, Popescu M, Otsuka A, Bezerianos A and Liu LC (2003). Magnetoencephalographic evidence of the inter-hemispheric asymmetry in echoic memory lifetime and its dependence on handedness and gender. *NeuroImage*, **19(3)**, 1061–1075.

Ioannides AA, Popescu M, Otsuka, Abrahamyan A and Deliège I (2002b). Using neuroimaging to study neural correlates of music over wide spatial and temporal scales. In C Stevens, D Burnham, G McPherson, E Schubert and J Renwick, eds, *Proceedings of the 7th International Conference on Music Perception and Cognition – ICMPC7, Sydney, July*, pp. 677–680. Australian Music and Psychology Society (AMPS), Sydney NSW and Causal Productions, Adelaide, SA. Published as CD Rom only.

Janata P (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, **13,** 1–17.

Jeannerod M (2004). Visual and action cues contribute to the self-other distinction. *Nature Neuroscience*, **7(5)**, 421–422.

Karni A, Meyer G, Jezzard P, Adams MM, Turner R and Ungerleider LG (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, **377(6545)**, 155–158.

Koelsch S, Gunter TC, Schroger E, Tervaniemi M, Sammler D and Friederici AD (2001). Differentiating ERAN and MMN: An ERP study. *Neuroreport*, **12(7)**, 1385–1389.

Koelsch S, Gunter TC, v Cramon DY, Zysset S, Lohmann G and Friederici AD (2002). Bach speaks: a cortical 'language-network' serves the processing of music. *Neuroimage*, **17(2)**, 956–966.

Koelsch S, Kasper E, Sammler D, Schulze K, Gunter T and Friederici AD (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, **7(3)**, 302–307.

Krumhansl CL (1990). *Cognitive foundations of musical pitch*. Oxford University Press, New York.

Krumhansl CL (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, **11(2)**, 45–50.

Kuhl PK (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, **5(11)**, 831–843.

Kutas M and Hillyard SA (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, **207(4427)**, 203–205.

Kutas M and Hillyard SA (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, **307(5947)**, 161–163.

Laskaris N, Liu LC and Ioannides AA (2003). Single-trial variability in early visual neuromagnetic responses: an explorative study based on the regional activation contributing to the N70m peak. *NeuroImage*, **20(2)**, 765–783.

Lehrdahl F and Jackendorff R (1983). *A generative theory of tonal music*. MIT Press, Cambridge, MA.

Levitin DJ and Menon V (2003). Musical structure is processed in 'language' areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *Neuroimage*, **20(4)**, 2142–2152.

Liégeois-Chauvel C, de Graaf JB, Laguitton V and Chauvel P (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding, *Cerebral Cortex*, **9,** 484–496.

Liegeois-Chauvel C, Giraud K, Badier JM, Marquis P and Chauvel P (2001). Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Annals of the New York Academy of Sciences*, **930,** 117–132.

Liu LC, Ioannides AA and Mueller-Gaertner HW (1998). Bi-hemispheric study of single trial MEG signals of the human auditory cortex. *Electroenceph Clin Neurophysiol*, **106,** 64–78.

Llinas RL (2001). *I of the vortex, from neuroscience to self*. MIT Press, Cambridge, MA.

Maess B, Koelsch S, Gunter TC and Friederici AD (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, **4(5)**, 540–545.

McAdams S, Winsberg S, Donnadieu S, De Soete G and Krimphoff J (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, **58,** 177–192.

Menon V, Levitin DJ, Smith BK, Lembke A, Krasnow BD, Glazer D, Glover GH and McAdams S (2002). Neural correlates of timbre change in harmonic sounds. *Neuroimage*, **17(4)**, 1742–1754.

Meyer LB (1956). *Emotion and meaning in music*. University of Chicago Press, Chicago, IL.

Molinari M, Leggio MG, De Martin M, Cerasa A and Thaut M (2003). Neurobiology of rhythmic motor entrainment. *Annals of the New York Academy of Sciences*, **999,** 313–321.

Moradi F, Liu LC, Cheng K, Waggoner RA, Tanaka K and Ioannides AA (2003). Consistent and precise localization of brain activity in human primary visual cortex by MEG and fMRI. *NeuroImage*, **18,** 595–609.

Morel A, Garraghty PE and Kaas JH (1993). Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, **335(3)**, 437–459.

Näätänen R (1992). *Attention and brain function*. Erlbaum, Hillsdale, NJ.

Osterhout L and Holcomb PJ (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, **31,** 785–804.

Overy K, Norton AC, Cronin KT, Gaab N, Alsop DC, Winner E and Schlaug G (2004). Imaging melody and rhythm processing in young children. *Neuroreport*, **15(11)**, 1723–1726.

Panksepp J (1995). The emotional sources of 'chills' induced by music. *Music Perception***, 13(2),** 171–207.

Pantev C, Hoke M, Lehnertz K, Lutkenhoner B, Anogianakis G and Wittkowski W (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory-evoked magnetic fields. *Electroencephalogr Clin Neurophysiol*, **69(2)**, 160–170.

Pantev C, Ross B, Fujioka T, Trainor LJ, Schulte M and Schulz M (2003). Music and learning–induced cortical plasticity. *Annals of the New York Academy of Sciences*, **999,** 438–450.

Parsons LM (2001). Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Annals of the New York Academy of Sciences*, **930,** 211–231.

Patel AD (2003). Language, music, syntax and the brain. *Nature Neuroscience*, **6(7)**, 674–681.

Patel AD and Balaban E (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature*, **404(6773)**, 80–84.

Patel AD, Gibson E, Ratner J, Besson M and Holcomb PJ (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, **10(6)**, 717–733.

Patterson RD, Uppenkamp S, Johnsrude IS and Griffiths TD (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, **36(4)**, 767–776.

Peretz I (1990). Processing of local and global musical information in unilateral brain damaged patients. *Brain*, **13,** 1185–1205.

Peretz I (2002). Brain specialization for music. *Neuroscientist*, **8(4)**, 372–380.

Pinker S (2000). *The language instinct: How the mind creates language*. HarperCollins Publishers, New York.

Popescu M, Otsuka A and Ioannides AA (2004). Dynamics of brain activity in motor and frontal cortical areas during music listening: A magnetoencephalographic study. *NeuroImage*, **21,** 1622–1638.

Recanzone GH, Schreiner CE and Merzenich MM (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience*, **13(1)**, 87–103.

Rivkin MJ, Vajapeyam S, Hutton C, Weiler ML, Hall EK, Wolraich DA, Yoo SS, Mulkern RV, Forbes PW, Wolff PH and Waber DP (2003). A functional magnetic resonance imaging study of paced finger tapping in children. *Pediatric Neurology*, **28(2)**, 89–95.

Rizzolatti G, Fadiga L, Gallese V and Fogassi L (1996). Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res*, **3(2)**, 131–141.

Roland PE, Skinhøj E and Lassen NA (1981). Focal activation of human cerebral cortex during auditory discrimination. *Journal of Neurophysiology*, **45,** 1139–11351.

Rowe JB and Passingham RE (2001). Working memory for location and time: Activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage*, **14(1/1)**, 77–86.

Sakai K, Hikosaka O, Miyauchi S, Takino R, Tamada T, Iwata NK and Nielsen M (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience*, **19(22)**, 10074–10081.

Samson S, Ehrle N and Baulac M (2001). Cerebral substrates for musical temporal processes. In Zatorre RJ and Peretz I, eds, *The biological foundations of music*, pp. 166–178. New York Academy of Science, New York.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A and Rupp A (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, **5(7)**, 688–694.

Schulte M, Knief A, Seither-Preisler A and Pantev C (2002). Different modes of pitch perception and learning-induced neuronal plasticity of the human auditory cortex. *Neural Plasticity*, **9(3)**, 161–175.

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

Stewart L, Henson R, Kampe K, Walsh V, Turner R and Frith U (2003). Brain changes after learning to read and play music. *NeuroImage*, **20(1)**, 71–83.

Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR and Dale AM (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *J Neurophysiology*, **91(3)**, 1282–1296.

Taylor JG, Ioannides AA and Muller-Gartner HW (1999). Mathematical analysis of lead field expansions. *IEEE Transactions on Medical Imaging*, **18,** 151–163.

Tervaniemi M, Kujala A, Alho K, Virtanen J, Ilmoniemi RJ and Naatanen R (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: A magnetoencephalographic (MEG) study. *Neuroimage*, **9(3)**, 330–336.

Todd NPM (1994). The auditory primal sketch: A multiscale model of rhythmic grouping. *Journal of New Music Research*, **23(1)**, 25–70.

Trainor LJ (1996). Infant preferences for infant-directed versus non-infant-directed play songs and lullabies. *Infant Behavior and Development*, **19,** 83–92.

Trevarthen C (1999). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. *Musicae Scientiae (Special Issue 1999–2000)*, 155–215.

Trevarthen C (2001). The neurobiology of early communication: Intersubjective regulations in human brain development. In AF Kalverboer and A Gramsbergen, eds, *Handbook on brain and behavior in human development*, pp. 841–882. Kluwer, Dordrecht, The Netherlands.

Trevarthen C (2002). Origins of musical identity: Evidence from infancy for musical social awareness. In R MacDonald, DJ Hargreaves and D Miell, eds, *Musical identities,* pp. 21–38. Oxford University Press, Oxford.

Trevarthen C (2004a). Brain development. In RL Gregory, ed., *Oxford companion to the mind*, 2nd edn, pp. 116–127. Oxford University Press, Oxford/New York

Trevarthen C (2004b). Language development: Mechanisms in the brain. In G Adelman and BH Smith, eds, *Encyclopedia of neuroscience*, 3rd edn, CD-ROM, Article Number 397. Elsevier Science, Amsterdam.

Trevarthern C and Malloch S (2002). Musicality and music before three: Human vitality and invention shared with pride. *Zero to Three*, 10–18.

Turner R and Jones T (2003). Techniques for imaging neuroscience. *British Medical Bulletin*, **65,** 3–20.

Tzourio-Mazoyer N, De Schonen S, Crivello F and Reutter B (2002) Neural correlates of woman face processing by 2-month-old infants. *Neuroimage*, **15,** 454–461.

Ullen F, Forssberg H and Ehrsson HH (2003). Neural networks for the coordination of the hands in time. *Journal of Neurophysiology*, **89(2)**, 1126–11235.

Von Helmoltz H (1853). Ueber einige Gesetze der Vertheilung elektrischer Stroeme in koerperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. *Ann Phys Chem*, **89,** 211–233, 353–377.

Wallin N (1991). *Biomusicology: Neurophysiological, neuropsychological, and evolutionary perspectives on the origins and purposes of music*. Pendragon Press, New York.

Warren JD, Uppenkamp S, Patterson RD and Griffiths TD (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences USA*, **100(17)**, 10038–10042.

Wolpert DM, Ghahramani Z and Jordan MI (1995). An internal model for sensorimotor integration. *Science*, **269(5232)**, 1880–1882.

Woods RP, Dodrill CB and Ojemann GA (1988). Brain injury, handedness, and speech lateralization in a series of amobarbital studies. *Annals of Neurology*, **23(5)**, 510–518.

Wright AA, Rivera JJ, Hulse SH, Shyan M and Neiworth JJ (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology General*, **129(3)**, 291–307.

Yonemoto K (2004). Language-related brain function during word repetition in post-stroke aphasics. *Neuroreport*, **15(12)**, 1891–1894.

Zatorre RJ, Evans AC and Meyer E (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, **14(4)**, 1908–1919.

Zeki S (1993). *A vision of the brain*. Blackwell Scientific Publications, Oxford.