A Neurotechnology Approach to the Analysis of Electroacoustic Music: A Proposition

Eduardo R. Miranda
Interdisciplinary Centre for Computer Music Research (ICCMR)
University of Plymouth, United Kingdom
http://cmr.soc.plymouth.ac.uk/
eduardo.miranda@plymouth.ac.uk

Abstract:
In this paper I propose an approach to the development of technology for the analysis of electroacoustic music, which is referred to as the *neurotechnology approach*. I propose the unprecedented possibility of developing tools for the analysis of electroacoustic music based on neurophysiologic models of our auditory system. These tools would reveal the representations that our brain produces at various stages of the auditory pathway, from the cochlea to the cortex. Firstly I describe the journey that sounds take through the auditory pathways from the ears up to the auditory cortices and then I enumerate the sorts of tools that could be developed based on the various stages of this journey and the types of information that they would elicit.

1 Introduction

There have been fantastic theoretical developments on establishing frameworks for the analysis of electroacoustic music. A glance at the proceedings of the Electroacoustic Music Studies conference series reveals a number of interesting approaches (http://www.ems-network.org). In comparison, despite tremendous advances in computing technology, the tools available for electroacoustic music analysis still seem to be in the stone-age: it lags far behind the theoretical work of scholars such as Smalley (1997), to cite but one example. That is, there are no suitable computing tools for signal analysis of pieces of electroacoustic music to support such theoretical framewoks. The great majority of tools available for the analysis of electroacoustic music are based on the Fourier transform paradigm (i.e., spectrogram), which in my view offers a rather limited representation of the signal. For instance, the Acousmographe, developed by Le Groupe de Recherches Musicales (GRM) in Paris is probably one of the best tools available for visualisation and annotation of electroacoustic music. Yet, the best it can offer in terms of signal analysis is the spectrogram.

In this paper I propose an approach to improve this scenario, which I refer to as the *neurotechnology approach*. In short, I envisage the unprecedented possibility of building

tools for the analysis of electroacoustic music based on neurophysiologic models of our auditory system. These tools would reveal the "representation"[1] that our brain produces at various stages of the auditory pathway, from the cochlea to the cortex. To the best of my knowledge, apart from the cochleogram[2] there have been no significant developments towards tools for the analysis of electroacoustic music based on neurophysiologic models. However, research into building computational models of auditory brain functioning is an area that is evolving rapidly. I would therefore expect to see progress at this front shortly.

The remaining of this paper is organised as follows: the following section describes the journey that sounds take through the auditory pathways from the ears up to the auditory cortices. I would refer to this journey as the journey from *hearing* to *active listening*. Then, I will enumerate the sorts of tools that could be developed based on the various stages of the auditory pathways and the types of information that they would elicit.

2 Auditory Pathways

The auditory pathways may be divided up into two types, or stages: *early auditory pathways* (Figure 1a), which lead from the outer ear[3] through a number of subcortical regions and terminate in the primary auditory cortex, and *cortical auditory pathways* (Figure 1b) that pass out activation from the primary auditory cortex to a number of other cortical areas. Whereas the early pathways are essentially the same for all types of sounds (i.e., all types of sounds go through the same path), the cortical pathways are dependent on the specific content of the auditory signal and brain plasticity. Also, it must be noted that early auditory pathways deal mostly with the analysis of single relatively short events, whilst cortical auditory pathways are primarily concerned with auditory Gestalten, sequencing, grouping, the building of representations and a number of other activities associated with music cognition. Cortical pathways enjoy greater plasticity than early pathways and their functioning are subject to learning and acculturation[4]. In this paper I will focus on the early auditory pathways.

---

[1] The term "representation" may not be the best to label this. Neuroscientists use the term "neural code".
[2] The cochleogram provides patterns of excitation at the basilar membrane of the cochlea in the inner ear as a function of time and frequencies in Barks. An excellent cochleogram tool is provided with the Praat software (http://www.fon.hum.uva.nl/praat/).
[3] This is where the arrow labelled "sound" is pointing to, at the bottom left-side of Figure 1a.
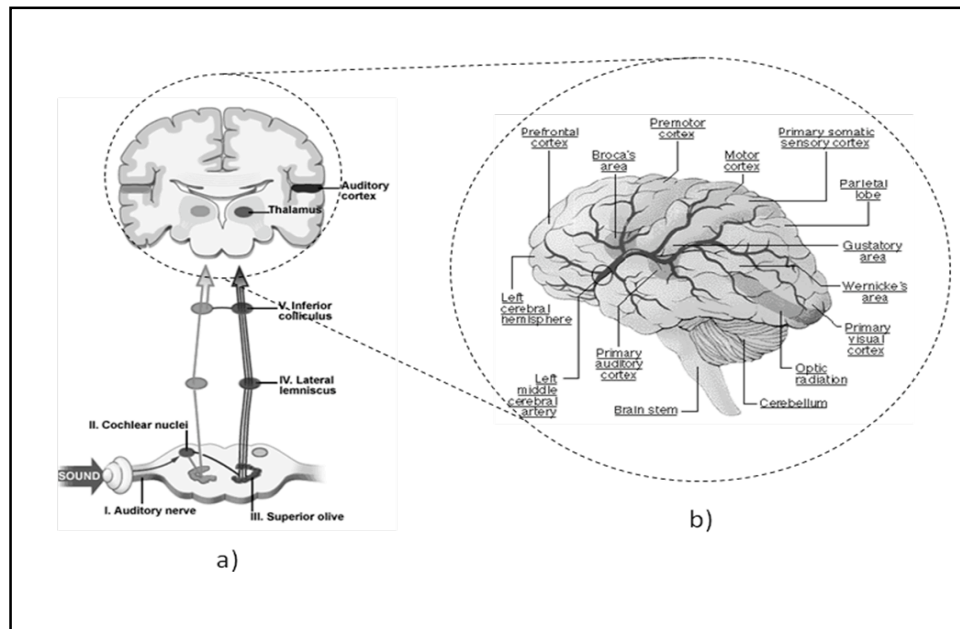[4] Acculturation is the process by which somebody absorbs the culture of a society from birth onward.

Figure 1:
The auditory pathways: (a) early auditory pathways; (b) cortical auditory pathways.

2.1 Early Auditory Pathways

Before I begin, I would like to clarify that while the description given below will be of the ascending early auditory pathway, there are also descending pathways that trace the route back from the primary auditory cortex all the way back to the cochlea. The specific functions of these descending pathways have not yet been fully established, although it is speculated that they are involved in top-down processes modulating the early processing of incoming sounds on the basis of previous sounds (Schofield and Cant, 1999).

Auditory signal processing in humans starts with the outer ear, which channels sounds towards the tympanic membrane, or eardrum (Figure 2). The tympanic membrane vibrates in response to air pressure changes and this vibration is relayed through the middle ear via three small bones: the malleus, the incus and the stapes. The last of these is connected to the oval window, which in turn leads into the cochlea, which is filled with fluid. The spiral-shaped cochlea forms the inner ear, and is responsible for transforming liquid motion into electrical neural signals. The movement of the liquid results in the movement of the basilar membrane and the subsequent stimulation of motion-sensitive hair cells. The basilar membrane gets wider towards the centre of the spiral, and also

decreases in stiffness. These two factors give it the crucial property of tonotopic[5] frequency selectivity: hair cells at given locations will respond maximally to particular frequencies[6] and do so in an ordered fashion, with the high frequency sounds stimulating cells near the outer edge of the spiral, and the low frequency sounds stimulating cells near the centre.
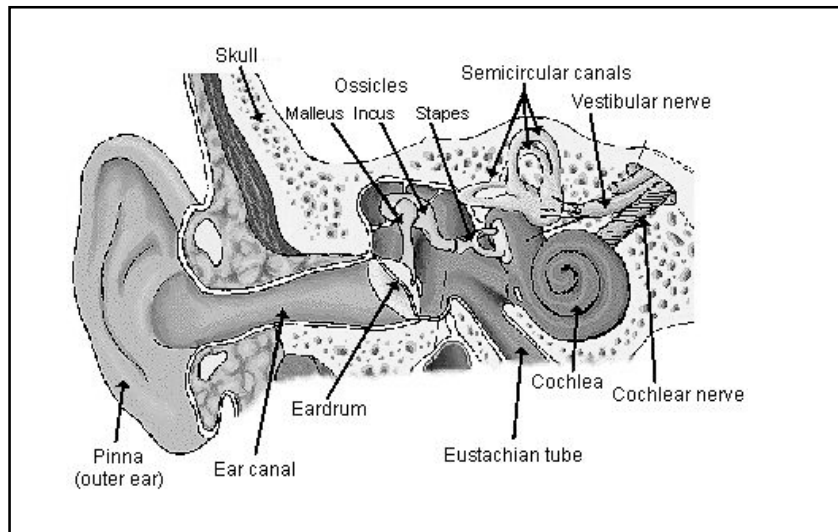

Figure 2:
The anatomy of the ear.

Spiral ganglion cells synapse onto the hair cells, and due to a pattern of innervating nearby cells, also show tonotopic frequency selectivity. They fire action potentials when stimulated, and their firing rate is approximately linearly related to the *sound intensity*, hence this information is also encoded. Altogether, the ear takes differences in air pressure, transforms them first into liquid motion, and then into tonotopically-organised neural signals that encode both *frequency* and *intensity* information.

The axons of the spiral ganglion cells of the cochlea form the auditory nerve, and from here the signal moves from the ear to subcortical brain regions. The auditory nerve terminates in the cochlear nuclei, which can be divided into one dorsal and two ventral regions. The ventral cochlear nuclei contain stellate cells, which are frequency-specific and encode *intensity* information within their firing rate, and bushy cells, which fire once at a stimulus onset, thus providing sound *onset timing* information. As there are also substantial connections between the cochlear nuclei of the two hemispheres (Shore et al., 1992), bushy cells also provide enough information to start to encode *horizontal position* information (since this is dependent on timing between the two ears), something which happens explicitly higher in the brainstem. The dorsal cochlear nucleus contains fusiform cells (otherwise found mostly in the cerebellum) which appear to be involved in

---

[5] Tonotopic refers to cells having the quality of being spatially organised by tone or frequency. Tones close to each other in terms of frequency activate topologically neighbouring cells.

[6] And to a lesser extent to adjacent frequencies, giving them a characteristic tuning curve.

vertical position encoding, and tuberculoventral cells, which are involved in identifying and suppressing the response of ventral bushy cells to echo sounds, thus allowing a very rapid discrimination between a *source sound and its echo*, which in evolutionary terms is likely to have been of survival benefit (Figure 3).
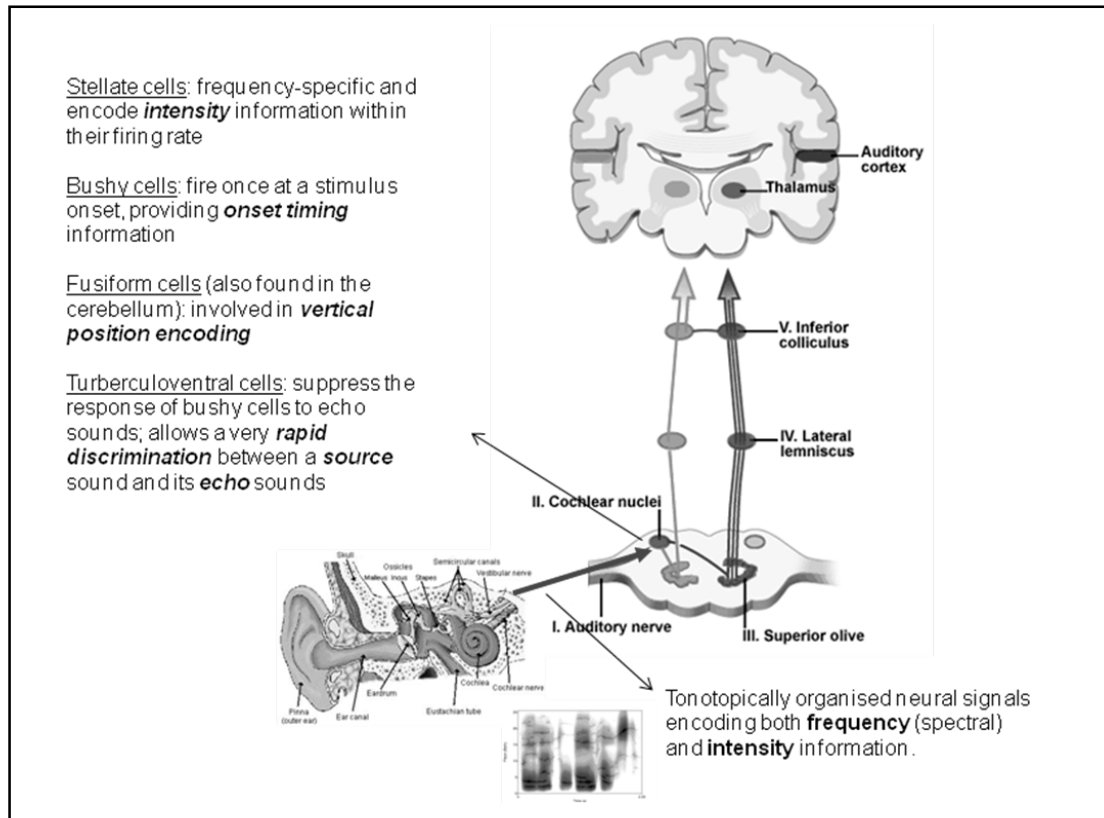

Figure 3:
The auditory nerve relays tonotopically organised neural signals to the cochlear nuclei. Different types of cells extract different types of information from these signals.

There are several pathways from this point onwards. The dorsal acoustic stria are a collection of axons that represent one of the main pathways, leading from the dorsal cochlear nucleus to both the lateral lemniscus (in the pons) and the inferior colliculus (in the midbrain). This is the most direct pathway up to this point. Similarly, the intermediate acoustic stria is a collection of axons that project mostly to the lateral lemniscus, where they synapse onto another set of neurons that project to the inferior colliculus. Hence we see that some of the axons from the cochlear nuclei terminate in the lateral lemniscus, some use it as a relay station, passing their signal to other neurons that move it on to the inferior colliculus, and some (the majority) simply pass right through the lateral lemniscus without synapsing at all, projecting directly to the inferior colliculus. Like the cochlear nuclei, the lateral lemnisci in both hemispheres are also connected, allowing some form of *binaural processing* at this stage.
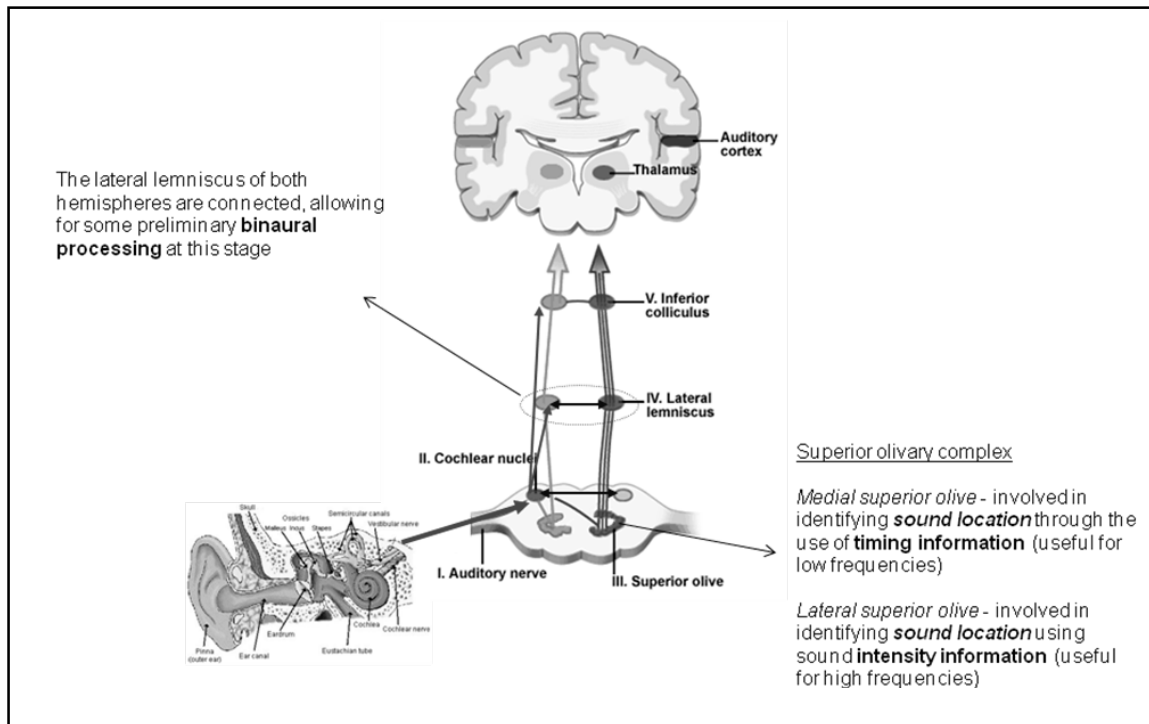
Figure 4:
Some connections terminate at the lateral lemniscus, but most use this area as a relay station, passing their signals on to other neurons that project to the inferior colliculus. The superior olivary complex provides two types of sound localisation mechanisms.

In addition to the dorsal pathways, the main ventral pathway projects first to the superior olivary complex in both hemispheres via the trapezoid body (Bear et al., 2001). The medial superior olive appears to be involved in identifying *sound location* through the use of timing information (which is most useful for low frequency sounds), while the lateral superior olive appears to identify location by using *sound intensity* information instead (which is most useful for high frequency sounds). Thus two different types of sound localisation mechanisms are available at this early point in the auditory pathway. These combined with frequency information allows for rapid processing of spatial/localisation information, which in evolutionary terms is crucial to survival (Figure 4).

The inferior colliculus (Figure 5) is where the various early auditory pathways once again converge and it appears to be engaged in further processing of *sound localisation*. The inferior colliculus is arranged in layers, and orthogonal to these layers the organisation reflects the tonotopic map that originated in the basilar membrane. Three additional important properties to note here are:

a) It receives inputs from the superior olivary complex in both hemispheres and the inferior colliculus in the opposite hemisphere
b) It receives somatosensory inputs

c) It has much greater processing power than lower structures, having some 400,000 neurons, as opposed to just 34,000 in the superior olivary complex (Worden, 1971)

Perhaps as a result of these properties, the inferior colliculus represents a sort of processing bottleneck in the auditory system. Lower inputs converge upon on it, and higher inputs, to the superior colliculus (which performs early *integration with the visual system*), to the reticular formation (involved in *autonomic processing*), the cerebellum (important centre of *motor activity*) and the thalamus, are projected to from it.
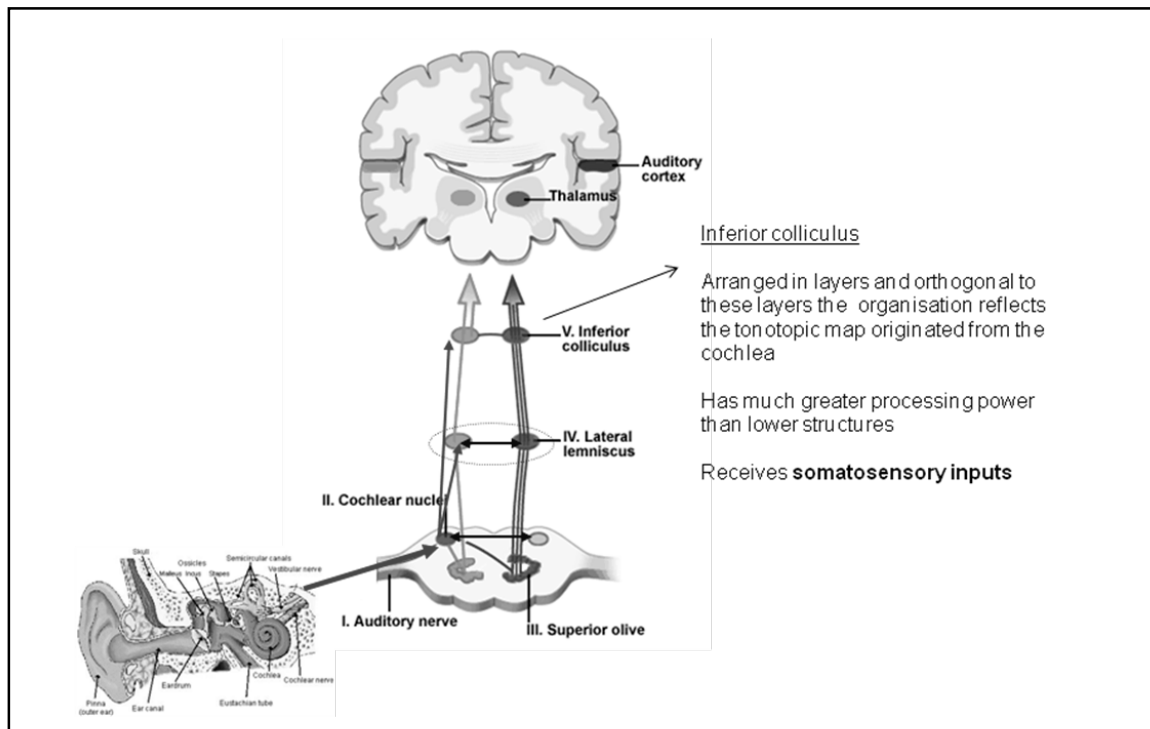

Figure 5:
Inferior colliculus performs early integration with the visual system (superior colliculus), the autonomic system (reticular formation) and motor activity (cerebellum).

The thalamus, or more precisely, a part of the thalamus referred to as the medial geniculate nucleus (MGN), represents the next stage in the main auditory pathway. It is considered to be the gateway to the cerebral cortex. The thalamus is largely responsible for control of *attention*; for example, it enables us to focus on a particular instrument from all other sounds in an orchestral piece. It seems that its role is to direct information to the cortex or suppressing it. The MGN receives inputs from the inferior colliculus and a number of other areas as well, including visual and somatosensory areas. It is here that sensorial information from *different modalities are combined* and where visual information, for example, can first modulate the main auditory signal. Neurons in the

MGN contain broader tuning curves than neurons earlier in the auditory pathway, but still exhibit some *frequency specificity* and tonotopic organisation. Some neurons also specifically encode *sound intensity* level. It is most probably at this point that hearing starts to become listening (Figure 6)[7].
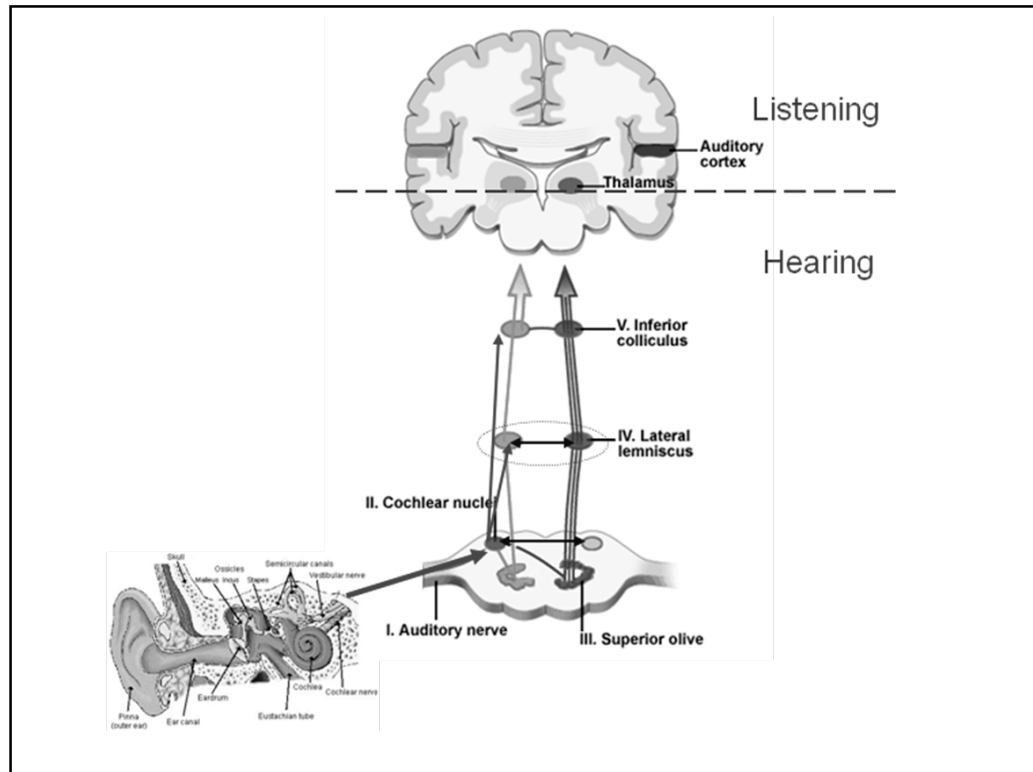

Figure 6:
At the thalamus, information from different modalities is combined in significant ways.
For example, it is where visual information can first modulate the main auditory signal.
The act of listening seems to begin to emerge at this point.

The MGN projects to the primary auditory cortex (referred to as A1), with some projections also going to the anterior and posterior auditory fields, and a smaller number to the secondary auditory cortex. The A1 is the interface between the early auditory pathways and the beginning of various cortical auditory pathways[8]. Located on Heschl's gyrus in the temporal lobe, this is a crucial neural area for auditory processing, building representations combining the extracted information in a process referred to as *auditory Gestalten* (Koelsch and Siebel, 2005). As in preceding areas of the early auditory pathway, A1 is tonotopically organised (Figure 7).

---

[7] I would venture to suggest that the notion of hearing becoming listening at this point of the auditory pathways may provide a neurophysiologic support to Pierre Schaeffer's (1966) discussion on different modes hearing and listening: *écouter, ouïr, entendre* and *comprendre*.

[8] I acknowledge that this notion of A1 as an interface between the two types of pathways is debatable. The vast majority of the literature considers the A1 as being part of the early auditory pathways.
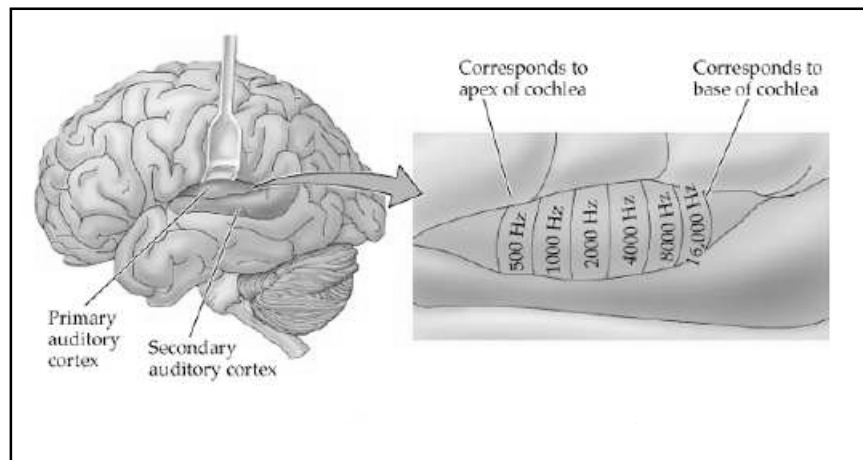
Figure 7:
Tonotopic organisation of auditory cortex.

Some studies of auditory imaging (Rauschecker, 2005), in which subjects are asked to imagine sound, reported that A1 was not active, in contrast to a situation of actually hearing the same sound, when A1 is active. Other studies, however, suggest that this may not always be the case (Halpern 2003; Zatorre 1999), and in fact it has been found that A1 appears to be active even during the anticipation of a sound (Voisin et al., 2006). All the same, it seems that A1 responds primarily to sound, either heard or imagined.

A variety of cortical networks develop from the A1, but a discussion on what happens from here onwards is beyond the scope of this paper. One important distinction that appears in the areas immediately surrounding A1 on the superior temporal gyrus (STG) is between a "what" pathway and a "where" pathway[9], with "what" information being processed anterior to A1 on the STG, and "where" information being processed posterior to A1; these areas have also been found using fMRI in a study by Arnott and colleagues (Arnot et al., 2005).

3 Envisaged Analysis Tools

Having described the journey that sounds take through the auditory pathways from the ears up to the auditory cortices, in this section I will enumerate the analysis tools that could be developed based on the various stages of this journey and the types of information that they would elicit.

---

[9] There is an analogue of this in visual processing.

I would envisage at least four analysis tools:

- cochlearnucleigram
- olivogram
- thalamogram
- auditory corticogram

The *cochlearnucleigram* would give information related to the activity of the cochlear nuclei (Figure 3). It would provide a powerful source separation tool based on minute onset and spatial information. The cochlearnucleigram would be able to provide precise onset information and trace the behaviour of the sounds in the horizontal and in the vertical planes.

The *olivogram* would give information related to the activity of the superior olivary complex (Figure 4). It would provide further information about sound localisation using two types of mechanisms: one based on timing information and another based on sound intensity. Identification of sound location through the use of timing information would focus on low frequency sounds whereas the use of intensity information would focus on high frequency sounds.

The *thalamogram* would give information related to the activity of the medial geniculate nucleus (MGN) in the thalamus (Figure 6). The thalamus controls attention; it enables us to suppress information in order to focus on particular aspects of the sounds we hear. Therefore, the thalamogram would reveal salient sound attributes that would be deemed more important than others in function of specific contexts or conditions. I would imagine the possibility of being able to specify such contexts as analyses parameters for simulating the focus of the thalamus under different contexts or conditions. This would reveal the impact of different sensorial modalities on the auditory signal. I would imagine a number of interesting analysis parameters here; for example, the amount of modulation from other sensorial modalities, the types of modalities to be considered and/or ignored, modalities priorities, and so on.

Finally, the *auditory corticogram* would give information related to the activity of the primary auditory cortex (Figure 6). It would take the thalamogram further by building representations and anticipations. I would envisage some sort of *thalamocortical control panel* here, which would allow for building representations combining different levels of attention to various sound features, influences from other sensorial modalities, and exposure. Ultimately, the thalamocortical control panel would allow us simulate and predict the kinds of representations that would emerge by forging different ontologies and cortical plasticities. I would envisage the possibility of furnishing the analysis system with different listening strategies based on exposure to different sound worlds.

4 Concluding Discussion

In this paper I proposed the possibility of devising tools for the analysis of electroacoustic music based on neurophysiologic models of our auditory system. These tools would reveal the representations that our brain produces at various stages of the auditory pathway, from the cochlea to the cortices.

One of the principal contributions of the burgeoning fields of auditory neuroscience and neuroscience of music to our understanding of both music and the brain is the elucidation of various neural pathways through which the auditory signal is processed. There are various stages in this pathway, each of which extracts different types of information from the signal. Basically, once the sound enters our ears, our auditory system decomposes the signal into a number of attributes on its way up to the brain cortex. Then, at the actual "listening" stage, the brain puts all these attributes back together; this is referred to as *binding*. Interestingly, some of these attributes might be modulated by other sensorial information before the binding takes place. For example, it is believed that the well-known phenomenon of synaesthesia[10] is caused by such modulations. There are various other brain activities that are not very well understood that may modulate these attributes, such as memory and expectation. The amount of information that flows in the sensory-brain-behaviour circuit is immense, but we have evolved strategies to react to sensations as quickly as possible. From an evolutionary perspective, we cannot afford the delay that is would take to wire from scratch billions of neurons for every leap of consciousness. Memory, expectations and a number of other mechanisms optimise brain functioning[11] and much of this optimisation takes place in the auditory pathways; for example, by means of the descending pathways, as briefly explained in section 2.1. A better understanding of these mechanisms would lead to incredibly powerful *auditory corticograms*.

To the best of my knowledge, there have been no significant developments towards the tools I have suggested in this paper. However, research into building computational models of auditory brain functioning is an area that is evolving rapidly. I am convinced that progress will begin to emerge once these models begin to outperform current approaches to research into *semantic audio analysis[12]*. With this paper I hope to prompt the curiosity of music technology investigators and developers towards the fascinating possibilities of the neurotechnology approach to the analysis of electroacoustic music.

---

[10] Synaesthesia is a joining of the senses. Sensations in one modality (e.g., hearing) produce sensations in another modality (e.g., colour) as well as its own.

[11] A recent book by David Huron (2006) presents an interesting introduction to the role of expectation in music cognition. Although Huron's work tends towards "note-based" music, this book is certainly is a good starting point.

[12] See, for instance, the activity of AES Technical Committee on Semantic Audio Analysis, http://www.aes.org/technical/saa/.

Acknowledgements

References

Arnott, S. R., Grady, C. L., Hevenor, S. J., Graham, S. and Alain, C. (2005). "The Functional Organization of Auditory Working Memory as Revealed by fMRI". *Journal of Cognitive Neuroscience* 17(5):819-83.

Bear, M. K., Connors, B. W. and Paradiso, M. A. (2001): *Neuroscience: Exploring the Brain*. Baltimore, MD: Lippincott Williams Wilkins (*2nd ed.)*.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: The MIT Press.

Koelsch, S. and Siebel, W. A. (2005). "Towards a neural basis of music perception", *Trends in Cognitive Science* 9(12):578-584.

Rauschecker, J. P. (2005). "Neural Encoding and Retrieval of Sound Sequences", *Annals of the New York Academy of Sciences* 1060:125-135.

Schaeffer, P. (1966). *Traité Des Objets Musicaux: essai interdisciplines*. Paris: Éditions du Seuil.

Schofield, B. R. and Cant, N. B. (1999). "Descending Auditory Pathways: Projections from the Inferior Colliculus Contact Superior Olivary Cells that Project Bilaterally to the Cochlear Nuclei", *Journal of Computational Neurology* 409(2):210-23.

Serafine, M.-L. (1988). *Music as Cognition:The Development of Thought in Sound*. New York, NY: Columbia University Press.

Shore, S. E., Godfrey, D. A., Helfert, R. H., Altschuler, R. A. and Bledsoe, S. C. Jr (1992). "Connection between the cochlea nuclei in guinea pig", *Hearing Research* 62:16-26.

Smalley, D. (1997). Spectromorphology: Explaining Sound-Shapes. *Organised Sound***, 2(2):107-126.

Voisin, J., Bidet-Caulet, A., Bertrand, O., Fonlupt, P. (2006). "Listening in silence activates auditory areas: a functional magnetic resonance imaging study", *Journal of Neuroscience* 26(1):273-278

Worden, F.G. (1971). "Hearing and the neural detection of acoustic patterns", *Behavioural Neuroscience* 16:20-30.