



Inesperata accident magis saepe quam quae speres.  
(Things you do not expect happen more often than  
things you do expect) Plautus (ca 200(B.C.))



Project no: 027787

## DIRAC

### Detection and Identification of Rare Audio-visual Cues

Integrated Project  
IST - Priority 2

#### DELIVERABLE NO: D4.6 Categorization of Level and/or Temporal Dynamics

*Date of deliverable:* 30.06.2007  
*Actual submission date:* 10.08.2007

Start date of project: 01.01.2006

Duration: 60 months

*Organization name of lead contractor for this deliverable:*  
**Leibnitz Institute for Neurobiology**

Revision [1]

Project co-funded by the European Commission within the Sixth Framework Program (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other program participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	



Inesperata accident magis saepe quam quae speres.  
(Things you do not expect happen more often than  
things you do expect) Plautus (ca 200(B.C.)

## D4.6 CATEGORIZATION OF LEVEL AND/OR TEMPORAL DYNAMICS

Leibniz Institute for Neurobiology (LIN)

***Abstract:***

The aim of the present deliverable was the development of behavioral learning paradigm for rodents in which the discrimination along a to-be-determined stimulus dimension (level dynamics and/or temporal dynamics) should be trained that will later allow transfer to visual stimuli in the experiments designed in work package 5. The work reported here has demonstrated that the discrimination of repetition frequency of sequences of pure tones fulfils the requirements for this paradigm: (1) transferability to visual stimuli, (2) sufficient speed and reliability of learning, (3) amenability to detailed electrophysiological analysis after transferal of the experiment to the visual domain.

## Table of Content

1.	Introduction.....	4
2.	Material and Methods.....	5
2.1	Animals.....	5
2.2	Animals Preparation.....	5
2.3	Recordings.....	5
2.4	Training Procedure.....	5
3.	Behavioral Analysis.....	7
4.	Neuronal Analysis.....	8
5.	Conclusion.....	10
	References.....	10
	Annex I	
	The Cognitive Auditory Cortex: Task-Specificity of Stimulus Representations .....	11

## 1. Introduction

A central aim of DIRAC is the investigation of neuronal mechanisms of audiovisual integration for rare-event processing in biological organisms to make the underlying processing strategies available for artificial cognitive systems also. Towards this aim, one of the research strategies of our consortium consists in designing an experimental paradigm for rodents, in which stimuli from one sensory modality (audition) are first learned to be discriminated according to a to-be-determined stimulus feature (dimension), and subsequently, this stimulus feature is transferred to stimuli from another modality (vision). If this can be achieved, this would represent an animal model, thereby amenable to detailed and mechanistic physiological analysis, in which a cross-modal (audiovisual) category transfer is employed. While the latter rationale is more comprehensively explored in the work summarized in report "Deliverable D-5-4", it was the aim of the present work to first establish the *auditory* training paradigm (to be rooted in work package 4 of DIRAC's program) so that it can subsequently be transferred to work package 5 and combined with the audiovisual category transfer.

Consequently, the main task here was to define the feature dimension to be selected for the auditory stimuli that has the principal prospect of being transferable to visual stimuli also. At the outset we hypothesized that such cross modal transfer might be feasible for stimulus dimensions associated with (1) either the *temporal structure* of stimuli or (2) with *intensity relations* between stimuli (cf. description of work packages 4 and 5 in the Technical Annex). Features associated with both domains might be sufficiently abstract to allow abstract transferal to a qualitatively different sensory modality, because both domains can be independently attributed to both auditory and visual stimuli.

The specific aim was therefore to decide whether the cross modal transfer experiments planned in work package 5 should rather employ as the to-be-transferred dimension (1) temporal dynamics, (2) level (intensity) dynamics or (3) both. The latter option was envisaged if neither temporal dynamics nor level dynamics would prove to be suited for establishing a sufficiently clear discrimination in the unimodal (auditory) training. However, in that case, the required experimental design for the work package 5 experiment would have to be more complicated as it had to control for the effects of both domains (temporal and level) and their possible interactions. With respect to the first two options, temporal dynamics would be preferable over level dynamics, mainly because previous neurophysiologic work of DIRAC partner LIN has accumulated a body of physiological mechanisms of temporal processing of auditory signals (for overview see Ohl and Scheich 2005; Scheich et al. 2007) that might support analysis of the *crossmodal* processing mechanisms also.

We therefore first focused on an experiment in which repetition rates of auditory pure tone sequences have to be discriminated. Based on previous work on the neuronal coding of amplitude modulations in our laboratory (e.g. Schulze et al. 2002; Deutscher et al. 2007) we selected 0.2 Hz and 5 Hz as the two repetition frequencies to be discriminated. The present report demonstrates that this stimulus configuration and the chosen training paradigm are suited to allow a stable and significant discrimination by Mongolian gerbils, thereby fulfilling the prerequisites required for studying audio-visual transfer in workpackage 5 (cf. deliverable D-5-4).

Although electrophysiological recording of brain activity was not a focus of the present deliverable, it was nevertheless necessary to demonstrate that the training experiment would work with animals prepared for recording. This was necessary because recording of brain activity is required in the visual experiment and the preparation for recording cannot be interjected between the auditory stimulation phase and the visual stimulation phase when the cross modal transfer should be studied. Therefore, we slightly simplified the chronic

preparation (in that we implanted electrodes over the auditory cortex only and left out additional preparation of the visual cortex).

## **2. Material and Methods**

### **2.1 Animals**

We used four adult Mongolian gerbils (*Meriones unguiculatus*) weighing 75-95 g. The animals were obtained from the Tumblebrook Farms, West Brookfield, MA, USA, and housed in a temperature-controlled room (22°C) under a 12-hrs light-dark cycle (light on 06:00 - 18:00 h) in standard laboratory cages with free access to food (rodent pellets and sunflower seeds) and water. All experiments were approved by the Ethics Committee of the State of Sachsen-Anhalt, Germany.

### **2.2 Animals Preparation**

Animals were chronically implanted with rectangular  $2 \times 5$  arrays of epidural electrodes for recording of the electrocorticograms over right primary auditory cortex. Arrays were configured from epoxy-insulated stainless steel microwires (diameter: 100  $\mu\text{m}$ ; interelectrode distance: 600  $\mu\text{m}$ ) (Barrie et al. 1996; Ohl et al. 2000a,b, 2001). Prior to implantation, the surface of the arrays were filed to concave shape matching the convexity of the cortical surface. Implantation was performed under deep ketamine anaesthesia (xylazine, 2 mg per 100 g body weight, i.p.; ketamine, 20 mg per 100 g body weight, i.p.). After the operation animals were allowed to recover for 3 days before recording sessions began. Positioning of arrays during operation, and validation of correct positioning after the operation, based on the analysis of the spatial organization of pure-tone-evoked P1- and N1-ERP-components, was performed as described in Ohl et al. (2000a).

### **2.3 Recordings**

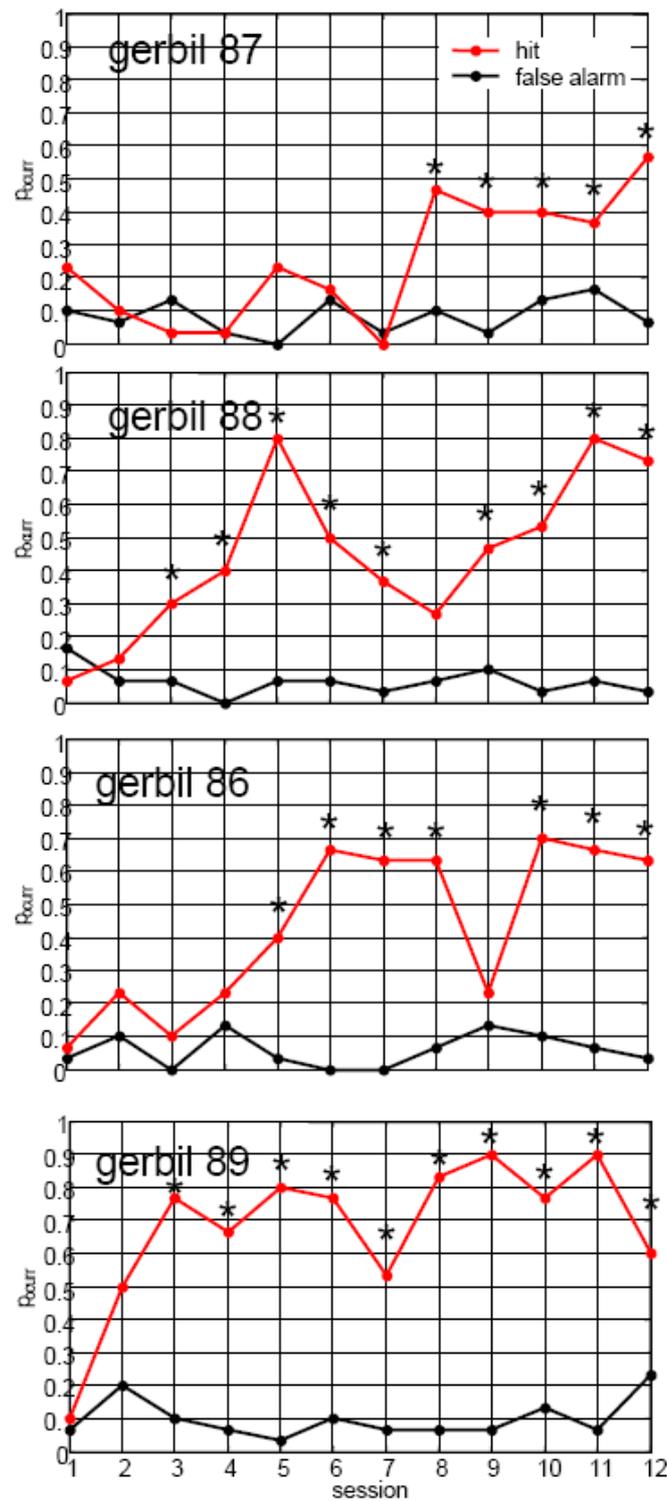
Electrocorticograms from the implanted arrays were recorded monopolarly against a mediofrontal reference electrode (amplification: 10,000; bandpass filtering: 0.1 Hz - 300 Hz) using the Plexon Multichannel Recording System (Plexon Inc., USA).

### **2.4 Training Procedure**

Gerbils were trained in a two-compartment shuttle-box (E10-15, Coulbourn Instruments) using a GO/(NO-GO) avoidance discrimination procedure. The compartments (16 cm  $\times$  18 cm  $\times$  22 cm) were separated by a 3-cm high hurdle. Mild electrical foot shocks (100 - 600  $\mu\text{A}$ ) applied through the floor grid served as unconditioned stimuli (US). Training was carried out in daily sessions consisting of 60 trials each, that is, 30 presentations of US-reinforced conditioned stimulus (CS+) and 30 presentations of non-reinforced conditioned stimulus (CS-) presented in randomized order after a variable intertribal interval of 15 - 19 s.

Each trial started with the presentation of a CS. After 6 sec the CS+ presentation was followed by a US unless the animal had crossed the hurdle during CS+ presentation, which would terminate CS+ presentation and trial and be counted as a hit response. The US lasted for up to 4 sec and could be terminated by the animal by crossing the hurdle. If the animal crossed the hurdle after presentation of the CS-, it was counted as a false alarm. Occurrences and times of hurdle crossings were monitored by a computer program. The CS stimuli consisted of short tone pulses (frequency: 2 kHz; intensity: 65 dB SPL) repeated with a repetition rate of 5.0 Hz in one experimental group (gerbils 86 and 89) or 0.2 Hz in a second group (gerbils 87 and 88). All stimuli were digitally synthesized (44.1-kHz sampling rate, 16-bit dynamic range) with 5 ms rise/decay times to reduce spectral splatter and presented after

digital-to-analog conversion using a commercial computer sound card from a loudspeaker mounted in the middle of the shuttle box roof.



**Figure 1.** Learning curves (hit and false alarm rates) of the four animals. The top two subjects "gerbil 87" and "gerbil 88" have been trained with 0.2 Hz as CS+ and 5.0 Hz as CS-. The bottom two subjects "gerbil 87" and "gerbil 88" have been trained with the reversed contingency. Asterisks indicate a significant difference between hit rate and false alarm rate (chi-squared,  $p < 0.05$ ).

### 3. Behavioral Analysis

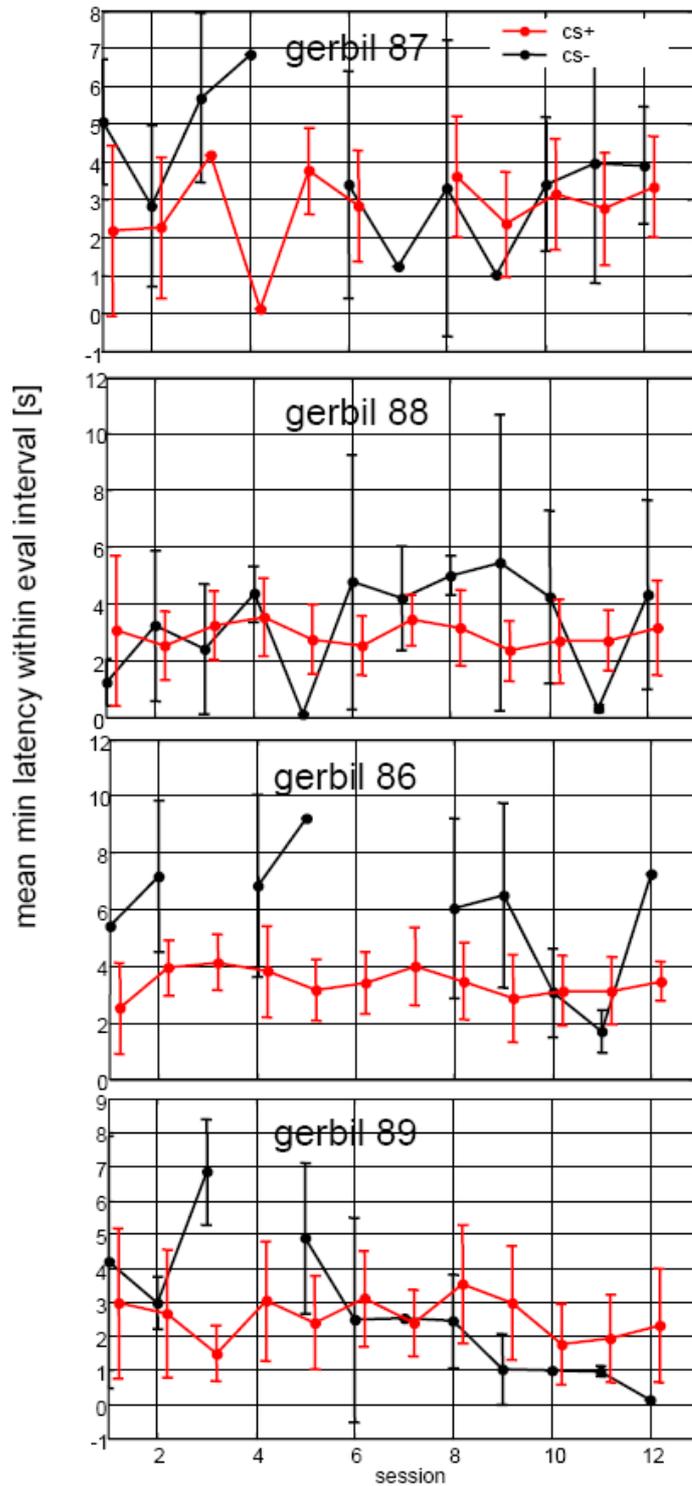
Of the four gerbils, two animals (gerbil 87 and gerbil 88) were trained using the slow repetition rate (0.2 Hz) as CS+ and the fast repetition rate (5.0 Hz) as the CS-, and two other animals (gerbil 86 and gerbil 89) were trained with the reversed contingency. Data were analyzed similar to our previous work (e.g. Ohl et al. 1999; Ohl et al. 2001).

Figure 1 shows the temporal development of the hit rate and false alarm rate over 12 training sessions for all four subjects. (No systematic difference between the two contingency schedules, i.e. slow or fast repetition frequency as the CS+, was apparent, or expected, although a rigorous demonstration of this point would require larger group sizes. Importantly, all four subjects reached a significant discrimination performance (chi-squared,  $p < 0.05$ ) between session 3 and 8.

Figure 2 displays the reaction time data for the four animal subjects. The comparison of the hit and false alarm rate data with the reaction time data supports the notion that gerbils can learn the discrimination of the chosen signal type and stimulus parameters with similar efficiency than previously studied auditory trainings (e.g. Ohl et al. 2001, 2004). The best-performing subject, gerbil 89, demonstrated reaction times in response to CS+ consistently  $< 4$  s (onset of the aversive US) for all training sessions. A similar pattern was evident in gerbils 86 and 88. But even the worst performer, gerbil 87, showed reaction times clustering around the 4 s (indicating either successful avoidance of the US or at least efficient escape shortly after US onset).

As the main result, it is demonstrated by the present experiments that the selected experimental paradigm is suited to reliably induce a behavioral state of significant discrimination of the temporal dynamics of the auditory input. This fulfills the requirement, imposed by the work plans of work packages 4 and 5, of establishing a uni-modal discrimination paradigm in which the relevant stimulus dimension of discrimination can be transferred to stimuli of another modality (cf. deliverable D-5-4).

**Figure 2 (next page).** Reaction times of the conditioned reactions (hurdle crossing in the shuttle box) in response to the CS+ and CS- stimuli for the four animal subjects.



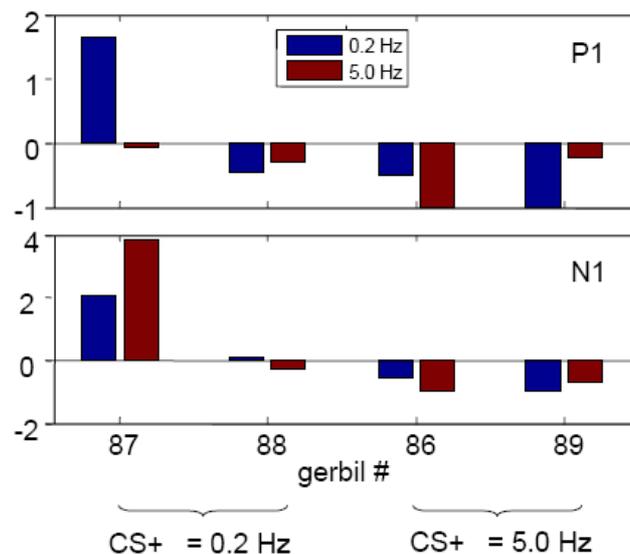
#### 4. Neuronal Analysis

Although recording electrodes (which are required for the crossmodal category transfer experiment described in workpackage 5) were mainly implanted to detect potential impediment of the behaviour in the shuttle box or the learning dynamics by this preparation, we nevertheless recorded the electrocorticogram from 10 electrodes (cf. sections 2.2. and 2.3) positioned epidurally over primary auditory cortex field AI (cf. Ohl et al. 2001).

**Figure 3** shows the relative changes

$$\Delta V = \frac{V_{post} - V_{pre}}{V_{pre}}$$

of ERP component peak amplitude before training  $V_{pre}$  and after training  $V_{post}$  for the four animal subjects. For the analysis, the response component to the 3<sup>rd</sup> pure tone pip in the sequence was used since the two stimuli cannot be discriminated with respect to the first tones in the sequences and discriminability increases with stimulus development. The 3<sup>rd</sup> pip occurs sufficiently late even for the fastest repetition rate to allow discrimination of the stimuli (cf. Schulze et al. 2002, Deutscher et al. 2006) and is present in the signal even for the slowest repetition rate.



**Figure 3.** Relative changes of ERP component peak amplitudes during training. Data are plotted from four animals, for the P1 and N1 component peak evoked by the 3<sup>rd</sup> pure tone pip in the sequence with 0.2 Hz repetition rate and 5.0 Hz repetition rate respectively.

Relative amplitude changes were found variable in the range from -1 to +2 (in a single case to +4). The direction of changes (relative increase or relative decrease) but not the absolute values tended to be correlated between the response change associated with the slow (0.2 Hz) repetition rate and the response change associated with the fast (5.0 Hz) repetition rate. In general, the patterns of changes observed in this experiment were not a straightforwardly interpretable as our recent results on learning-induced single unit retuning (cf. Scheich et al. 2007; and deliverable D-4-3). We obtained a similar heterogeneous pattern of results for the strength of synchronization between stimulus and field potential (as quantified by the vector strength of so-called "LFP-triggers" (e.g. Eggermont 1998) The heterogeneous patterns of results were expectable, however, as the number of subjects required for a full electrophysiological analysis typically exceeds that required for a significant behavioural analysis. The present results nevertheless indicate a potential relevance, not discussed so far, for research activities in workpackage 2 focussing on the amplitude-modulation structure of input (especially complex auditory input like speech).

## 5. Conclusion

The main aim of the present deliverable was to establish an auditory training paradigm with the following boundary conditions:

- (1) Auditory stimuli should be discriminated with respect to a stimulus dimension that can later be transferred to visual stimuli
- (2) The paradigm should be implemented in a species that is able to learn the discrimination with sufficient speed
- (3) The paradigm should allow detailed electrophysiological analysis once it is transferred to workpackage 5 and combined with the crossmodal transfer experiment.

The present work has demonstrated that all 3 requirements could be met. The discrimination of tone repetition frequency provides a paradigm in which the stimulus dimension relevant for discrimination can principally be transferred to visual stimuli. Implementation of that paradigm for the rodent species of the Mongolian gerbil allows for high speed of learning (in comparison to monkey) and detailed intracranial electrophysiological investigation (in comparison to human).

## References

- Deutscher A, Kurt S, Scheich H, Schulze H (2006) Cortical and subcortical sides of auditory rhythms and pitches. *NeuroReport* **17**: 853- 856.
- Eggermont JJ (1998) Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J Neurophysiol.* **80**: 2743-2764.
- Ohl FW, Wetzel W, Wagner T, Rech A, Scheich H (1999) Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learning & Memory* **6(4)**: 347-362.
- Ohl FW, Scheich H, Freeman WJ. (2000) Topographic analysis of epidural pure-tone-evoked potentials in gerbil auditory cortex. *J Neurophysio.* **83**: 3123-3132.
- Ohl FW, Schulze H, Scheich H, Freeman WJ (2000) Spatial representation of frequency-modulated tones in gerbil auditory cortex revealed by epidural electrocorticography. *J Physiol (Paris)* **94**: 549-554.
- Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. *Nature* **412**: 733-736.
- Ohl FW, Scheich H (2004) Fallacies in behavioural interpretation of auditory cortex plasticity. *Nature Reviews Neurosciences* (reviewed online comment, doi:10.1038/nrn1366-c1).
- Ohl FW, Scheich H (2005) Learning-induced plasticity in the auditory cortex. *Curr Opin in Neurobiol* **15**: 470-477.

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW (2007) The cognitive auditory cortex: task-specificity of stimulus representations. *Hear Res* **229**: 213-224.

Schulze H, Hess A, Ohl FW, Scheich H (2002) Superposition of horseshoe-like periodicity and linear tonotopic maps in auditory cortex of the Mongolian gerbil. *Eur J Neurosci* **15**: 1077–1084.

## **Annex I**

### **The Cognitive Auditory cortex: Task-Specificity of Stimulus Representations**

Research paper

# The cognitive auditory cortex: Task-specificity of stimulus representations

Henning Scheich<sup>a</sup>, André Brechmann<sup>b</sup>, Michael Brosch<sup>a</sup>, Eike Budinger<sup>a,c,\*</sup>,  
Frank W. Ohl<sup>d,e</sup>

<sup>a</sup> Leibniz Institute for Neurobiology, Department of Auditory Learning and Speech, Brenneckestraße 6, D-39118 Magdeburg, Germany

<sup>b</sup> Leibniz Institute for Neurobiology, Special Lab. Non-invasive Brain Imaging, Brenneckestraße 6, D-39118 Magdeburg, Germany

<sup>c</sup> Clinic of Neurology II, Otto-von-Guericke-University Magdeburg, Leipziger Straße 44, D-39120 Magdeburg, Germany

<sup>d</sup> Institute of Biology, Otto-von-Guericke-University Magdeburg, Germany

<sup>e</sup> Leibniz Institute for Neurobiology, Research Group Neuroprosthetics, Brenneckestraße 6, D-39118 Magdeburg, Germany

Received 28 November 2006; received in revised form 17 January 2007; accepted 31 January 2007

Available online 12 February 2007

## Abstract

Auditory cortex (AC), like subcortical auditory nuclei, represents properties of auditory stimuli by spatiotemporal activation patterns across neurons. A tacit assumption of AC research has been that the multiplicity of functional maps in primary and secondary areas serves a refined continuation of subcortical stimulus processing, *i.e.* a parallel orderly analysis of distinct properties of a complex sound. This view, which was mainly derived from exposure to parametric sound variation, may not fully capture the essence of cortical processing. Neocortex, in spite of its parcellation into diverse sensory, motor, associative, and cognitive areas, exhibits a rather stereotyped local architecture. The columnar arrangement of the neocortex and the quantitatively dominant connectivity with numerous other cortical areas are two of its key features. This suggests that cortex has a rather common function which lies beyond those usually leading to the distinction of functional areas. We propose that task-relatedness of the way, how any information can be represented in cortex, is one general consequence of the architecture and corticocortical connectivity. Specifically, this hypothesis predicts different spatiotemporal representations of auditory stimuli when concepts and strategies how these stimuli are analysed do change. We will describe, in an exemplary fashion, cortical patterns of local field potentials in gerbil, of unit spiking activity in monkey, and of fMRI signals in human AC during the execution of different tasks mainly in the realm of category formation of sounds. We demonstrate that the representations reflect context- and memory-related, conceptual and executional aspects of a task and that they can predict the behavioural outcome. © 2007 Elsevier B.V. All rights reserved.

**Keywords:** Anatomy; Behaviour; Categorization; Discrimination; Functional magnetic resonance imaging

**Abbreviations:** AI, primary auditory field; AAF, anterior auditory field; AC, auditory cortex; BF, best frequency; BOLD, blood oxygenation level dependent; CM, caudomedial field; FM, frequency modulated/modulation; fMRI, functional magnetic resonance imaging; IC, inferior colliculus; icv, inferior cerebral vein; mca, middle cerebral artery; MGB, medial geniculate body; RM, reference memory; SC, superior colliculus; VTA, ventral tegmental area; WM, working memory

\* Corresponding author. Address: Leibniz Institute for Neurobiology, Department of Auditory Learning and Speech, Brenneckestraße 6, D-39118 Magdeburg, Germany.

*E-mail addresses:* [scheich@ifn-magdeburg.de](mailto:scheich@ifn-magdeburg.de) (H. Scheich), [brechmann@ifn-magdeburg.de](mailto:brechmann@ifn-magdeburg.de) (A. Brechmann), [brosch@ifn-magdeburg.de](mailto:brosch@ifn-magdeburg.de) (M. Brosch), [budinger@ifn-magdeburg.de](mailto:budinger@ifn-magdeburg.de) (E. Budinger), [frank.ohl@ifn-magdeburg.de](mailto:frank.ohl@ifn-magdeburg.de) (F.W. Ohl).

## 1. Introduction

The auditory pathway, like all sensory pathways in mammalian species, terminates in a specific part of the neocortex, a unique six-layered structure that covers almost the complete surface of the forebrain. The organizational principle of neocortex is presumably most significant for mammalian brain evolution as it has allowed a thousand-fold expansion of neocortex from primordial mammals to man concomitant with acquisition of numerous functions but without obviously changing its fundamental architecture. It seems to have grown essentially by multiplication

of small columnar networks of neurons, *i.e.* modules which represent structural–functional units (*e.g.*, Jerison, 1973; Szentagothai, 1978; Creutzfeld, 1983; Mountcastle, 1997).

Neocortex not only comprises sensory terminal areas but also contains the sources of voluntary motor pathways and any higher multisensory or cognitive function that mediates between sensory experience and controlled behaviour. Interestingly, in spite of this mosaic of extremely diverse functions, chiefly reflecting the thalamic inputs and long-range subcortical outputs, convergent evidence from embryological development, from adult cytoarchitectonics and from the study of extrinsic and intrinsic connectivities suggests that structural differences across neocortex mainly consists of quantitative details of a stereotyped architecture (*e.g.*, Creutzfeldt et al., 1977; Gatter and Powell, 1978; Rockel et al., 1980; Douglas and Martin, 1994; Buxhoeveden and Casanova, 2002; Rakic, 2002; Silberberg et al., 2002). Thus, commonalities of structure, and therefore presumably of some aspects of information processing, are larger than the differences, especially in comparison to the plethora of diverse organizational types of corresponding nuclei at subcortical levels. In a search for functional parcellation of neocortex especially with modern imaging methods, dissimilarities rather than these similarities have been emphasized (*e.g.*, König et al., 2005, chapter 1; but see also Eggermont, 1998).

In trying to explore functional specializations of auditory cortex (AC) with respect to subcortical auditory stations, it seems important to draw reasonably general conclusions on these similarities across neocortex. The mechanistic strategies how an auditory scene is processed cortically cannot be entirely different from the processing of a visual scene, from the organization of motor programs, or from the orchestration of cognitive decisions, but might be vastly different from subcortical auditory processing. Processing in stations of the subcortical auditory pathway can be conceptualized by a succession of convergent or divergent ascending input connections of neurons, thus by a hierarchy of bottom-up processing of cochlear information. It will be argued below and supported by data that, starting with primary AC, auditory cortical fields receive a majority of their input connections from other cortical fields including non-auditory areas. This suggests that their ascending auditory input comes under the massive influence of top-down information. In this global perspective, sound processing by the AC cannot simply represent the continuation, however exquisite, of a certain analysis of sound properties (auditory feature analysis). Claims in this direction have already been weakened by a comparison of neuronal sound feature distinctions in different auditory cortical fields which revealed that neurons show rather quantitative than qualitative differences in their responses, *i.e.* similarities of fields in this respect seem to be larger than the differences (Eggermont, 1998).

Based on previous accounts (*e.g.*, Scheich et al., 1994, 2005; Brechmann and Scheich, 2005; Ohl and Scheich, 2005; Fritz et al., 2005), and recent results (Selezneva

et al., 2006; Brechmann et al., 2007) summarized below, we propose that the role of AC is to actively generate sound stimulus representations that are task-specific. This proposition suggests that spatiotemporal activation patterns of neuronal ensembles in AC, passively generated by a given stimulus and basically reflecting all features of a stimulus, can be modified according to the context and the procedural and cognitive demands of a listening task, *i.e.* also reflect semantic aspects of a stimulus. Namely, detection or recognition of a particular sound feature of a stimulus will generate a representation of that stimulus different from the representation during the discrimination of features of that stimulus from similar features of another stimulus (*e.g.*, Scheich et al., 2005). Other representations will result when multiple comparisons are made or when a category of stimuli with common abstract relationships of features is formed (*e.g.*, Ohl et al., 2001; Selezneva et al., 2006). Thus, cortical representations of sound stimuli seem to be constructs related to the cognitive operations performed with these stimuli and they even change with the individual performance of a given task as shown below.

Furthermore, in tasks that are based on complex procedural guidelines for execution and that require a large amount of training, even the procedural contingencies of the task become represented in AC in addition to the changed stimulus representations (Brosch et al., 2005).

All these influences are unlikely to originate in the AC but seem to be generated by interaction of the AC with multiple other brain areas, *e.g.*, multisensory lateral prefrontal areas where task-relevant information is selected (Brass and von Cramon, 2004). The new aspect is that the AC is not the pattern analysis slave of hierarchically higher order cortical regions but, by its multiple cortical interactions, seems to be in flexible local control of its own analysis mechanisms. In a more general sense, task-specific representations of information, which are constructs containing a combination of actual information and of memory-related, conceptual, and executional information, may be a principle that is shared by cortical regions of apparently different functions.

## 2. Architecture of AC facilitates task-specific functions

The AC, including the primary field AI as well as several non-primary fields, shares a columnar arrangement of neurons across the six cortical layers with most areas of the neocortex like other sensory, motor, and so-called associative cortex areas. Cortical columns, often either described as functional macrocolumns (*e.g.*, ocular dominance columns in visual cortex, Hubel and Wiesel, 1977; frequency columns in AC, Scheich et al., 1993) or as anatomical microcolumns (*e.g.*, Mountcastle, 1997; Buxhoeveden and Casanova, 2002), are characterized by the arrangement of large pyramidal neurons in deep layer V (Vb) with apical dendrites that span almost the whole thickness of cortex orthogonal to the layers (*e.g.*, Goldschmidt et al., 2004). They are surrounded by various types of non-pyramidal

neurons in all layers and by smaller pyramidal neurons in infragranular (V–VI) and supragranular layers (I–III), but particularly in layer III. Layer III pyramidal neurons have shorter apical dendrites, but like for the layer V pyramidal neurons, apical dendrites are all roughly parallel to each other and they all reach the surface layer I (e.g., Szentagothai, 1978; Winer, 1992). Because of this geometry, apical dendrites of the different pyramidal cells more or less spatially overlap in the supragranular layers. Thus, the different pyramidal cells could directly integrate nearly all corticocortical inputs because these inputs predominantly terminate also in the supragranular layers (e.g., Bannister, 2005). In addition to the input integration across the layers by the apical dendrites, incoming inputs within the layers are integrated by the basal dendrites of the pyramidal cells. The outputs of the pyramidal cells via their extrinsic axonal projections are also rather layer-specific. The laminar pattern of these axonal terminations within a given target area as well as the laminar location of cells of origin within a given source structure indicates the type of corticocortical connectivity, *i.e.* the feedforward, lateral, or feedback type (e.g., Felleman and van Essen, 1991; Rouiller et al., 1991).

Although the functional impact of the cortical architecture has not yet become fully transparent, the study of the cortical and subcortical input sources to the different layers, for example, of the AC and the output targets from the different layers of the AC sheds much light on this problem. Thalamocortical inputs to the AC from the ascending auditory pathway cover only a minor fraction of all inputs, even to the primary field AI, but other forebrain and especially cortical sources make up the lions share of the available inputs (e.g., Martin, 2002; Winer et al., 2005; Winer and Lee, 2007; Fig. 1). The core thalamocortical input, which is largely restricted to mid granular layer IV, is yet extremely powerful in generating stimulus-specific activity in a column (e.g., Miller et al., 2001; Swadlow et al., 2002; Metherate et al., 2005). This input obviously constitutes a functional division between the corticocortical inputs to the superficial supragranular layers and the commissural inputs to the deeper infragranular layers, particularly layer V (e.g., Winer, 1992; Budinger and Heil, 2005).

Some principles of the way in which cortical options of task-specific auditory processing are anatomically possible are suggested by our recent studies of auditory cortical connectivities of AC in the Mongolian gerbil (Budinger et al., 2000a; Budinger et al., 2000b; Budinger et al., 2006; Fig. 1). Besides the qualitative information about which subcortical and cortical structures are connected with the various auditory cortical fields, we also obtained initial quantitative data about the input connections into AI (summarized in Fig. 1). This was achieved by determining distribution and densities of tracer-labelled cells in various brain regions after injections of retrograde tracers (usually dextranamines) which were made into AI.

About one third (34%) of all neurons, projecting to AI, are part of ascending inputs from subcortical structures.

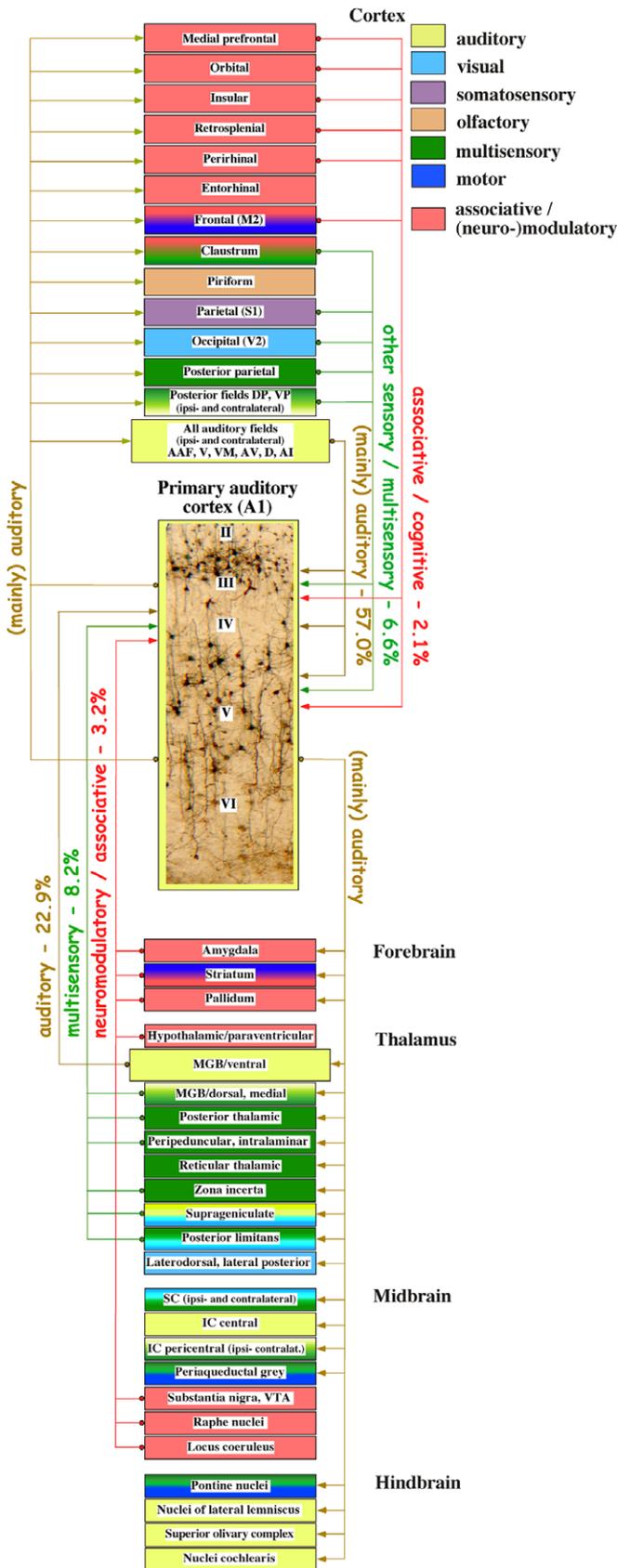
Their projections terminate chiefly in layer IV but also in other layers. The majority of these subcortical inputs, *viz.* 23% of all inputs into AI, is of specific auditory modality and originates exclusively in the ventral division of the medial geniculate nucleus. Further 8% of the total inputs are of mainly multisensory and further 3% are of mainly associative or modulatory subcortical modality. These inputs stem from the thalamus (e.g., from other sensory and multisensory thalamic nuclei), from neuromodulatory structures of the midbrain, and from deep forebrain structures like the amygdala, striatum, and pallidum.

On the other hand, about two thirds (66%) of all neurons, projecting to AI, are part of descending inputs from other cortical regions. These projections terminate chiefly in the supragranular, but also in infragranular and granular layers. Again, the majority of these inputs, *viz.* about 57% of all inputs into AI, is of mainly auditory modality and stems from surrounding auditory and auditory-related fields in the ipsilateral and contralateral hemisphere. The remaining 9% of the inputs into AI are from a variety of other sensory (visual occipital, parietal somatosensory cortex), multisensory (e.g., posterior parietal cortex), motor (frontal cortex), and associative/cognitive cortical regions (e.g., prefrontal, orbital, insular, perirhinal cortex). All these cortical regions also receive largely reciprocal input from AI which chiefly originates from supragranular and to some extent also from infragranular layers in AI.

In conclusion, the majority of inputs into AI is (1) of cortical origin and (2) terminates mainly in supragranular layers where (3) they can be integrated directly by the apical dendrites of pyramidal neurons from all layers. Thus, reciprocal corticocortical feedforward and feedback connections are most likely the dominant factor that influences the activity of the neurons in AI. These connections presumably do not directly elicit sound-stimulus-driven firing in AI, which is most likely the role of the auditory thalamic inputs. Instead, the cortical inputs may substantially modulate and thereby specify this activity with respect to any task demand. This includes mainly the inputs from the other non-primary auditory cortical fields which are likewise under a similar or even stronger modulatory influence of various non-auditory corticocortical and subcortical connections (Budinger et al., 2000a; Budinger et al., 2000b). In the following some types of these modulatory influences will be illustrated. In addition, it will be shown that, in the case of extensively trained monkeys, the non-auditory influences can become strong enough to drive auditory neurons in AC.

### 3. Categorization of FM direction in gerbils

To explore task-relatedness of sound representation in AC various detection or discrimination tasks exploiting overt stimulus properties could be used. But it is generally not *a priori* certain whether neuronal processing, relevant for solving a given task, really is localized in AC or is already mediated by subcortical processing transmitted to



of sound properties that indeed do require the formation of an abstraction in AC.

In linear frequency modulated (FM) tones various stimulus dimensions, like onset and offset frequency, covered frequency range, steepness of modulation, duration, and intensity, can be varied, but any FM tone will always fall into one of two natural categories, *viz.* “rising” or “falling”, depending on the sign of the variation of instantaneous frequency with time. The discrimination of the modulation direction of FM tones in gerbils depends on AC and even more specifically on right AC as shown by bilateral (Ohl et al., 1999) and unilateral AC lesion experiments (Wetzel et al., 1998). In shuttle box go – no go experiments gerbils could neither learn to discriminate rising from falling FM tones nor could they remember the discrimination learned before the lesion without right AC, even though they could learn to discriminate pure tones without AC. This discrimination presumably involves an interaction, among others, between medial prefrontal cortex and AC. This is because the first successful discriminations that lead to an avoidance of footshocks in the shuttle box are accompanied by a transient dopamine release in medial prefrontal cortex of gerbils and because the discrimination behaviour can be blocked by local application of D1 receptor antagonists (Stark et al., 2004).

Following demonstrations that gerbils could discriminate modulation direction using symmetrically rising and falling FM tones, it was investigated whether they could transfer the learned discrimination of direction to other exemplars of FM tones, *i.e.* to form the concept of modulation direction with the categories “rising” and “falling” independent of other parameters (Ohl et al., 2001). The shuttle box paradigm was combined with chronic epidural recording of surface potentials over AC using an 18 electrode array (Fig. 2a). As expected, learning to discriminate

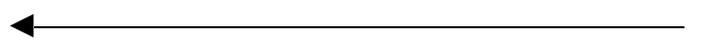


Fig. 1. Schematic diagram of the connections of the primary auditory field (AI) of the Mongolian gerbil as established by the retrograde and anterograde transport of biocytin (*e.g.*, Budinger et al., 2000a; Budinger et al., 2000b) and dextranamine (*e.g.*, Budinger et al., 2006) injected into AI. Each box represents a brain structure which is connected with AI. Different colours are used to identify the main modality of these structures (legend). The origin of a projection is represented by a filled circle, the target by an arrow. Yellow, green, and red colour is used to identify the type of information which is most likely processed by these projections, *i.e.* auditory, other sensory/multisensory, or associative/cognitive/(neuro) modulatory information. The ratios (in percent) of inputs assigned to one of these information streams on the total input into AI are also indicated. Ratios were estimated from the relative number of retrogradely labelled cell bodies in the structures contributing to the respective stream. Connections between all other structures than AI, for example between structures of the auditory pathway, are not illustrated (for review and models see, *e.g.*, Weinberger, 2004; Suga and Ma, 2003). Connections of AI with the medial prefrontal cortex include separate connections with anterior cingulate, prelimbic, and infralimbic areas; striatum includes caudate putamen and accumbens nucleus; pallidum includes globus pallidus, ventral pallidum, and nucleus basalis of Meynert. Golgi-staining of AI neurons – courtesy of Judith Mylius, IFN.

AC resulting in some form of observable modification of local representations of sounds there. We have demonstrated in Mongolian gerbils some forms of categorization

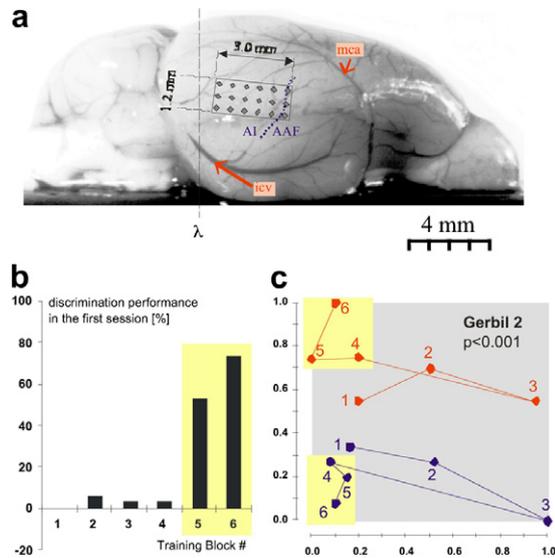


Fig. 2. Preparation (a), psychophysical (b), and electrophysiological results (c) of the gerbil FM tone categorization experiment. (a) A  $3 \times 6$  array of epicortical electrodes was chronically implanted over primary auditory cortex, field AI (caudal to field AAF), to cover the tonotopic representation of frequencies from 0.25 to 16 kHz. (b) Abrupt onset of categorization (in training block 5), *i.e.* correct go/no-go responses to novel stimuli introduced in the first session of a training block and (c) mutual similarity relationships between marked states activation patterns (see text) measured for the categories rising (red) and falling (blue) at each of six consecutive training blocks (numbers). Displayed is a nonlinear projection of the 18-dimensional marked state vectors onto the 2 dimensional plane (in arbitrary units) allowing visualization of the similarity relationships. The spatial distance between any two points is a measure of the dissimilarity of the corresponding activation patterns. The yellow rectangles emphasize the formation of similarity clusters arising with categorization indicative of a map-like organization of neural activity. (Modified from Ohl et al., 2001.)

a new pair of FM tones in a different frequency range required several sessions like in the first pair. But after 2–6 training sessions with additional novel FM tone pairs, the next novel pair presented was correctly discriminated from the start. This transfer of the learned classification responses to novel, previously unheard, stimuli indicated the formation of the correct respective categories. After this transfer individuals categorized any new FM tone, *i.e.* the formed categories were stable for the rest of the training experience (unless the rules of the training were altered).

The emphasis of the analysis of the simultaneous multi-channel recordings was on comparisons of single trials. Reaction times in each individual, especially during the categorization stage, varied considerably such that it was not clear whether any AC activity, indicating direction categorization, was stably related to FM stimulus onset and might still be identifiable after averaging over multiple trials. The method consisted of calculating an 18-dimensional state vector of power in the beta and gamma band of the brain potentials across the recorded channels. Using a sliding time window, epochs were identified after stimulus onset in the time series of such vectors (“trajectories”) such

that the Euclidian distance between vectors associated with rising and falling FM tones, respectively, was maximized. As a result of this procedure, an epoch with reliably short delay after the stimulus presentation was identified in which the state vectors for rising and falling FM sweeps in naive and trained animals deviated strongly. These state vector epochs corresponded to delays of mid-latency evoked potentials and thus to the directly stimulus-related differences of evoked activity.

However, in animals trained on FM tone discrimination additional later epochs of state vector deviation occurred and became more and more frequent with training on further FM pairs, until they were reliably present during every trial when animals categorized. These “marked states” identified in a range of 100 ms to 2 s after stimulus onset, *i.e.* in a period before the animals actually jumped, had no fixed latency from trial to trial in each animal and between animals (Ohl et al., 2001). By the correspondence of their occurrence with the emergence of discrimination performance they potentially qualified as states in which a cognitive analysis of stimuli could occur. Indeed, as shown in Fig. 2c, an analysis of the similarity relations between state vectors found during these marked states in each animal across trials and trained FM pairs revealed the development of an AC system property that paralleled the individual transition to categorization behaviour (Fig. 2b). Points in the two-dimensional map in Fig. 2c correspond to state vectors (spatial activity patterns during a marked state) and similarity between any pair of vectors is coded by the proximity of the corresponding points within this map. That is, similar vectors (patterns) are mapped into closely neighbouring points, more dissimilar vectors are mapped to spatially more separated points in this graph. It is evident from this representation of data that cortical activity states during the initial training with 4 pairs are dissimilar between rising and falling sweeps as well as between different pairs. However, when the animal behaviourally categorized modulation direction starting with the 5th pair, states of the same category (“rising” or “falling”) became very similar to each other while state pairs corresponding to different categories became maximally dissimilar independent of the respective stimulus similarities measured in physical stimulus dimensions. In every animal this state transition of the dynamics of AC activity occurred exactly when the animal started to behaviourally categorize the FM sweeps, *i.e.* after training experiences with 2–6 FM pairs.

The results suggest an isomorphism between the psychological phenomenon of category formation and the neuronal dynamics of the AC system. When a category is formed from different objects, many properties of exemplars may be non-specific for the category because they are also properties of exemplars of other categories (*e.g.*, in apples and pears). The psychological dynamics of forming the category seems to consist of emphasizing the category-specific properties and deemphasize the non-specific properties which make exemplars more similar than they really are.

Obviously the AC dynamics reflect this process as during discrimination training state patterns are dissimilar for different exemplars, but finally become similar, *i.e.* “run into an attractor” in the terminology of dynamics.

#### 4. Directional categorization of FM in human AC

The results obtained on cortical mechanisms of FM directional discrimination and categorization with the gerbil model were followed up by functional magnetic resonance imaging (fMRI) in human subjects. It was examined whether the reliance of FM directional discrimination on the AC in the right hemisphere is a more general phenomenon of the mammalian brain. Furthermore, due to flexible instruction-guided experimental options in human subjects, the question of task-dependence of stimulus representation could be more specifically addressed. Normal human subjects will instantly discriminate symmetrically rising and falling FM pairs and, by instruction, will sort series of widely varied FMs into the two categories by a mouse-key press. Thus, they do not need to be trained to establish the two categories by pairwise comparison of exemplars and by feedback as to a different behavioural meaning.

In a first attempt to capture directional categorization-related representations of FMs with fMRI in human AC, a comparison was made between mere exposure to a series of FM sweeps (unknown to the subject) and subsequent categorization of the same stimuli upon instruction (Brechmann and Scheich, 2005). FM sweeps are relatively powerful stimuli eliciting strong blood oxygenation level dependent (BOLD) responses in the primary and secondary auditory cortical areas that have been previously distinguished on the dorsal surface and lateral aspect of the superior temporal gyrus (Gaschler-Markefski et al., 1998; Scheich et al., 1998; Baumgart et al., 1999; Brechmann et al., 2002; Deike et al., 2004). Using a low-noise fMRI scanning procedure, the comparison of exposure-related

activation with silent intervals between stimulus blocks showed across subjects either an even distribution of activity between the two hemispheres or a more left-lateralized AC activation (Fig. 3A). In contrast, a subsequent categorization of the same stimuli in the same subjects lateralized the activation to the right AC (Fig. 3B).

These results are, of course, not yet evidence that the categorization of direction of the FM stimuli caused the rightward lateralization but could be due to any task involving these stimuli in comparison to mere exposure. Therefore, in a second experiment, FM stimuli were designed which allowed an additional type of categorization, namely a categorization according to long or short durations of the FM sweeps. The comparison of the two experiments (directional categorization *vs.* durational categorization of the same stimulus material) provided the evidence that activations, at least in the relation of these tasks, were task-specific. The task-specific effect was localized to secondary auditory cortical areas (T3) on the *planum temporale* of the subjects. In right T3, the activity caused by directional categorization was significantly stronger than that caused by durational categorization and *vice versa* in left T3 (Brechmann and Scheich, 2005).

The localization of specific activity related to directional categorization in a subarea of the AC on the *planum temporale* allowed to spatially isolate another aspect of the task-relatedness of the stimulus representation. Subjects, for the first time confronted with the directional categorization task, varied widely in their performance measured by the sensitivity index  $d'$  using signal detection theory (Swets et al., 1961). When the performance was compared with the BOLD activation a significant negative correlation was found in right T3 (Fig. 4a). Subjects with the highest performance showed the most spatially confined activation in this area. This, at first glance counterintuitive result could have several reasons. Therefore, we specifically tested in initially bad performers whether they improved on the categorization task by training. This was indeed the case

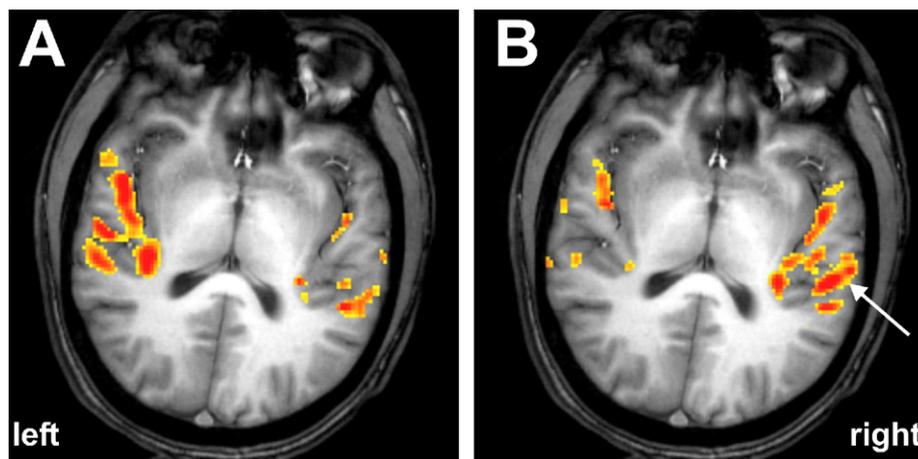


Fig. 3. Pattern of fMRI activation in the auditory cortex of a single subject for exposure to FM (A) and for categorization of FM-direction (B). Arrow points to a cluster of activation in the right AC belonging to the territory T3 which is specifically involved in the processing of FM-direction. (Modified from Brechmann and Scheich, 2005.)

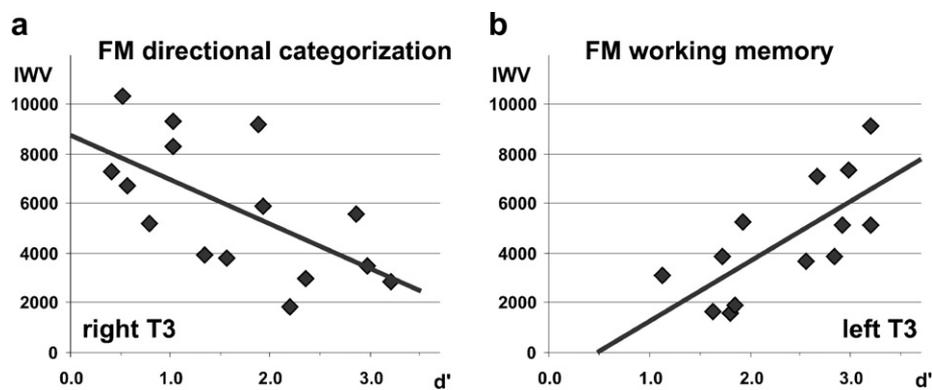


Fig. 4. Correlation of fMRI activation and sensitivity index  $d'$ . Directional categorization of FM led to an inverse correlation of performance with activity in right T3 on *planum temporale* ( $r = -0.63$ ;  $p = 0.013$ ) (a). In contrast, a working memory task, in which pitch and direction of FM had to be compared in a 2-back paradigm, revealed a positive correlation of performance with activation in left T3 ( $r = 0.72$ ;  $p = 0.006$ ) (b). IWV: Intensity weighted volume, i.e. activated volume multiplied by the mean BOLD response amplitude. (Modified from Brechmann et al., 2007.)

and repeated fMRI showed that their degree of right T3 activation decreased concomitantly. Thus, individual improvement with learning showed basically the same negative correlation with activation as the distribution of performances across a group of subjects. The effect is obviously due to proficiency that is sensitive to learning.

While this result suggests that the activity in right T3 is task-specific to the directional categorization, one could also speculate that this result is related to the findings on category formation in gerbil right AC. The dynamics of similarity of electrical states in gerbil AC, concomitant with categorization proficiency, and the negative correlation of BOLD activity in human AC, concomitant with task proficiency, could have similar underlying mechanisms. A concept explaining both effects might be the initial involvement of a network of various classes of auditory cortical neurons that are sensitive to FM sweeps and that might be individually tuned to quite different parameters of the sweep like frequency range, steepness of modulation duration, and direction. With higher proficiency, the overall activation might become reduced to already existing selective neurons that distinguish the direction of FM sweep independent from a wide range of other parameters. Alternatively, a network may be formed by neurons that have become most sensitive to that distinction by training.

In any event, the negative correlation of performance and T3 activation in the fMRI categorization experiments is a task-specific correlate of plasticity as shown by further fMRI experiments using the same class of FM sweeps but a different task (Brechmann et al., 2007). A series of FMs, varying in direction and frequency, can also be used for working memory (WM) tasks. In this case, each actually occurring FM sweep had to be compared with a sweep that occurred two events earlier (2-back matching to sample WM task). This procedure had to be carried out sequentially with all stimuli and matching events were indicated by mouse-key press. As control condition we applied a reference memory (RM) task using the same series of stimuli as in the WM task. The subjects had to identify one partic-

ular FM sweep which was learned in the beginning of the session. The number of possible matches was the same in both tasks. To make the identification easier in the RM task only six FM sweeps with varying direction and frequency range were used in both experiments.

The results of these experiments were different from those in the FM directional categorization task; not only with respect to the performance-related activation but also with respect to the hemisphere that was dominantly involved. The WM task produced significantly stronger activation than the RM task and this effect was exclusively found on the left *planum temporale* (left T3). Furthermore, the WM task-related activity on left T3 showed a positive correlation with task performance of subjects (Fig. 4b). Subjects with the best performance exhibited most activation in left T3 while right T3 showed no correlation with performance. While this left T3 correlation with performance is a clear indication of the task-specificity of the effect, it represents a strong contrast to the task-specific converse correlation in right T3 obtained with the directional categorization task (Fig. 4a).

Also for this left T3 effect some neuronal explanation is available. The solution of WM tasks has been related to properties of delay-specific neurons originally found in prefrontal cortex (Fuster and Alexander, 1971) but later also in other association and sensory cortical areas including AC (Gottlieb et al., 1989; Pasternak and Greenlee, 2005; Ranganath and D'Esposito, 2005; Sakurai, 1994). The activity of these neurons lasts longer than the duration of the activating stimuli and is maintained until following stimuli occur. The time course of that maintenance in single neurons seems to be stimulus-related and task-specific leading to the term "WM-signature neurons" (Constantinidis and Procyk, 2004; Passingham and Sakai, 2004). Quite in contrast to the mechanistic neuronal concept of a categorization, where a memory of individual members of a category does not help, it may be assumed that a 2-back WM task is best performed when sufficient members of WM-signature neurons maintain the memory of preceding stimuli in order to secure the

recognition of a match with an ongoing stimulus. This could lead to a recruitment of many suitable WM-signature neurons in best performers of the task. Various correlations of fMRI and PET (positron-emission tomography) activations in AC with task performance of subjects have also been obtained in other laboratories. There, seemingly identical tasks upon close comparison reveal different strategies of their solution which may explain the different correlations (e.g., Ohl and Scheich, 2005).

## 5. Categorization of tone sequences in monkey AC

The results described above demonstrated that the plasticity of stimulus representation in AC was task-dependent. In the gerbil categorization experiments and in the fMRI experiments with human subjects the levels of observation were field potentials or BOLD responses, which seem to be intimately related to field potentials in cortex (Logothetis et al., 2001). Thus, the evidence relied on global network properties rather than on spike discharges of cortical neurons. Furthermore, except for the WM task with FM, the tasks were limited to the categorization of single event stimuli, the interaction aspect of multiple stimuli or their serial order was not important.

The following experiments in macaque AC served to extend the insight into cortical categorization mechanisms (1) to the level of information provided by the firing of neurons and (2) to categorization that depends on the sequence aspect of stimuli (Selezneva et al., 2006). Tone sequences at particular frequency intervals are characteristic of melodies. One of the most important cues for recognizing melodies is the contour of tone sequences with rising or falling frequency steps (up or down steps) or constant frequency (flat) steps (Dowling, 1978). If melodies are transposed in frequency, they are still recognized as the same, a process requiring categorization of such steps. In some sense this categorization in the frequency domain is related to the FM directional categorization, which however does not involve relationships of separate events.

Two monkeys were trained on a frequency contour categorization of tone sequences which covered the following procedure: Upon a light signal the animals could grasp a bar and thereby actively elicit a tone sequence after a constant delay. The sequence always started with three identical tones followed by three identical tones of lower frequency (down sequence) or first followed by 3–5 tones of higher frequency and then by 3 tones of lower frequency (up-down sequence; Fig. 5). The release of the bar upon identification of a down step was immediately rewarded. Initial tones varied over a wide range of frequencies and tone steps of one or one half octave varied randomly across trials. Both monkeys took more than two years, including a number of training steps, to master the task (Brosch et al., 2004).

Recordings of single units and clusters of 3–5 units (in the following summarized as units) were made with multiple microelectrodes in AI and the caudomedial field CM.

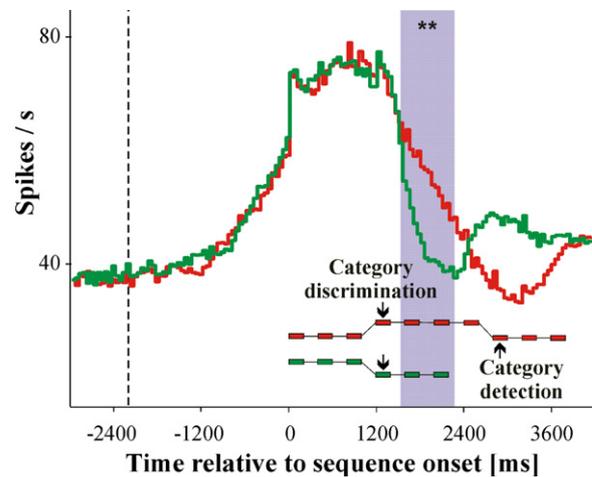


Fig. 5. Average time courses of 40 tonically firing multiunits in the caudomedial auditory cortex of primates in the two tone sequence conditions (inset with same time scale): Sequences with a downward frequency step only are marked in green. Sequences with an upward frequency step followed by a downward frequency step are marked in red. Dashed vertical line indicates bar grasping. In the period after the 4th tone indicated in blue (\*\*) the decrease of firing represented by the red and the green curves was significantly different (Wilcoxon test,  $p < 0.01$ ). The motor reactions (bar release) occurred approximately at the minimum of the curves, i.e. earlier in down sequences (green). Thus, the slopes of the red and green curves are predictive of the subsequent behavioural responses. (Modified from Selezneva et al., 2006.)

The recordings not only identified units that responded in a specific way to tone steps (Selezneva et al., 2006), but also units that responded to various procedural contingencies of the training paradigm and units that reflected the cognitive contingencies of the task (Brosch et al., 2005).

Remarkably, in both AI and CM many units were found that responded with short latency to the light signal, or fired concomitantly with bar grasping, or with the release of the bar, or to the consumption of the liquid reward. Many units responded to one or several of such procedural contingencies and to the tones as well. Thus, at least the latter units were auditory units but were obviously under non-auditory influence. The task-specificity of such unit responses to procedural contingencies became evident by a second task that was trained in order to scrutinize the specificity. The monkeys easily learned to switch to a task with the same procedural contingencies but in which the bar grasping elicited flickering of the light signal instead of the tone series. Bar release upon detection of this flicker was rewarded. When units during the auditory task were encountered that reliably responded to the bar grasping or to bar release, the task was switched to the visual modality. In none of these cases were the responses of the same units maintained. Units in AC responding to the procedural contingencies of the auditory task did not respond to the same contingencies during the visual task. This suggests that the underlying multimodal connectivities of AC were functional only in the context of the auditory task. The significance of these procedure-related activations might therefore be a facilitation of auditory task execution,

namely in terms of prediction of events as will be explained in more detail.

How tone-responses of units would distinguish the frequency steps during the auditory task was the central question of the neuronal analysis (Selezneva et al., 2006). Most units in AI and CM showed phasic responses to all tones of the sequence within their frequency receptive fields. On average the spike rate of all responsive units increased to stepped tones and was stronger to the tones representing a down step than to tones representing up steps or flat steps. This highly significant effect was seen independent of best frequency (BF) of neurons, which suggests that the ensemble of phasic units preferred the rewarded down steps of tones. Nevertheless, this average difference of response increase to stepped tones, even though being prominent, raises a fundamental problem: How can a monkey in an ongoing trial decide whether the first change in the tone sequence is an up step or a down step? There is no direct comparison of these two tone responses possible and a monkey could use only an internal reference, *i.e.* some kind of a threshold to distinguish smaller up step-related increases of discharge from larger down-step related increases of discharge.

This consideration made us search for a neuronal code that could signal the direction of change without an internal reference. In the same units we compared in each trial the response to a stepped tone with the response to the immediately preceding tone and then we analyzed the discharge differences across trials. This analysis produced a binary-type response difference (Selezneva et al., 2006). Up steps generated no increase of discharge rate over that of the preceding tone while the rewarded down steps significantly increased the discharge rate. A reliable response bias of auditory cortical neurons to one direction of a frequency change has not been observed in AC of naïve monkeys (Brosch et al., 1999; Bartlett and Wang, 2005). Interestingly, this effect was independent of whether the monkeys correctly released the bar upon a downward step or made a mistake. Thus, phasic auditory units were on average better in identifying tone steps than the monkey's performance. This result motivated a search in the obtained unit data for additional information that could explain this discrepancy.

In the search for neuronal mechanisms that could be responsible for the monkey's behavioural mistakes in spite of sufficient neuronal information about the direction of frequency steps we found multiple sites in the field CM where units showed relevant slow modulation of tonic firing (Fig. 5; see also Selezneva et al., 2006). These units started to increase firing rate as soon as the monkey grasped the bar and reached a plateau of a high firing rate at the occurrence of the first tone. Immediately after the 4th tone the firing rate steeply decreased and reached a minimum with bar release, *i.e.* earlier in down sequences and later in up-down sequences. This time course of firing reflected the two mistakes observed in error trials. In "false alarm" trials of up-down sequences, in which the

monkey erroneously released the bar after the rising frequency step, the decrease of firing after the 4th tone was steep and indistinguishable from correct "hit" trials of down-sequences. Conversely, in "miss" trials of down-sequences, the tonic firing decreased more slowly and similarly as in the "correct rejection" trials of up-down sequences. Thus, this type of firing did not reflect the categorical stimulus properties of rising and falling steps like phasic responses but rather the correct and false interpretation of these stimuli leading to the corresponding behaviour of the monkey.

Even though we do not yet know whether this slow modulation of firing by units is a read-out of the information, provided by the phasic responses, it is anticipatory of the monkey's behaviour. The steepness of decrease of firing after the 4th tone was significantly different for up and up-down sequences and predicted whether the monkey would release the bar earlier or later (Fig. 5). This poses the intriguing problem why the 4th tone was the single most important event for this type of units. A solution is obtained by reconsidering the structure of the tone sequences which inadvertently contains information that the monkey could learn to use for cognitive decisions.

The monkeys could learn that the first tone step always occurred with the 4th tone and that the correct categorization of this tone predicted all subsequent events. If this tone did not represent a down step, a second step would occur later that always was a down step, *i.e.* did not require categorization but only detection of a change. This key role of the 4th tone in the paradigm might explain why the tonic activity of units was focussed on this and no later events. In this sense, this type of activity might represent a cognitive state of the AC. The increase of activity until the 4th tone might represent the time window for the categorical decision on that tone in comparison to the preceding tones and subsequently the decrease of activity with variable time course might represent the result of this decision translatable into behaviour (Selezneva et al., 2006).

In some sense these tonic dynamics of activity are reminiscent of the marked states obtained in gerbil AC for categorization of FM direction. Both types of dynamics do not depend on the stimuli *per se* but on the cognitive and behavioural conclusions that are drawn from these stimuli. Both types of dynamics predict the subsequent behaviour.

There can be no question that the described response types in monkey AC after the training of frequency contour categorization are highly task-specific. All three phenomena described here reflect special aspects of the task: (1) The response bias of phasic units to the rewarded direction of frequency steps, (2) the non-auditory responses reflecting the procedural framework in which the auditory task was embedded, and (3) the slow modulation of activity reflecting the cognitive structure of the task.

## 6. Conclusion

The experimental results presented here from gerbils, monkeys, and humans converge on the conclusion that stimulus representations in AC, as reflected by several correlates of neuronal activity, are modified according to the structure of auditory tasks and especially of trained tasks. This may be explained by the extensive corticocortical connectivity of AC and of functional areas of cortex in general which provide flexible options for task-related influences. Changes of representation not only involve modifications of the direct response to an auditory stimulus itself like the increased response to the occurrence of a rewarded frequency change in monkeys. Also, additional responses to stimuli of other modalities and to self-generated events can be elicited in AC if they determine the behavioural context or the temporal contingencies of the auditory task. Furthermore, some slow changes of field potentials, unit firing, and of neuronal activity, presumably underlying the fMRI signals, seem to represent special cognitive states of AC and predict successful categorization of stimuli (gerbil), mistaken categorization (monkey), or the performance level of categorization (human subjects). The topographic and dynamic spatial aspects of the changes of stimulus representations may vary from AC including AI (gerbil, monkey) to exclusively covering *planum temporale* areas (human subjects), but differentially in the right and left hemisphere.

Not all aspects of these changes of representation are new scientific results (e.g., Fritz et al., 2005), but the convergence of all phenomena in a concept of task-specific changes seems to be a novel proposition that overcomes previous conflicts of results. For instance, shifts of BF of neurons towards another frequency within their receptive fields, that has become significant by simple associations or discrimination learning, was proposed as a general mechanism in AI how representations of stimulus frequencies become distinct (for review see Weinberger, 2004 but also see Irvine and Wright, 2005). This, however, is neither compatible with the result of other frequency discrimination tasks (Ohl and Scheich, 1996, 1997, 2004, 2005), nor with the results of the described frequency categorization task in the monkey (Selezneva et al., 2006). In the latter case, neurons increased their firing exclusively to the rewarded falling frequency steps independent of their BF. If neurons generally shift their BF towards a new behaviourally salient frequency, then neurons, whose BF were originally above a falling frequency step, would after training indeed respond with an increase of discharge to the falling frequency step. Thus, they would respond to their new BF. But also neurons, whose BF were originally below a rising frequency step, would respond with an increased discharge to that step due to the shift of their BF. But increased responses to rising steps did not occur in this task.

There is increasing evidence that the AC is under multisensory influence which has led to the suggestion that in natural situation all sensory processing in neocortex might

be multisensory (for review see, e.g., Ghazanfar and Schroeder, 2006). Most of the multisensory influences on the AC described here cover modulation of auditory responses by other modalities. The underlying assumption was that auditory percepts are modified by this non-auditory modulation of the responses which may shape, e.g., auditory discrimination. But the non-auditory influences may be more powerful. In one of the earliest reports of visual impact on AC it was shown that association learning of auditory and visual stimuli after training led to responses of AC to the visual stimuli alone (Cahill et al., 1996). This indicates that an auditory memory might be evoked by an associated other stimulus perhaps leading to auditory “imagination”. Human studies with mere imagination of auditory stimuli have shown activation of AC (Halpern and Zatorre, 1999). Yet another domain of non-auditory activation of AC is the firing of neurons in the above described monkey experiments. Presumably, this activity is procedurally related to the execution of the sequential task and to the cognitive task aspects involved. It is difficult to assume that the perceptual aspects of tones are changed by this activity which is clearly separate from the tone responses. Conversely, this type of activity suggests that not all modes of changed firing in AC may lead to changed auditory percepts. Taken together, these various types of non-auditory influences on auditory cortical activity seem to be covered by a concept of task-specificity, which shapes the auditory aspects of a task and ties it together with the procedural and cognitive aspects.

Finally, there is a large body of literature showing that AC functions in humans are lateralized not only for speech processing but also for other types of auditory patterns (Zatorre et al., 2002; Hutsler and Galuske, 2003; for review). A fruitful concept unifying to some extent the speech-related left AC specializations with non-speech specializations is a higher temporal resolution of auditory signals in left AC than in right AC. In contrast, a higher spectral resolution is assumed for right AC which could underlie dominant right AC involvement in some aspects of music processing (Zatorre et al., 2002) and of speech prosody processing (Brechmann and Scheich, 2005). While this division of labour in temporal *vs.* spectral stimulus analysis seems to be compatible with the available results in human brain imaging and with the neuropsychology of unilateral brain lesions, the notion (Zatorre et al., 2002) that the lateralization derives from stimulus properties that, in order to be determined, require superior resolution in the temporal or spectral domain appears to be too limited, even though it gains some appeal from general signal analysis theory. Any auditory signal has spectral and temporal properties that could be analysed to an unlimited degree and consequently both AC hemispheres would always be involved in the analysis. But whether a high temporal resolution is an issue will be determined in specific situations, e.g., by a given discrimination task where two stimuli show minor differences in temporal properties even though the recognition of each stimulus is not primarily

temporally demanding. The same reasoning applies to the spectral resolution. Furthermore, the results described here on processing of linear FM, both in gerbil and human AC, may not be ideal examples of refined spectral resolution in right AC. The specialization for the directional discrimination and categorization of sound in right AC does not require high spectral resolution and rather reflects spectrotemporal analysis of a change of frequency. The analysis of these FM stimuli themselves is not the domain of right AC, rather the identification of the direction of their frequency change, as shown by the WM task using the same type of stimuli. There, the fMRI activation became dominant in the left AC, as shown by the task-specific correlation of activity with performance. While these examples weaken the conviction that functional hemispheric specializations of AC are related to pure sound analysis they emphasize the notion that the hemispheres have different auditory task-specificities.

## References

- Bannister, A.P., 2005. Inter- and intra-laminar connections of pyramidal cells in the neocortex. *Neurosci. Res.* 53, 95–103.
- Bartlett, E.L., Wang, X., 2005. Long-lasting modulation by stimulus context in primate auditory cortex. *J. Neurophysiol.* 94, 83–104.
- Baumgart, F., Gaschler-Markefski, B., Woldorff, M.G., Heinze, H.J., Scheich, H., 1999. A movement-sensitive area in auditory cortex. *Nature* 400, 724–726.
- Brass, M., von Cramon, D.Y., 2004. Selection for cognitive control: A functional magnetic resonance imaging study on the selection of task-relevant information. *J. Neurosci.* 24, 8847–8852.
- Brechmann, A., Scheich, H., 2005. Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cereb. Cortex* 15, 578–587.
- Brechmann, A., Baumgart, F., Scheich, H., 2002. Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study. *J. Neurophysiol.* 87, 423–433.
- Brechmann, A., Gaschler-Markefski, B., Sohr, M., Yoneda, K., Kaulisch, T., Scheich, H., 2007. Working memory specific activity in auditory cortex: Potential correlates of sequential processing and maintenance. *Cereb. Cortex*, doi:10.1093/cercor/bhl160.
- Brosch, M., Schulz, A., Scheich, H., 1999. Processing of sound sequences in macaque auditory cortex: response enhancement. *J. Neurophysiol.* 82, 1542–1559.
- Brosch, M., Selezneva, E., Bucks, C., Scheich, H., 2004. Macaque monkeys discriminate pitch relationships. *Cognition* 91, 259–272.
- Brosch, M., Selezneva, E., Scheich, H., 2005. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.* 25, 6797–6806.
- Budinger, E., Heil, P., 2005. Anatomy of the auditory cortex. In: Greenberg, S., Ainsworth, W. (Eds.), *Listening to Speech*. Lawrence Erlbaum, Mahwah, New Jersey, pp. 91–113.
- Budinger, E., Heil, P., Scheich, H., 2000a. Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). III. Anatomical subdivisions and corticocortical connections. *Eur. J. Neurosci.* 12, 2425–2451.
- Budinger, E., Heil, P., Scheich, H., 2000b. Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). IV. Connections with anatomically characterized subcortical structures. *Eur. J. Neurosci.* 12, 2452–2474.
- Budinger, E., Heil, P., Hess, A., Scheich, H., 2006. Multisensory processing via early cortical stages – connections of the primary auditory field with other sensory systems. *Neuroscience* 143 (4), 1065–1083.
- Buxhoeveden, D.P., Casanova, M.F., 2002. The minicolumn hypothesis in neuroscience. *Brain* 125, 935–951.
- Cahill, L., Ohl, F., Scheich, H., 1996. Alteration of auditory cortex activity with a visual stimulus through conditioning: a 2-deoxyglucose analysis. *Neurobiol. Learn. Memory* 65, 213–222.
- Constantinidis, C., Procyk, E., 2004. The primate working memory networks. *Cogn. Affect. Behav. Neurosci.* 4, 444–465.
- Creutzfeldt, O.D., 1983. *Cortex Cerebri*. Springer, Berlin, Heidelberg.
- Creutzfeldt, O.D., Garey, L.J., Kuroda, R., Wolff, J.-R., 1977. The distribution of degenerating axons after small lesions in the intact and isolated visual cortex of the cat. *Exp. Brain Res.* 27, 419–440.
- Deike, S., Gaschler-Markefski, B., Brechmann, A., Scheich, H., 2004. Auditory stream segregation relying on timbre involves left auditory cortex. *Neuroreport* 15, 1511–1514.
- Douglas, R.J., Martin, K.A.C., 1994. The canonical microcircuit: A co-operative neuronal network for neocortex. In: Allowitz, B. et al. (Eds.), *Structural and functional Organization of the Neocortex*. In: *Exp. Brain Res. Ser.*, Vol. 24. Springer, Berlin, Heidelberg, pp. 131–141.
- Dowling, W.J., 1978. Scale and contour: two components of a theory of memory for melodies. *Psychol. Rev.* 85, 341–354.
- Eggermont, J.J., 1998. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. *J. Neurophysiol.* 80, 2743–2764.
- Felleman, D.J., van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Fritz, J., Elhilali, M., Shamma, S., 2005. Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hear. Res.* 206 (1–2), 159–176.
- Fuster, J.M., Alexander, G.E., 1971. Neuron activity related to short-term memory. *Science* 173, 652–654.
- Gaschler-Markefski, B., Baumgart, F., Tempelmann, C., Woldorff, M.G., Scheich, H., 1998. Activation of human auditory cortex in retrieval experiments: an fMRI study. *Neural Plast.* 6, 69–75.
- Gatter, K.C., Powell, T.P.S., 1978. The intrinsic connections of cortex of area 4 of the monkey. *Brain* 101, 513–541.
- Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285.
- Goldschmidt, J., Zuschratter, W., Scheich, H., 2004. High-resolution mapping of neuronal activity by thallium autometallography. *Neuroimage* 23, 638–647.
- Gottlieb, Y., Vaadia, E., Abeles, M., 1989. Single unit activity in the auditory cortex of a monkey performing a short term memory task. *Exp. Brain Res.* 7, 139–148.
- Halpern, A.R., Zatorre, R.J., 1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9, 697–704.
- Hubel, D.H., Wiesel, T.N., 1977. Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* 198 (1130), 1–59.
- Hutsler, J., Galuske, R.A., 2003. Hemispheric asymmetries in cerebral cortical networks. *Trends Neurosci.* 26, 429–435.
- Irvine, D.R.F., Wright, B.A., 2005. Plasticity of spectral processing. *Int. Rev. Neurobiol.* 70, 435–472.
- Jerison, H.J., 1973. *Evolution of the brain and Intelligence*. Academic, New York, London.
- König, R., Heil, P., Budinger, E., Scheich, H. (Eds.), 2005. *The auditory cortex – a synthesis of human and animal research*. Lawrence Erlbaum, Mahwah, New Jersey.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.
- Martin, K.A., 2002. Microcircuits in visual cortex. *Curr. Opin. Neurobiol.* 12, 418–425.
- Metherate, R., Kaur, S., Kawai, H., Lazar, R., Liang, K., Rose, H.J., 2005. Spectral integration in auditory cortex: mechanisms and modulation. *Hear. Res.* 206, 146–158.

- Miller, L.M., Escabi, M.A., Read, H.L., Schreiner, C.E., 2001. Functional convergence of response properties in the auditory thalamocortical system. *Neuron* 32, 151–160.
- Mountcastle, V.B., 1997. The columnar organization of the neocortex. *Brain* 120, 701–722.
- Ohl, F.W., Scheich, H., 1996. Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field A1) of the alert Mongolian gerbil. *Eur. J. Neurosci.* 8, 1001–1017.
- Ohl, F.W., Scheich, H., 1997. Learning-induced dynamic receptive field changes in primary auditory cortex of the unanaesthetized Mongolian gerbil. *J. Comp. Physiol. [A]* 181, 685–696.
- Ohl, F.W., Scheich, H., 2004. Fallacies in behavioural interpretation of auditory cortex plasticity. *Nature Rev. Neurosci.* online (doi: 10.1038/nrn1366-c1).
- Ohl, F.W., Scheich, H., 2005. Learning-induced plasticity in animal and human auditory cortex. *Curr. Opin. Neurobiol.* 15 (4), 470–477.
- Ohl, F.W., Wetzel, W., Wagner, T., Rech, A., Scheich, H., 1999. Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learn Memory* 6, 347–362.
- Ohl, F.W., Scheich, H., Freeman, W.J., 2001. Change in pattern of ongoing cortical activity with auditory category learning. *Nature* 412, 733–736.
- Passingham, D., Sakai, K., 2004. The prefrontal cortex and working memory: physiology and brain imaging. *Curr. Opin. Neurobiol.* 14, 163–168.
- Pasternak, T., Greenlee, M.W., 2005. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6 (2), 97–107.
- Rakic, P., 2002. Evolving concepts of cortical radial and areal specification. *Prog. Brain Res.* 136, 265–280.
- Ranganath, C., D'Esposito, M., 2005. Directing the mind's eye: prefrontal, inferior and medial temporal mechanisms for visual working memory. *Curr. Opin. Neurobiol.* 15, 175–182.
- Rockel, A.J., Hiorns, R.W., Powell, T.P., 1980. The basic uniformity in structure of the neocortex. *Brain* 103 (2), 221–244.
- Rouiller, E.M., Simm, G.M., Villa, A.E.P., de Ribaupierre, Y., de Ribaupierre, F., 1991. Auditory corticocortical interconnections in the cat: evidence for parallel and hierarchical arrangement of the auditory cortical areas. *Exp. Brain Res.* 86, 483–503.
- Sakurai, Y., 1994. Involvement of auditory cortical and hippocampal neurons in auditory working memory in the rat. *J. Neurosci.* 14, 2606–2623.
- Scheich, H., Heil, P., Langner, G., 1993. Functional organization of auditory cortex in the Mongolian gerbil (*Meriones unguiculatus*). II. Tonotopic 2-Deoxyglucose. *Eur. J. Neurosci.* 5, 898–914.
- Scheich, H., Simonis, C., Ohl, F.W., Thomas, H., Tillein, J., 1994. Mapping of stimulus features and meaning in gerbil auditory cortex. In: Allowitz, B. et al. (Eds.), *Structural and functional Organization of the Neocortex*. In: *Exp. Brain Res. Ser.*, Vol. 24. Springer, Berlin, Heidelberg, pp. 252–267.
- Scheich, H., Baumgart, F., Gaschler-Markefski, B., Tegeler, C., Tempelmann, C., Heinze, H.J., Schindler, F., Stiller, D., 1998. Functional magnetic resonance imaging of a human auditory cortex area involved in foreground-background decomposition. *Eur. J. Neurosci.* 10, 803–809.
- Scheich, H., Ohl, F.W., Schulze, H., Hess, A., Brechmann, A., 2005. What is reflected in auditory cortex activity: Properties of sound stimuli or what the brain does with them? In: König, R., Heil, P., Budinger, E., Scheich, H. (Eds.), *The Auditory Cortex – A Synthesis of Human and Animal Research*. Lawrence Erlbaum, Mahwah, New Jersey, pp. 383–408.
- Seleznova, E., Scheich, H., Brosch, M., 2006. Dual time scales for categorical decision making in auditory cortex. *Curr. Biol.* 16 (24), 2428–2433.
- Silberberg, G., Gupta, A., Markram, H., 2002. Stereotypy in neocortical microcircuits. *Trends Neurosci.* 25 (5), 227–230.
- Stark, H., Rothe, T., Wagner, T., Scheich, H., 2004. Learning a new behavioral strategy in the shuttle-box increases prefrontal dopamine. *Neuroscience* 126 (1), 21–29.
- Suga, N., Ma, X., 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci.* 4, 783–794.
- Swadlow, H.A., Gusev, A.G., Bezdudnaya, T., 2002. Activation of a cortical column by a thalamocortical impulse. *J. Neurosci.* 22, 7766–7773.
- Swets, J.A., Tanner, W.P., Birdsall, P.G., 1961. Decision processes in perception. *Psychol. Rev.* 68, 301–340.
- Szentagothai, J., 1978. The Ferrier Lecture, 1977. The neuron network of the cerebral cortex: a functional interpretation. *Proc. R. Soc. Lond. B Biol. Sci.* 201 (1144), 219–248.
- Weinberger, N.M., 2004. Specific long-term memory traces in primary auditory cortex. *Nat. Rev. Neurosci.* 5, 279–290.
- Wetzel, W., Ohl, F.W., Wagner, T., Scheich, H., 1998. Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neurosci. Lett.* 252, 115–118.
- Winer, J.A., 1992. The functional architecture of the medial geniculate body and the primary auditory cortex. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy*. Springer, New York, pp. 222–409.
- Winer, J.A., Lee, C.C., 2007. The distributed auditory cortex. *Hear. Res.* 229, 3–13.
- Winer, J.A., Miller, L.M., Lee, C.C., Schreiner, C.E., 2005. Auditory thalamocortical transformation: structure and function. *Trends Neurosci.* 28, 255–263.
- Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46.