Insperata 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: D5.1
Setup of Experimental Paradigm for Studying Audiovisual Fusion in Rodent Cortex

Date of deliverable: 31-12-2006
Actual submission date: 11.02.2007

Start date of project: 01.01.2006 Duration: 60 months

Organization name of lead contractor for this deliverable: **Leibniz 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)	

D5.1 SETUP OF EXPERIMENTAL PARADIGM FOR STUDYING AUDIOVISUAL FUSION IN RODENT CORTEX

Leibniz Institute for Neurobiology (LIN)

Abstract:
Two experimental paradigms, one employing rodents, the other employing human subjects have been developed for the study of neuronal mechanisms underlying two forms audiovisual information fusion. In the rodent experiment we were able to demonstrate the category transfer from information learned in the auditory modality to the visual modality and identify first potential physiological correlates of that transfer. In the experiment employing human subjects we were able to demonstrate the influence of visual inducer stimuli on the performance auditory temporal order judgements. Both experimental approaches together provide a framework in which various psychophysical phenomena of audiovisual integration can be treated in a coherent fashion and rooted in a physiological basis.

Table of Content

1. Introduction

The phenomenon of ("high-level") category transfer from one modality to another plays a fundamental role in the theoretical underpinning of DIRAC's approach of the multi-sensory fusion problem, as we envisage crossmodal category transfer as one means by which a cognitive system can meaningfully deal with rare and unexpected events in one modality. Putting experience of a rare, unexpected event in relation to informed cognitive structures established by experience with other modalities might serve as a role model for top-down processing of rare events, for which no immediate models exist in the framework of the original feature space (modality).

The task was therefore to first create an experimental paradigm that allows analysis of cortical mechanisms of audiovisual interaction. In addition to the initially planned experiment studying "high-level" transfer of information learned in the auditory modality to the visual modality to be conducted in rodents (gerbils), we have also developed an experiment to study audiovisual interaction in humans. The experiment on gerbils exploits the precise spatiotemporal resolution of intracerebral recording of electrocortiograms and focuses on integration dynamics on timescales of tens of seconds and longer. The experiment on humans (limited to scalp recordings of EEGs) takes advantage of the easier access to phenomenological aspects of perception given by the possibility of report by subjects and focuses on integration dynamics on timescales of tens of seconds and shorter.

It was the main task of the first year to build the experimental setups and hardware, develop necessary software and demonstrate feasibility of experiments by conducting informative pilot experiments.

2. Transfer of Categorization Learned in the Auditory Modality to the Visual Modality

2.1 Experimental Strategy

The task was to develop an experimental paradigm in which gerbils learn a temporal categorization task with auditory stimuli and then are required to transfer the learned contingencies from the auditory modality to the visual modality, while brain activity is recorded in parallel. The paradigm should be modelled according to a previous experiment which elucidated fundamental neuronal mechanisms of category formation (Ohl et al. 2001). Brain activity should then be inspected for signs of directed information flow between auditory and visual cortices.

2.2 Granger Causality Analysis

To tackle the latter problem we have used Granger causality analysis which provides pairwise directional information of causal influences between time series.

The basic idea behind Granger causality analysis is the insight that causes always precede their consequences and consequences might be related to their causes in the past. If one considers the bivariate autoregressive process

$$X_1(t) = \sum_{j=1}^{p} A_{11}(j) X_1(t-j) + \sum_{j=1}^{p} A_{12}(j) X_2(t-j) + E_1(t)$$

$$X_2(t) = \sum_{j=1}^{p} A_{21}(j) X_1(t-j) + \sum_{j=1}^{p} A_{22}(j) X_2(t-j) + E_2(t)$$

in which $X_1(t)$ and $X_2(t)$ denote the time series from two data channels the question about which one of the time series might be causative for the other can be given the following operational definition: If the variance of the prediction error E_1 is reduced by the inclusion of the X_2 terms in the first equation, then X_2 is said to "Granger-cause" X_1. In other words, Granger causality would be indicated if the off-diagonal coefficients of the coefficient matrix $A_{mn}(j)$, j=1, …, p are not uniformly zero.

In the general case we assume that the k-dimensional (vector) process X(t) can be described by MVAR process of order p

$$X(t) = \sum_{i=1}^{p} A(i) X(t-i) + E(t).$$

This equation can be transformed to the frequency domain

$$A(f) X(f) = E(f)$$

where

$$A(f) = -\sum_{j=0}^{p} A(j) \exp(-i\, 2\pi\, j)$$

and A(0) is the negative identity matrix,

$$A(0) = -I\ .$$

Foramally, the inverse of the Fourier-transformed coefficient matrix can be considered the transfer matrix H(f) of a linear system, since

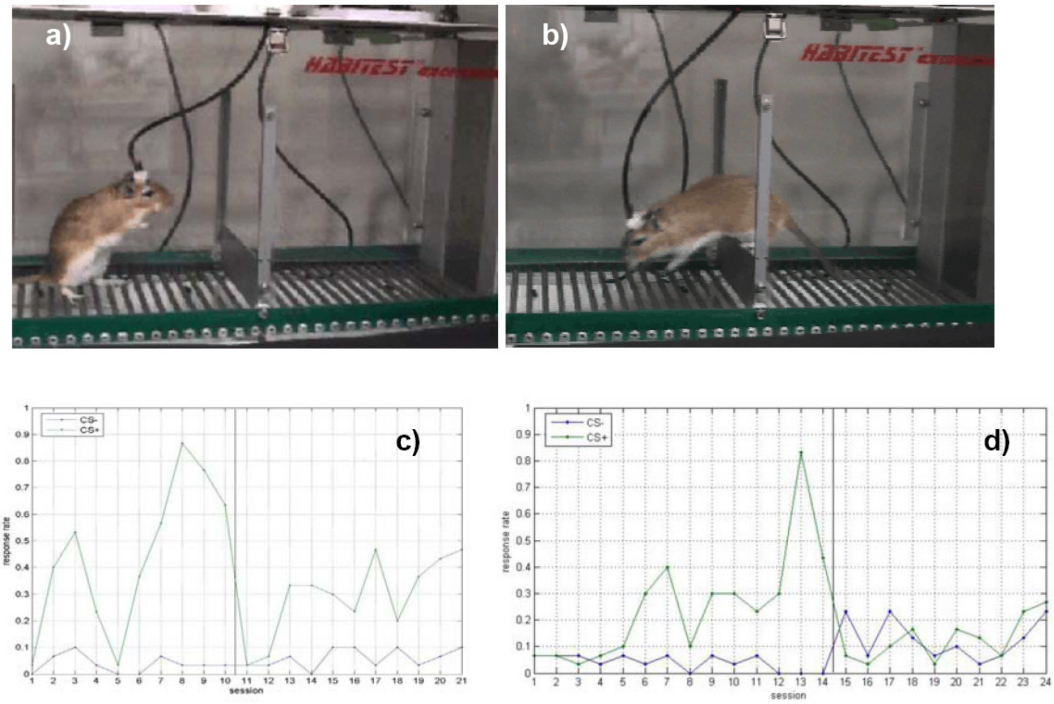

Figure 1. Experimental preparation and sample data of the experiment on audio-visual information transfer of rhythm categories. Panels a) and b) show a gerbil prepared for multichannel recording of unit and local field potential activity from auditory and visual cortex performing a NO-GO and GO situation, respectively, in a shuttle box paradigm. Panel c) shows a significant transfer of conditioned GO response rate (ordinate) over time (abscissa) from a phase of auditory training (left of vertical line) to a phase in which only visual stimuli (LEDs above the shuttle box, not shown) are presented. Green and blue curves show conditioned GO response rates for target stimuli (hit rate) and non-target stimuli (false alarms), respectively. Panel d) illustrates a control situation in which the behavioural meaning of the categories "slow" and "fast" are reversed during transition from the auditory to the visual training phase. In this case no significant transfer develops over time as indicated by the nonsignificant difference between the blue and green curve right of the vertical line.

$$X(f) = A^{-1}(f)E(f) = H(f)E(f).$$

The Granger-causal influence of channel j to channel i, can then be represented by the directed transfer function (DTF) from channel j to channel i, defined as

$$\Theta_{ij}^2(f) = \left|H_{ij}(f)\right|^2$$

If the DTF from channel j to channel i is normalized by the joint influences from all other channels to channel i, the normalized DTF (Kaminski and Blinowska 1991) is obtained

$$\gamma_{ij}^2(f) = \frac{\left|H_{ij}(f)\right|^2}{\sum_{m=1}^{k}\left|H_{im}(f)\right|^2}$$

which specifies the normalized Granger-causal influence.

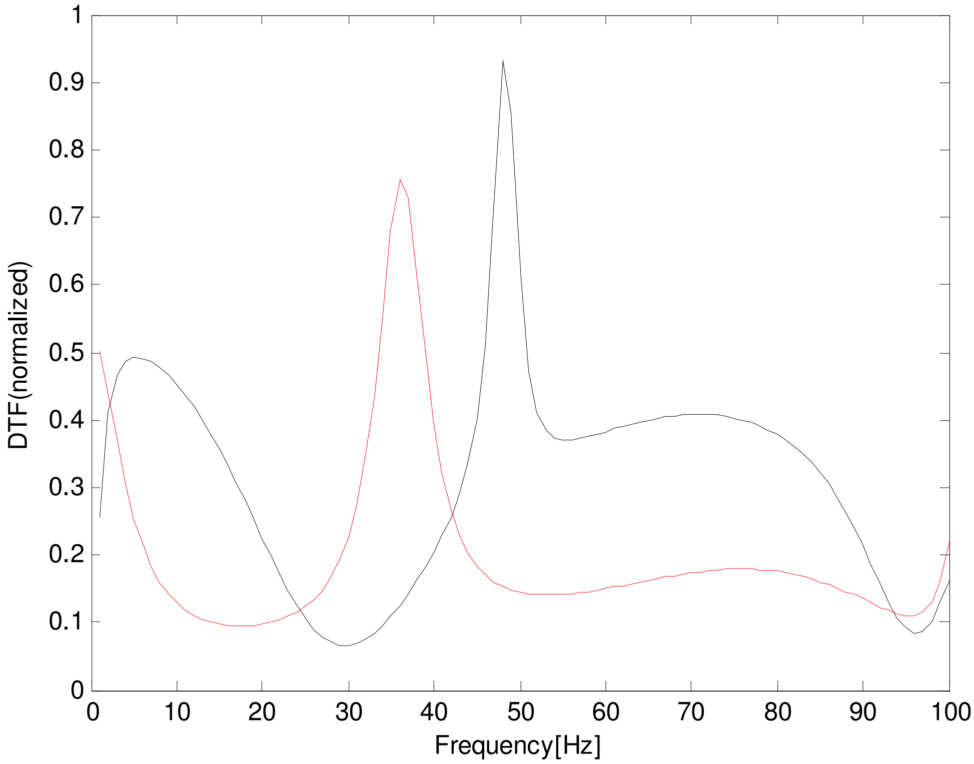

Figure 2. Normalized directed transfer functions for a pair of recording channels, one placed in primary auditory cortex, field A1, the other in primary visual cortex, field V1. (red) DTF: A1->V1, (black) V1->A1.

2.3 Results

It was demonstrated that gerbils (rodents) can be trained to categorize auditory stimuli into two classes, viz. "slow" rhythms and "fast" rhythms. The exact physical category boundaries can be largely shifted in signal space as was expected by previous work (Ohl et al. 2001; Ohl and Scheich 2005). After training, gerbils were able to transfer this concept of temporal structure to visual stimuli (blinking LEDs), as indicated by the results of classical psychophysical transfer experiments with congruent and incongruent transfer conditions (see Fig. 1).

Furthermore, we were able to combine the psychophysical experiment with the parallel intracerebral recording of electrocorticograms and apply Granger causality analysis to track physiological evidence of "high-level" cross-modal transfer of learned concepts. Figure 2 shows an example of asymmetric directed transfer functions based on recordings from two electrodes, one implanted in auditory cortex and one implanted in visual cortex. Significant Granger-causality was found for both directions auditory->visual and visual->auditory. Most interestingly, these functions showed power maxima at different frequencies. While the auditory->visual causality peaked in the high beta range (>35 Hz), the visual->auditory causality typically peaked in the gamma range (50-60 Hz).

While these experiments clearly demonstrate the feasibility of the proposed experimental approach more cases are needed to be able to draw statistically valid conclusions about the cortico-cortical interaction and its role in mediating cross-modal category transfer.

3. Temporal Order Judgements

3.1 Introduction and Background

Audiovisual information integration is a key feature of human communication by speech. Uncovering its neuronal mechanisms could not only improve technical communication systems (from recreational aspects like foreign language movie synchronization to clinically relevant aspects like the design of hearing aids) (cf. Zampini et al., 2003) but is expected to allow fundamental insights in how the brain combines information provided by separable input channels generally. Previous studies have shown the importance of spatial coincidence and temporal synchrony for binding and modulation of crossmodal perception, also showing that time and space are difficult to investigate completely independent of each other (for reviews see Driver & Spence, 1998; Shimojo & Shams, 2001; Bertelson & de Gelder, 2003; Grondin 2003). In workpackage 5 of DIRAC we will focus on temporal aspects.

Ideas about the nature of the temporal organisation of perception in the psychological context go back to Fechner (1860) who stated that the conscious processing is based on oscillations. In the year 1864 von Baer coined the term "moment" as the minimum duration of a sensory experience. Later research has indicated that this duration could also be interpreted as a time window for integrative processes (Fig 3). An empirical foundation of the existence of "moments" was provided by Lalanne (1876) with his observation of common fusion frequencies (of about 18 Hz, 55 Hz) for acoustic, tactile and visual temporally modulated stimuli. Later studies implicated greater variability of moment durations, typically in the range of 20 to 40 ms (e.g. Hirsh and Sherrick 1961; Pöppel 1997). Using greater care, recent studies suggest that there is no invariant moment. Zampini et al. (2003) were able to show that the just noticeable difference (JND) in audiovisual stimulation was increased from 22 ms to 62 ms when the response dimension for the subjects was changed from indicating which modality came first (similar to the study by Hirsch und Sherrick, 1961) to spatial judgements, where subjects indicated whether the "left" or "right" stimulus came first. In addition, these authors could demonstrate a spatial confound in most prior experiments aimed at measuring temporal order judgements (TOJs), by showing that JNDs are overestimated when the auditory and visual stimulus come from different locations (as it is the case in many experiments involving an LED in the center of the visual field as the visual stimulus and auditory stimuli delivered by headphones) as compared to a situation when they are presented from the same spatial location. Measurement of TOJs is further complicated by known target-distractor interactions (Shimojo and Shams 2001; Morein-Zamir et al. 2003) and the possibility for recalibration of the TOJs by context and experience (e.g. Spence and Squire 2003; Fujisaki et al 2004; Vroomen et al. 2004; Navarra et al. 2005).

Based on Fechner's original proposal, several investigators with the advent of EEG measurement techniques have sought to establish a relationship between psychophysical JND and TOJ parameters on the one side and certain EEG oscillations on the other (e.g. alpha: Stroud 1995; beta-gamma: Pöppel 1997). Gho and Varela (1988) could show that the perceived temporal order of two asynchronous light flashes does indeed depend on the relative phase of stimulus presentation to the ongoing cortical alpha activity.

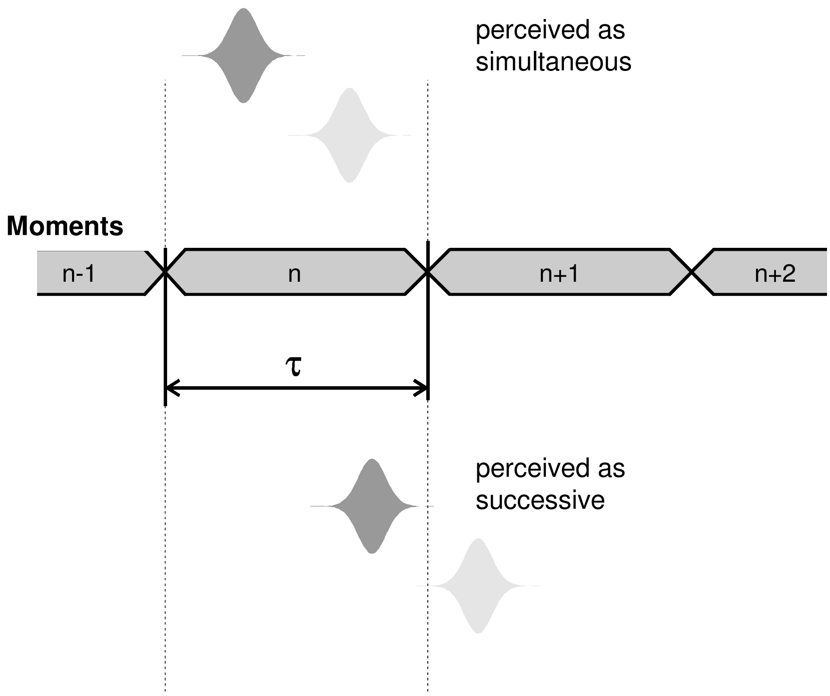

Figure 3. Schematic depiction of the moment model by von Baer (1864) or Stroud (1955) in relation to a TOJ task. Temporal structure of perception is conceived as a sequence of "moments" of fixed duration τ. If two stimuli, close in time, fall within the same "moment" they cannot be perceived as asynchronous.

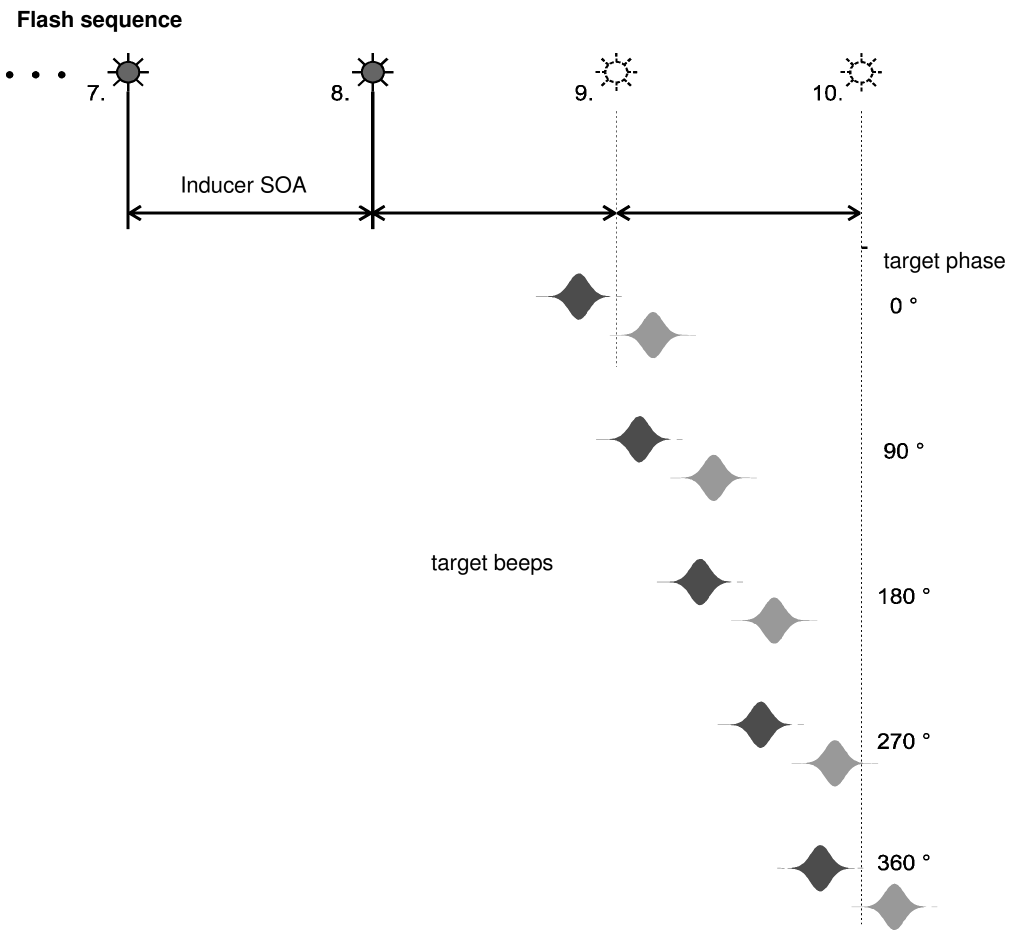

Figure 4. Design of the experiment. Depicted are schemata of the pure tone pips at various relative phases to the inducer flashes, of which the last four are shown.

3.2 Experiment

Based on these studies we have first started to conduct a psychophysical study of the performance of subjects in an intramodal (auditory) TOJ task influenced by an induced rhythm in another modality (vision). We presented two different tones (700 and 1700 Hz with Gaussian shaped envelopes to avoid spectral splatter) monaurally over headphones (60 dB SPL) as a target and assumed to induce cortical oscillations by presenting prior to target onset a sequence of whole visual field flashes at 10 Hz (Fig. 4). Tones were presented beginning with 9^{th} flash and relative phases of the tone presentation times were related to this instant as indicated in the figure. Relative phases were changed in steps of 25 ms yielding phase shifts of 0^o, 90^o, 180^o and 270^o and 360^o. Stimulus onset asynchronies between the high and low pitch tones were -40, -24, -8, 0, 8, 24 and 40 ms.

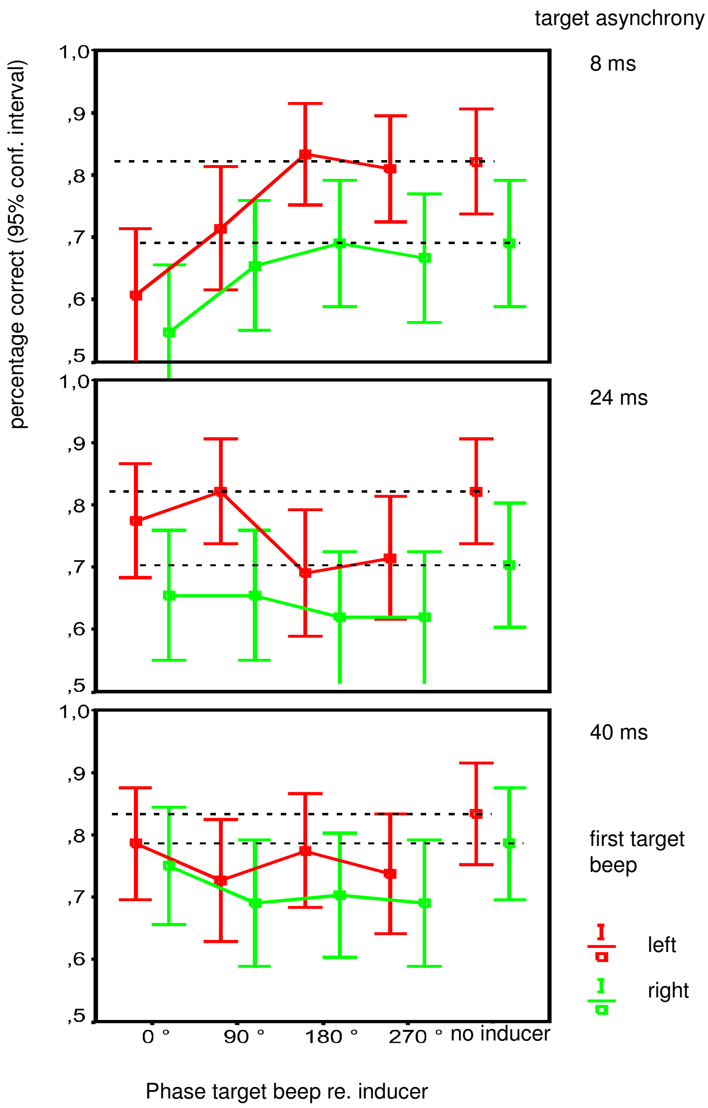

Figure 5. Performance in an auditory TOJ task during presentation of a visual inducer stimulus. Percent correct values and 95% confidence intervals are plotted as a function the target phase relative to the inducer for different SOA target values. Data from left and right monaural presentation are plotted separately in red and green color, respectively.

The main results of this experiment are depicted in Fig. 5. First, it was demonstrated that that performance in an intramodal (auditory) TOJ task is influenced by the presentation of a 10 Hz "inducer stimulus" in another (visual) modality. Second, performance for stimuli presented to the left ear was consistently higher than for stimuli presented to the right ear, in line with an often reported superiority of performance of the left ear (for review see Nicholls 1996). Third, for short stimulus onset asynchronies (8 ms) a conspicuous phase dependence could be noted with stronger impairment of TOJ performance for smaller relative phase values. Note that for phase 0 the two tone pips flank the visual inducer flash creating a situation somewhat complementary to that used by Scheier et al. (1999) who flanked visual stimuli of a visual TOJ task by auditory stimuli (although with longer crossmodal SOAs of 50 ms) and found improved TOJ performance in the condition with flankers.

Under the hypothesis that our experimental procedure indeed induced stable alpha oscillations, results obtained are more in line with those by Gho and Varela (1988) of a dependence of the TOJ performance from the phase of the ongoing alpha rhythm than with those indicating the existence of invariant moments. In a footnote of their paper, VanRullen and Koch (2003) stated that they were unable to reproduce the result by Gho and Varela. We have now modified the experimental design to be compatible with scalp EEG recording which will allow us to directly test our initial hypothesis of alpha induction by the flash.

Our further aim in workpackage 5 is to develop a suitable theoretical framework that (1) allows proper comparison of the different experimental approaches used by the different groups and (2) roots the described phenomenology about factors influencing TOJ in physiology.

4. Conclusion

We have successfully developed two experimental paradigms, one using rodents, the other using human subjects, for the study of neuronal mechanisms underlying two forms of audiovisual information fusion. The two paradigms exploit the complementary advantages of both rodent and human electrophysiology.

The rodent experiment allows superior spatial and temporal resolution of electrophysiological data due to the possibility of intracerebral recording. This experiment has demonstrated the feasibility of "high-level" category transfer from the auditory modality to the visual modality. Potential physiological correlates of this transfer have been identified using Granger causality analysis. The human experiment allows more convenient assessment of a subject's perception and was used to study the influence of visual inducers on performance in an auditory temporal order judgement task. A dependence of this influence on the relative phase between auditory target and visual inducer could be demonstrated.

Both experimental setups developed her define together a suitable environment for future development a theoretical framework in which various animal and human psychophysical data can be uniformely compared and rooted in a physiological. We expect that these data fundamentally extend current concepts of the role of cortical dynamics not only for crossmodal interaction, but for perception and cognition in general. The fundamental implications of LIN's work for the understanding brain processes of perception and categorization and its dissemination for engineering applications was featured in an overview article in the journal Nature (Abbott, 2006, p. 127).

5. Reference

(references marked with * are work by the author's group)

*[1] Abbott A (2006) In search of the sixth sense. *Nature* **442**: 125-127.

* Fillbrandt A, Deliano M, Ohl FW (2007) Audiovisual category transfer - an electrophysiological study in rodents. Proceedings of the 7th Meeting of the German Neuroscience Society/31st Göttingen Neurobiology Conference 2007 (submitted).

Kaminski M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern 85: 147-157.

* Ohl, F.W., Scheich, H. and Freeman, W.J. (2001) Change in pattern of ongoing cortical activity with auditory category learning. *Nature* **412**: 733-736.

* Ohl FW, Scheich H (2006) Neuroprothetik. Hitech im Gehirn. *Gehirn und Geist* **10**: 64-67.

Baer KV v (1864) Welche Auffassung der lebendigen Natur ist die richtige? Und wie ist diese Auffassung auf die Entomologie anzuwenden? *In Schmitzdorf H (Ed) Reden gehalten in wissenschaftlichen Versammlungen, und kleine Aufsätze vermischten Inhalts*: 237-283 St Petersburg: Verlag der Kaiserlichen Hofbuchhandlung.

Driver J, Spence C (1998) Multisensory perception: beyond modularity and convergence. *Curr Biol* **10**: 731-735.

Fechner GT (1860) Elemente der Psychophysik, Vol 22. Breitkopf und Haertel, Leipzig.

Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* **7**:773-778.

Gho M, Varela F (1988) A quantitive assessment of the dependency of the visual temporal frame upon the cortical rhythm. *Journal of Physiology (Paris)* **83**: 95-101.

Grondin S (2003) Processing time between visual events. *Arq Bras Oftamol* **66**: 96-104.

Hirsh IJ, Sherrick CE (1961) Perceived order in different sense modalities. *J Exp Psychol* **62**: 423-432.

Lalanne L (1876) Sur la durée de la sensation tactile. *Compt rend Acad de Sci Paris* **82**: 1314-1316.

Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision : examing temporal ventroliquism. *Cogn Brain Res* **17**: 154-163.

Navarra J, Vatakis A, Zampini M, Humphreys W, Soto-Faraco S, Spence C (2005) Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cogn Brain Res* **25**: 499-507.

Nicholls MER (1996) Temporal processing asymmetries between the cerebral hemispheres: evidence and implications. *Laterality* **1**:97-137

Pöppel E (1997) A hierarchical model of temporal perception. *TRENDS in Cognitive Sciences* **1**: 56-61.

[1] This is not a paper by the author but a feature article that portrays among other things the work of the author on cortical dynamics.

Scheier CR, Nijwahan R, Shimojo S (1999) Sound alters visual temporal resolution. Paper presented at the Investigative Ophtalmology and Visual Science meeting Fort Lauderdale FL.

Shimojo S, Shams L (2001) Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology* **11**: 505-509.

Spence C, Squire S (2003) Multisensory Integration: Maintaining the Perception of Synchrony. *Current Biology* **13**: 519-521.

Spence C, Shore DI, Klein RM (2001) Multisensory prior entry. *J Exp Psychol Gen* **130**: 799-832.

Stroud JM (1955) The fine structure of psychological time. In Quastler H (Ed) *Information Theory in Psychology*. The Free Press, Glenco Il.: 140-207

VanRullen R, Koch C (2003) Is perception discrete or continuous? *TRENDS in Cognitive Sciences* **5**: 207-213.

Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn Brain Res* **22**: 32-35.

Zampini M, Shore DI, Spence C (2003) Audioviusal temporal order judgments. *Exp Brain Res* **152**: 198-210.