

**Representation of visual stimuli in the inferotemporal cortex**  
**Invariance of perception despite modulation**

Summary of PhD Thesis

Kaposvári Péter

Supervisor: Gyula Sáry MD PhD  
Department of Physiology, University of Szeged

2011

## **Introduction**

Invariant sensory processing enables the sense of locality among the information units of the surroundings, but at the same time it also needs to be variable to adapt to a continuously changing environment. Robust coding guarantees the invariability and it also makes the sensory system able to learn by plasticity. Current research based on the investigation of the inferior temporal cortex of the macaque monkey in our laboratory supports the same view. The inferotemporal cortex (ITC) is a crucial member of the ventral stream. Here, the processed visual information becomes more and more complex, from the simple orientation discrimination in V1, (Hubel & Wiesel 1962) to the complex shape selectivity in the monkey ITC (Gross et al., 1969). The information and cellular projections converge in the ventral stream, thus the receptive field of the cells increases as we move from V1 towards ITC. It becomes binocular and up to 70 degrees large in area TE. This fact and the large receptive field which include the fovea with the position invariance of the cell's responses suggest an abstract way of the coding. The main feature of the ITC is shape selectivity, or shape preference. This feature was strongly consistent despite the change of the stimuli parameters, while some changes could cause modulation of response attributes, such as firing rate or latency. This thesis reports on one hand about the processing of Kanizsa type of illusory contours in the inferotemporal cortex and, on the other hand, about the changes in neural responses caused by auditory stimulation in this visual area.

# **Experiment I.: The representation of Kanizsa type illusory contours in the monkey inferior temporal cortex**

## **Introduction**

Perceived boundaries without physical differences between shape and background are called illusory or subjective contours (IC). They show that the perception is an active process; the sensory system gets the information in a flexible way. Thus, the investigation of these kinds of contour's processing can reveal mechanisms of perceptual organization, e.g., mechanisms of figure-ground segregation or contour integration.

Recently it has been shown how cells react to ICs based on abutting gratings pattern (Vogels & Orban 1987) in the monkey ITC (Sary et al., 2007). In the present experiment we tested whether ITC neurons respond to shapes defined by another type of illusory contours, the ones of the Kanizsa type (Kanizsa 1976).

## **Materials and Methods**

### **Stimuli**

The stimulus set contained four different conditions: line drawing (LD), silhouette (SIL), Kanizsa IC (KIC) and control (CTR) images. We used 9 black discs with white background as basic stimulus to test the baseline activity. They functioned as inducers when the shapes were presented on it in the Kanizsa condition. There were 20 stimuli in the different conditions.

Kanizsa type shapes with illusory contours were produced by cutting the overlapping parts of the shape from the discs. This way we generated perceived edges between the “pacmen” producing different Kanizsa type images.

In the control condition, the “pacmen” from the IC condition were rotated out from their original position (Brodeur et al., 2006), thus not producing KIC any longer whilst retaining the geometrical details of the inducers.

The line drawing condition was used to “hunt for” and isolate the cells.

### Stimulus sequence and single-cell recording protocol

A trial started with the onset of the fixation spot on the computer screen on dark background. If the animal fixated long enough, a white background was presented to level the baseline activity. After 300 ms nine black discs (the inducers without the notches) appeared for 400 ms followed by the stimuli themselves for 500 ms.

### Data analysis

We analysed spike counts in two 300 ms time windows off-line. The first window was set from -300 ms to 0 ms (0 = stimulus onset time) with the black discs inducers serving for the estimation of the baseline activity, and a second was used to calculate the stimulus-evoked response (100-400 ms).

To make a statistical comparison between the net response levels, we took the responses given to the most effective stimulus in every condition. This way, neuronal activities of the cells in response to the best (quasi optimal) images from our stimulus set in the different conditions could be compared.

The latency was calculated by using the “Poisson spike train analysis” only in case the cell had a significant response to at least one of the stimuli.

To assess the stimulus selectivity of the cells, tuning curves were constructed as follows: for each neuron and condition, first, the six selected stimuli according to their absolute net responses in one condition were ranked and then this sequence was followed in the other conditions.

## **Results I.**

### Single cell recording

The responses of 129 cells from the two monkeys tested for LD, KIC, SIL and CTR were analysed. There was no difference between the results of the two animals so the data were pooled. From these cells 99 (76.7 %) were found to be selective for the KIC (ANOVA). The following results were derived from these cells.

### Responsiveness and selectivity

On the population level, the mean  $\pm$  SEM baseline activity was  $11.7 \pm 1.11$  spikes/s and mean responses given to the LD, SIL, KIC and CTR were as follows:  $20.83 \pm 1.80$ ,  $28.10 \pm 2.13$ ,  $24.23 \pm 2.02$  and  $13.31 \pm 1.7$  spikes/s, respectively.

Responses given to LD, SIL, KIC and CTR differed significantly (ANOVA,  $F=35.6$ ,  $p<0.001$ ). A post-hoc (Fisher) test revealed that mean response levels to CTR were lower than in any other condition. This shows clearly that ITC cells react well in the different conditions, including KIC. In addition, we found that responses to SIL were larger than to LD.

We also wanted to show whether ITC cells were capable of responding selectively to KIC stimuli. We found that 76.7% of the recorded neurons were selective for the presented KIC stimuli (ANOVA). The tuning curves in the KIC and CTR conditions differed significantly. When the stimuli were ranked based on their responses to KIC (see methods), the tuning curve for KIC showed a significant slope (ANOVA,  $F=55.21$ ,  $p<0.001$ ) while the corresponding plot for CTR ran practically parallel to the X-axis (ANOVA,  $F=1.05$ ,  $p<0.383$ ). This indicates shape selectivity for KIC.

To show how this shape selective-response of the ITC neurons relate to the responses evoked by the same stimuli having real contours, we performed the same analysis regarding selectivity described for KIC and CTR above for the responses to KIC, LD and SIL. The selectivity in the KIC, LD and SIL conditions differed significantly, showing the flattest slope for LD. As stated above, KIC evoked significantly larger responses from ITC cells than LD (Wilcoxon,  $P<0.001$ ). To exclude the possibility of having tuning curve differences based on markedly different firing rates, we repeated the same analysis, but with the responses ranked according to the responses to LD. The results were similar, there was a highly significant difference between the tuning curves obtained for LD and KIC ( $F = 17.14$ ,  $p < 0.001$ ).

Since Kanizsa images often make the impression of a floating plane above the inducing elements, and ITC cells are sensitive to the inner features of a particular stimulus (Vogels & Biederman 2002) we tested whether the activity level evoked by KIC were similar to those to SIL of the same shapes.

The responses evoked by SIL were greater than those evoked by KIC (Wilcoxon,  $P < 0.001$ ), but we found a highly significant correlation between responses given to SIL and KIC (Spearman Rank Order Correlations, 0.77,  $p < 0.05$ ). Still, although the slope of the tuning curves obtained for KIC and SIL are more similar to each other than for KIC and LD, there was a significant difference between the curves (ANOVA,  $F=20.61$ ,  $P < 0.001$ ). It is worthwhile noting that the largest ratio of cells having similar selectivity was found in the KIC-SIL comparison.

#### Response latencies

The latency values obtained by the Poisson spike train analysis in the different conditions were 136.47 ms  $\pm$ SD 34.96 for LD, 135.39 ms  $\pm$ SD 30.33 for SIL and 147.99 ms  $\pm$ SD 36.06 for KIC. We found significant differences as estimated by ANOVA ( $F=7.30$ ,  $P < 0.001$ ,  $n = 76$ ), the latency values for KIC were longer than those measured in the two other conditions.

#### Discussion I

We presented shapes having real contours (LDs, SILs) and KICs to fixating monkeys while recording the neuronal activity in the ITC, and compared the responses to KICs and real contours. Our main findings can be summarized as follows:

- a) The neurones in the ITC are responsive to KICs.
- b) The neuronal response onset latencies for KICs are longer than those for shapes with real contours (LDs and SILs).
- c) The IT cortical cells exhibit clear shape selectivity for KIC stimuli.
- d) The shape selectivity of the ITC neurones differs from that for LDs or SILs, but the degree and similarity of selectivity were the closest between KICs and SILs.

a) KICs proved to be very effective stimuli with the mean response level to KICs somewhat even higher than that for real contours.

KIC stimuli elicited much higher responses than the CTR stimuli, which suggests that the responses to the KICs are due not only to the inducers (the pacmen), but also to the KICs inducing the shape from the background, even if real contours are absent.

b) The response latencies to the KICs were longer than those to the stimuli with real contours, LDs or SILs.

Processing of ICs is delayed in the early parts of the visual pathway as well. V1 neurones are responsive to the orientation of ICs defined by abutting gratings and KICs (Ramsden et al., 2001; Lee & Nguyen 2001), but the responses emerge later than those for real contours. This delay and the fact that the responses in V1 emerge 30 ms later than those in V2 suggest possible intercortical interactions during the processing of ICs.

c) Our data indicate that neuronal responses from the ITC to KICs can be used to construct tuning curves with a significant slope, while the same curves for CTR stimuli are essentially flat. This shows that ITC neurones have a stimulus preference and can reliably discriminate between stimuli even in the absence of real contours.

d) In our previous study we found that the selectivity of ITC cells was different for coloured stimuli and ICs, but IC selectivity did not differ from that of SILs (Sary et al., 2007). The present study demonstrated that shape selectivity for KICs was different from those for the SILs and LDs. This is somewhat surprising, since KICs usually give the impression of a floating surface, which lack internal visual surface information, in a similar way how SILs do.

Our results confirm the role of ITC in visual shape perception and generalize the ITC function in object and shape coding into the IC domain.

The response latency of the neurons recorded by us is longer in the case of the illusory contours (see also (Sary et al., 2007)) than of the real ones and of the silhouettes that represent the surfaces. This means additional processing, i.e. those higher order visual areas do not rely simply on the feed forward information flow when identifying shapes.

## **Experiment II: Auditory modulation of the inferior temporal cortex neurons in rhesus monkey**

### **Introduction**

There is a growing number of the evidence to support the idea that there are no separated unimodal systems, which exchange information only in the association areas to process them (Budinger et al., 2006; Schiller 1996). It is also known from everyday experience and from psychophysical studies how intricately visual and auditory information act together in shaping our knowledge of our surroundings and our acting in them.

Thus, we found it advisable to study the effects of auditory stimulation in that area, regarded as unimodal, in the anterior part of the inferotemporal cortex (TE). Inferotemporal cortex ITC is considered to be the last unimodal visual stage in the ventral visual stream (Tanaka 1996; Ungerleider & Mishkin 1982), consisting of several subregions, like anterior, middle and posterior part (Tamura & Tanaka 2001).

This is not the first study to test the effect of auditory stimuli in ITC, but they used the sound as a warning signal or joined it with act. In one of these studies (Iwai et al., 1987), one third of the recorded neuronal population in the ITC responded to pure tones. Another one (Watanabe & Iwai 1991) reported neurons of the posterior part of inferotemporal region (TEO) reacting to auditory signals, while others obtained similar results in ITC in split-brain monkeys (Ringo & O'Neill 1993). However, these studies might have found attention or action dependent modulated responses. Baylis et al., reported neurons receiving auditory signals mostly in TS and TAa areas of the temporal cortex (Baylis et al., 1987). They did not find considerable auditory activity in IT with poorly described auditory stimulus exactly.

In this study, we tested whether the neurons of the TE area of inferotemporal cortex, responsive to visual stimulation, could react to sole auditory stimulation.

## **Material and Methods**

### **Stimuli**

Colourful, complex images were shown on a standard computer display. Visual stimuli had a viewing angle of  $5 \times 6 \text{ deg}^2$ , and mean luminance was set at  $\sim 8 \text{ cd/m}^2$ . Auditory stimuli were presented from a computer loudspeaker positioned on top of the stimulation display (app. 60 dB, 440, 466, 493, 523, 554, 587, 622, 659, 698, 739, 783, 830, 880, 932, 987, 1046, 1108, 1174, 1244 and 1318 Hz, respectively).

Three stimulation conditions were used: one unimodal condition contained only the visual images (VIS); the other unimodal condition had only the auditory stimuli (AUDIO). For bimodal stimulation (condition: AUDIOVIS) the visual images were paired with those of the auditory stimuli. These pairings were fixed, i.e., an image was always presented with the same sound. A screen with a grey background, without the sound served as control.

### **Stimulus sequence**

A simple fixation task was used in this experiment. Trials started with the onset of a fixation spot, which was followed by a grey background (500 ms). Visual stimuli (VIS) appeared in this background and were presented for 500 ms. In the AUDIO condition, the grey background stayed on for another 500 ms, but a sound was presented instead of the images. In AUDIOVIS condition, sound was presented with the images for 500 ms. The animals were rewarded for fixating on the fixation spot within  $\sim 200 \text{ ms}$  after the disappearance of the stimuli. Neurons were isolated with the use of the VIS condition. Then, some stimuli (10 for monkey Ch and 6 for monkey Z) evoking stronger or weaker responses were selected. Next, during the registration, visual stimuli (VIS), sounds alone (AUDIO) and combined audiovisual stimuli (AUDIOVIS) were presented in a semi-random order, at least ten times each.

### **Data analysis**

For the calculation of net firing rates, the baseline activity (given to the grey background, measured in a 400 ms time window before stimulus onset) was trial wise subtracted from the responses (measured between 100 ms and 500 ms after stimulus onset).

T-test was used to test responsiveness of the cells when comparing the background and the firing activity during the stimulus presentation.

Friedmann ANOVA non-parametric test with a post-hoc Sign-test for dependent pairs was used to test the selectivity. Tests were classified significant if the corresponding type I error was  $< 0.01$ .

For finding a modulatory effect in the AUDIOVIS condition, we made the following analysis: we selected a neuron responding to at least one member of the stimulus pairs in the VIS and in the AUDIOVIS conditions. Net responses were compared with a t-test. A cell was regarded as “modulated” if there was a significant difference in the responses given to the stimulus pairs.

Latency measurements were performed using a modified Poisson spike train analysis in each case in which a stimulus triggered a response.

## **Results II.**

### General responsiveness

From the isolated units 92 neurons were responsive at least in one of the three conditions (monkey 1: 67, monkey 2: 25). In general, neurons responded well in the VIS and also in the AUDIOVIS condition (mean net responses  $23.5 \pm 18.5$  and  $22.8 \pm 19.3$  spikes/s, respectively), showing the well-known stimulus preference of ITC units (mean SI:  $0.573 \pm 0.227$  and  $0.573 \pm 0.226$  in the VIS and AUDIOVIS conditions). There was no significant difference between the mean neuronal response latency in the VIS condition ( $137.8 \text{ ms} \pm 35.8$ ) and the AUDIOVIS condition ( $136.9 \text{ ms} \pm 36.3$ ).

### Response/modulation in ITC by auditory signals

There were 9 cells (10% of the sample), which responded at least to one of the auditory signals in the AUDIO condition, having a mean firing rate of  $7.15 \text{ spikes/s} \pm 4.05$ . There was no difference in the firing rate between the different conditions in this sub-sample. The mean latency time to the AUDIO was  $137.70 \pm 32.3$  which did not differ significantly from the latency times measured in the VIS ( $129.43 \pm 25.4$ ) or AUDIOVIS ( $128.13 \pm 26.3$ ) condition in the same cells as tested with Friedman ANOVA.

In 22 neurons (~ 24 % of the sample), although responses given in the AUDIO condition did not reach significance level (except 5 neurons), the responses differed between the VIS and the AUDIOVIS conditions at least in one of the stimulus pairs (16 cells 1 pair, 5 cells 2 pairs, 1 cell 3 pairs).

We could not find a universal pattern for the modulation, i.e., response rates could either increase or decrease in the bimodal condition as compared to responses in VIS.

Response latency values on the population did not differ significantly between VIS and AUDIOVIS. Some cells showed an increase, others a decrease in the latency when changing stimulation condition.

## **Discussion II.**

We found a number of neurons in area TE of the ITC responded to auditory stimulation and also showed the modulation of responses induced by visual stimuli when they were presented simultaneously to auditory stimulus. This is clear evidence that auditory modality does have an effect on the ITC neuronal visual responses and that it impacts information processing at a stage previously regarded as unimodal.

A connectivity study of Saleem et al., (2000) reported anatomical connections between area TE and the SPS, thus, there might be a “pre-wired” pathway for sound stimuli that reaches the inferotemporal cortex.

Earlier reports exist on sound related activity in the inferotemporal cortex of the monkey. In the next section, we compare these studies with our present study. In general, with the exception of Baylis et al., (1987), who did not present sufficient data concerning the auditory stimulus parameters and stimulating conditions, none of the papers discussed below presents pure auditory effects in area TE. Other research group (Desimone & Gross 1979) failed to find cells in this area. However, we must note that they used anaesthetized animals. This might explain the differences between their and our studies, even if our animals had to perform only a simple fixation task.

In a study (Iwai et al., 1987), the precise recording location, i.e., TEO or TE was not defined, and the sound was presented earlier than the visual stimuli. The response latencies for the auditory signal were long (on average over 500 ms), excluding a direct auditory effect on ITC

cells, because of the much shorter latency values observed in these neurons. As the authors also stated, the sound related signals were more probably a manifestation of some attentional mechanisms.

Watanabe & Iwai (1991) performed their recordings in areas like the auditory area (AA), STP and TEO of the ITC. On one hand, the closest area to TE, the TEO did not respond to the isolated auditory signals. AA and STP contained several sound responsive neurons, which is not surprising, since these areas are known to be auditory and polysensory (Watanabe & Iwai 1991). Cells in area TEO did not react to the sounds only, but also caused a modulation of the visual responses. Again, this modulation seems to be attention-dependent, since the monkey was warned by the sound. Similarly, in the study of Ringo & O'Neill (1993), split brain monkeys were exposed to alerting tone signals, while recording in ITC. Some neurons were responsive (17 cells from 308 units) to the warning signals, but warning preceded the visual stimuli. The authors also stated their results resembled those of Iwai et al., (1987).

In our study the sounds did not carry biological relevance. The animals were fixating, receiving their reward for keeping their eye on the fixation spot. For this reason, what happened on the screen, or what sound was presented during the visual stimulation, was irrelevant for the monkeys. Finally, we believe that the results we present are due to the sensory processing of the auditory stimuli and not to some task related or attentional effect.

In this study, we wanted to clarify whether or not ITC neurons in TE area responded to auditory signals presented together with visual stimuli. This implicates that correlated sensory modalities might converge and have modulatory effects in sensory areas earlier considered as strictly unimodal, even if the stimuli belong to different modalities.

## **General discussion**

The experiments presented here fit in the history of the laboratory. Shape processing of the TE was investigated in a systematic way. Shapes with reduced surface information were tested (Kovacs et al., 2003). Colour, texture, shading, contrast and inner contours were systematically removed from the stimuli. Shape selectivity of the cells was not changed, but the responses were bigger for the coloured shapes in comparison to the reduced version of them. The same result was found with achromatic stimuli. The shape selectivity was not changed, but the response rate decreased in the black and white condition (Tompa et al.,

2005). In a preliminary experiment we found possible selectivity difference between colour shapes and line drawing (Kaposvári et al., 2009). It seems that the selectivity parameter of the shape representation in TE for the reduced visual stimuli is strongly invariant.

To test the existence of Biedermann's theory in ITC, in another preliminary experiment shape selectivity differences were not found in contour deleted, recoverable and unrecoverable stimuli versus line drawing (Sáry et al., 1998). To test the neuronal responses for more degraded stimuli, illusory shapes made by abutting gratings were used (Sary et al., 2006). The result of this experiment was compared to the present study about illusory contour. To follow this line I reported two experiments. The first described the processing of Kanizsa illusory shapes; the second discussed the modulation effect of sounds in the TE.

Our experiments give additional information to complete the view of information processing regarding the role of ITC in perception. Everyday experience suggests the holistic way to the representation of the environment. This concept is also supported by the anatomical connections among the different sensory areas.

We found electrophysiological evidence of pure tone effect in a crucial stage of high-level visual perception. We used a simple fixation task, the stimulus had no biological relevance, and thus the modulation seems to have been caused by solely passive information process. But this process is usually not passive; the sensory integration has an essential function in the representation, perception and planning of an action. To perform these tasks the system needs the feed forward information process as well. These connections convert the sensory system to an active perceptual system. In the second experiment, we tested the effect of Kanizsa type illusory shape in ITC. These kinds of illusion represent well the active organisation of the information in the early stages of the visual system. However, we gained some results for the rules regarding KIC perception in the more complex area. ITC. They show that the ITC cells need to process the KIC shapes, and they process them in a different way, as compared to shapes with real contours. It seems the ITC is not a passive stage of the visual system; on the contrary, it takes an active part in information processing.

## **Acknowledgement**

I would like to thank prof. Benedek György, the former head of the Department of Physiology, and prof. Jancsó Gábor, the present head of the Department who made my work possible in the Institute.

It is a pleasure to thank my supervisor, Sárosi Gyula. He supported my work in a number of ways.

I received an excellent training at the Department of Neurophysiology at the KULeuven; I am grateful to prof. Rufin Vogels and Joris Vangeneugden.

I would like to thank my former colleagues, Köteles Károly and Tompa Tamás who helped a lot during the experiments and analysis.

I am grateful to the former and present students: Bognár Anna, Csete Gergő, Csibri Péter, Kincses Bálint, Kis Alexandra, Ferge Petra, Frankó Edit, Nász Veronika, Utassy Györgyi, Varga Zsuzsa.

I also have to thank those colleagues who helped with my administration duties, maintained the experimental setup and took care for the animals: Jász Anikó, Siposné Dósa Gabriella, Hermann Kálmán, Lízli Péter, Rácz László.

I have to thank all my colleagues in the Department of Physiology who provided an excellent atmosphere for my work and all my friends who never failed to help me whenever needed.

Finally, my greatest thanks go to my parents. Their support made all these possible.

### **The applicant's relevant publications**

- I. SÁRY, G., KOTELES, K., KAPOSVARI, P., LENTI, L., CSIFCSAK, G., FRANKO, E., BENEDEK, G., TOMPA, T., (2008) The representation of Kanizsa illusory contours in the monkey inferior temporal cortex, *Eur. J. Neurosci.* 2008. Nov. ;28. (10):2137. -46., 28, 2137-2146.
- II. KAPOSVÁRI P. CSIBRI, G. CSETE, T. TOMPA AND GY. SÁRY, Auditory modulation of the inferior temporal cortex neurons in rhesus monkey, *Physiological Research* (in press)

### Posters

- I. GY. SÁRY, T. TOMPA, K. KÖTELES, P. KAPOSVÁRI AND GY. BENEDEK, Beyond physical boundaries: coding of two types of illusory contours in the macaque inferotemporal cortex, *MITT* 2008
- II. GY. SÁRY, T. TOMPA, P. KAPOSVÁRI AND GY. BENEDEK, Beyond physical boundaries: coding of abutting-line and Kanizsa-type illusory contours in the monkey inferotemporal cortex, *MET* 2008
- III. P. KAPOSVÁRI, P. CSIBRI, G. CSETE, T. TOMPA AND GY. SÁRY, Sounds modulate neurons in the visual cortex of the monkey, Belgrad, International symposium, One hundred years of Ivan Djaja's 2010
- IV. P. KAPOSVÁRI, P. CSIBRI, G. CSETE, T. TOMPA AND GY. SÁRY, Sounds modulate neurons in the visual cortex of the monkey, *MITT* 2011

### Presentations

- I. KAPOSVÁRI PÉTER, CSETE GERGŐ, CSIBRI PÉTER, TOMPA TAMÁS, SÁRY GYULA, Auditorikus hatások a látórendszerben (Bimodális integráció a majom inferotemporális kérgében), *Látás szimpózium 2010* (presentation)