

# Auditory Modulation of the Inferior Temporal Cortex Neurons in Rhesus Monkey

P. KAPOSVÁRI<sup>1</sup>, P. CSIBRI<sup>1</sup>, G. CSETE<sup>1</sup>, T. TOMPA<sup>1</sup>, GY. SÁRY<sup>1</sup>

<sup>1</sup>Department of Physiology, University of Szeged, Szeged, Hungary

Received January 3, 2011

Accepted February 16, 2011

On-line July 19, 2011

## Summary

We performed a systematic study to check whether neurons in the area TE (the anterior part of inferotemporal cortex) in rhesus monkey, regarded as the last stage of the ventral visual pathway, could be modulated by auditory stimuli. Two fixating rhesus monkeys were presented with visual, auditory or combined audiovisual stimuli while neuronal responses were recorded. We have found that the visually sensitive neurons are also modulated by audiovisual stimuli. This modulation is manifested as the change of response rate. Our results have shown also that the visual neurons were responsive to the sole auditory stimuli. Therefore, the concept of inferotemporal cortex unimodality in information processing should be re-evaluated.

## Key words

Monkey • Inferotemporal cortex • Auditory modulation

## Corresponding author

Gy. Sály, Dept. of Physiology, University of Szeged, 10. Dóm tér, Szeged, Hungary. Fax: +36 62 545 842. E-mail: [sary@phys.szote.u-szeged.hu](mailto:sary@phys.szote.u-szeged.hu)

## Introduction

It is more and more evident that we cannot regard the brain as made up of independent unimodal streams of information joining only at the association areas (Budinger *et al.* 2006, Schiller 1996). We also know from everyday experience and from psychophysical studies how intricately visual and auditory information acts 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 par excellence unimodal, 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 and Mishkin 1982), consisting of several subregions, like anterior, middle and posterior part (Tamura and Tanaka 2001). In one study (Iwai *et al.* 1987), one third of the recorded neuronal population in the ITC responded to pure tones. Another study (Watanabe and 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 and O'Neill 1993). Baylis and his co-workers reported neurons receiving auditory signals in TS and TAa areas of the temporal cortex (Baylis *et al.* 1987). However, these studies might have found attention or action dependent modulated responses. 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

### Subjects

Two adult *Rhesus* monkeys (monkey Ch, a male and monkey Z, a female; 6.8 kg and 6 kg, respectively) participated in the study. The monkeys worked under a controlled water access paradigm in daily recording sessions, 5 days a week. During weekend, they were allowed to drink water *ad libitum* and fruits and vegetables were added to their diet. All procedures conformed to the guidelines of the National Institutes of Health Guide for the Care and Use of Laboratory Animals and the guidelines of the University of Szeged Animal Care and Use Committee.

### Surgery

Prior to the recordings, the monkeys underwent two aseptic surgeries (Sáry *et al.* 2008) performed under general anesthesia to receive a headpost and recording chamber over the ITC on the skull, and a scleral search coil in the eye (Judge *et al.* 1980). Anesthesia was induced with Calypsol (Ketamine, 20 mg/kg i.m.) and maintained with 1.5 % Halothane in a 2:1 mixture of N<sub>2</sub>O and O<sub>2</sub>. Atropine (0.05 mg/kg i.m.), was injected to reduce bronchial secretion and salivation. Body temperature, heart rate, respiratory rate, end-tidal CO<sub>2</sub> and peripheral O<sub>2</sub> saturation were monitored throughout the surgery. Analgesic was given after surgery (Nalbuphine, 0.15 mg/kg) for 5 days.

### Stimuli and recordings

Colorful, complex images (Sary *et al.* 2004) were shown on a display (Philips Brilliance 17A having 17" diameter; refresh rate, resolution and viewing distance were 74 Hz, 800x600 pixels and 57 cm, respectively). Visual stimuli had a viewing angle of 5x6 degs, and mean luminance was 7.9 cd/m<sup>2</sup>. 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 stimulus (AUDIO). For bimodal stimulation (condition: AUDIOVIS) the visual images were paired with one of the auditory stimuli. This pairing was fixed, i.e., an image was always presented with the same sound. A screen with a grey background, without the sound served as control (see stimulus sequence).

Standard electrophysiological equipment was used (FHC inc., Bowdoin, ME, USA) with electrodes having impedance values of 1-3 MΩ (FHC). During the recording sessions, the monkeys sat with their heads fixed. A custom-made software running on a PC recorded the eye movements (200 Hz sampling rate), delivered the reward, controlled the animals' behavior, and presented the stimuli. Another PC collected the electrophysiological data. The background luminance in the sound-proof experimental room was kept constant at a level <1 cd/m<sup>2</sup>. For a detailed description of the above methods see Kovacs *et al.* (2003).

### Recording site

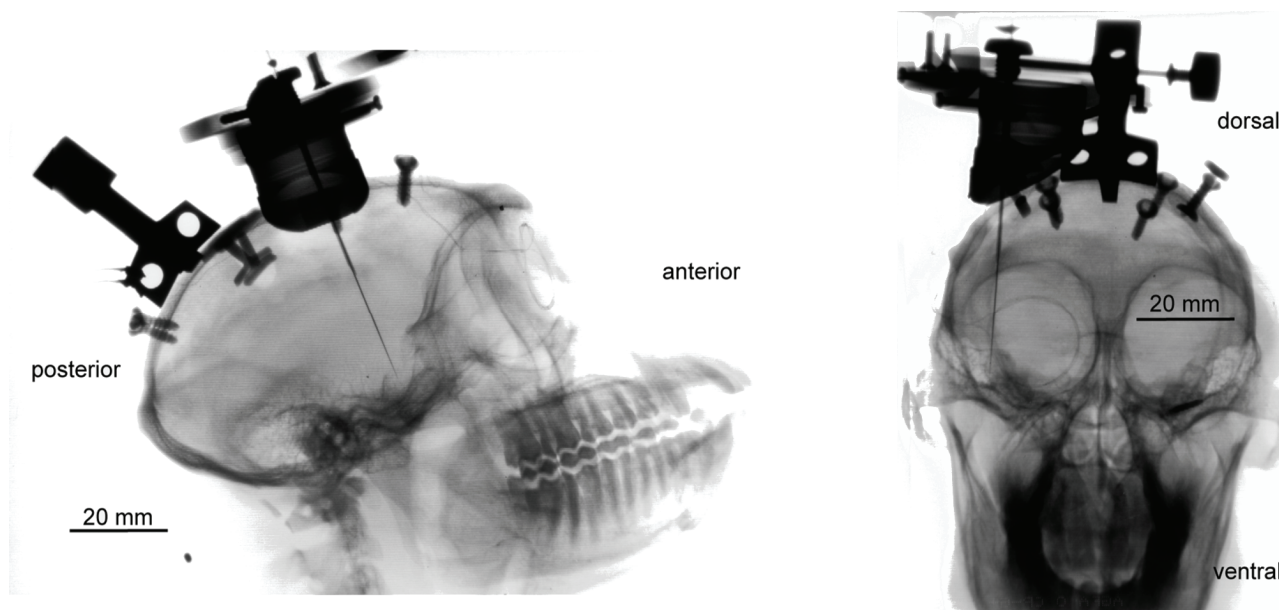
Although the histological verification of the recording site is not available yet, we strongly believe that the recordings were done in the anterior part of the ITC for the following reasons: recorded neurons showed responses typical of this region (response latency, strong presence of stimulus selectivity). The recording chamber was mounted on the basis of the atlas of the monkey brain (Paxinos *et al.* 1999), and on previous MRI. On the bases of measures obtained from the MRI images, the stereotaxic atlas coordinates were corrected if necessary: the recording chamber was centered at stereotaxic coordinates at AP: +16 mm, ML: 22 mm. Our recordings were made in the appropriate area by reaching the subarachnoidal space at the end of the recording session. Neuronal activity was collected from two cortical regions, before the electrode reached the bone: the lower bank of the superior temporal sulcus (STS), and the bottom of the TE. They are known as TEa and TE1 (Baylis *et al.* 1987). We clearly recognized the white matter between these two parts of the TE, and also the gap between the two banks of the STS. Finally, X-ray images made in the coronal and sagittal planes in monkey Z with an *in situ* electrode showed our electrode positioned into area TE of the ITC (Fig. 1).

### Stimulus sequence

The animals were trained in a simple fixation task (Sary *et al.* 2006). Fixation window had a side length of 0.8 deg, the fixation spot had a radius of 3 pixels. 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. The animals were rewarded for fixating on the fixation spot within 100-350 ms after the disappearance of the stimuli. Breaking the fixation resulted in an abort; only fully completed trials were included in the analysis. Neurons were isolated using 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

Neuronal responses were analyzed off-line, with



**Fig. 1.** X ray images from monkey Z. **A:** Lateral view from the skull, showing the surgical screws, headpost and recording chamber. The electrode – not lowered completely – points to area TE of the inferotemporal cortex. **B:** Frontal view of the skull. The electrode is lowered at app. lateral 22 mm from the midline.

the use of the net firing rates. For this, the baseline activity (given to the gray 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). Only the cells responding at least in one of the three conditions (VIS, AUDIO and AUDIOVIS) were further analyzed. Since inferotemporal neurons might exhibit stimulus-preference, a selectivity index (SI) was calculated, based on the following formula:

$$SI = R_{max} - R_{min} / R_{max} + R_{min}$$

where  $R_{max}$  is the maximal and  $R_{min}$  is the minimal response to a stimulus from the stimulus set, respectively. The closer this index is to “1”, the bigger the difference between a “preferred” and a “non-preferred” stimulus is, i.e., the more selective the neuron is.

For finding a modulatory effect in the AUDIOVIS condition, we made the following analysis: we selected 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.

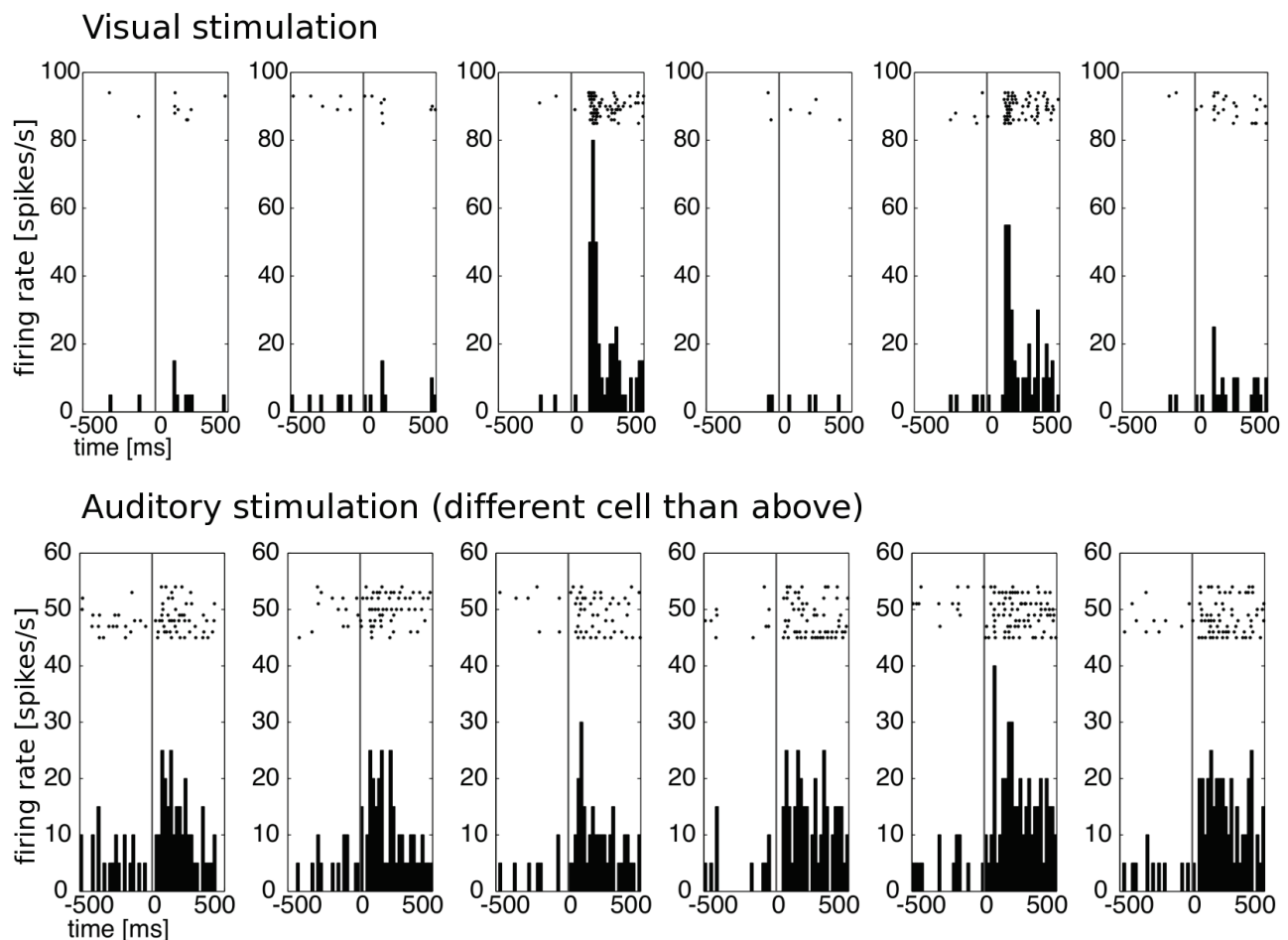
Friedmann ANOVA non-parametric test with post-hoc Sign-test for dependent pairs were used for statistical evaluation. Tests were classified significant if

the corresponding type I error was  $<0.01$ . Latency measurements were performed using a modified Poisson spike train analysis (Legendy and Salcman 1985, Sary *et al.* 2006) in each case in which a stimulus triggered a response.

## Results

### General responsiveness

From the isolated units 92 neurons were responsive at least in one of the three conditions (monkey Z: 67, monkey Ch: 25). Since there was no difference between the results of the two animals (regarding response amplitude, response latency, SI indices), data were pooled. 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 (Kobatake *et al.* 1998, Wang *et al.* 2000, Liu and Jagadeesh 2008) (mean SI:  $0.573 \pm 0.227$  and  $0.573 \pm 0.226$  in the VIS and AUDIOVIS conditions, respectively, Fig. 2). There was no significant difference of the mean neuronal response latency in the VIS condition ( $137.8 \text{ ms} \pm 35.8$ ) versus the AUDIOVIS condition ( $136.9 \text{ ms} \pm 36.3$ ). 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  $7.15 \text{ spikes/s} \pm 4.05$  (Fig. 2). There was no difference in the firing rate between the different



**Fig. 2.** Individual inferotemporal neuronal responses to visual (upper panel) and auditory (bottom panel) stimulus. Peristimulus time histograms show the mean firing rates after ten presentations of each stimuli. **A:** ITC neuron responding only to two of the six visual stimuli. The grey background appeared at  $-500$ , stimulus onset was at  $0$ , and stimuli disappeared at  $500$ . Vertical axis indicates firing rates in spikes/s. **B:** ITC neuron responding to auditory stimuli.

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.

#### *Response/modulation in ITC by auditory signals*

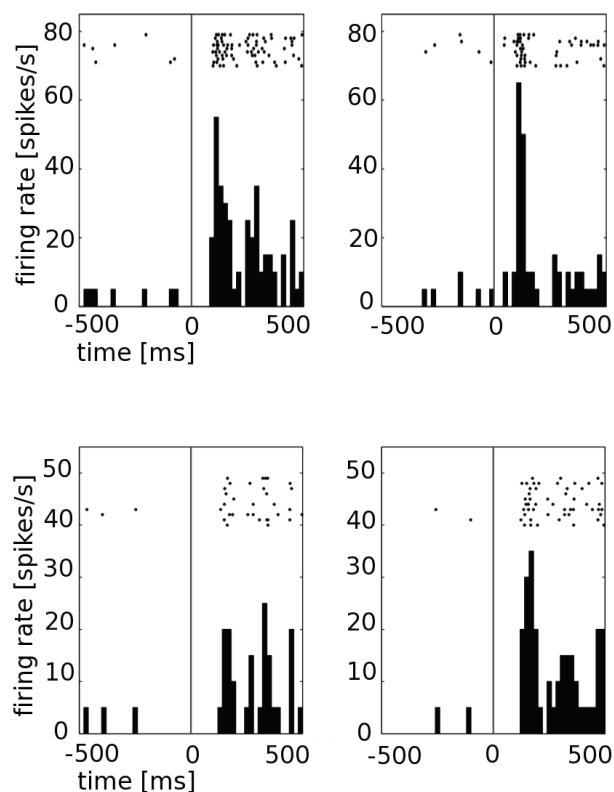
In 22 neurons ( $\sim 24\%$  of the sample), although responses given in the AUDIO condition did not reach significance level (except 5 neurons), responses differed between the VIS and the AUDIOVIS conditions at least in one of the stimulus pairs (16 cells 1 pair, 5 cells 2 pairs, 1 cell 3 pairs).

Fig. 3 shows examples, where responses given to a VIS stimulus were modulated in the bimodal, AUDIOVIS condition. 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. The distribution of

response changes between the conditions is shown in Fig. 4. Response latency values on the population did not differ significantly between VIS and AUDIOVIS. Some cells showed an increase, others a decrease of the latency when changing stimulation condition. The distribution of latency differences is presented in Fig. 4.

## Discussion

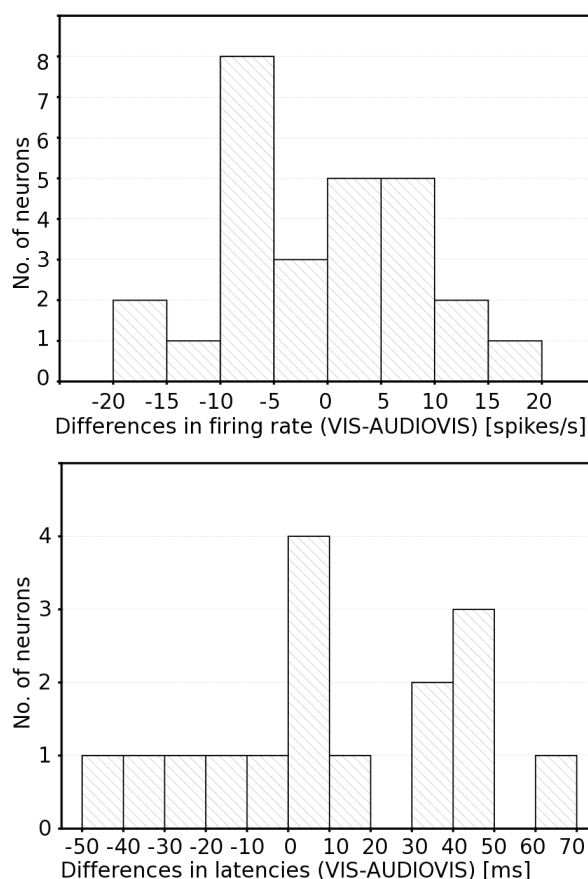
Here we report that a number of neurons in area TE of the ITC responded to auditory stimulation and showed modulation of the responses induced by visual stimuli, when they were presented simultaneously to auditory stimulus. This is clear evidence that other modalities (e.g., auditory) are having an effect on the ITC neuronal visual responses and impacts the 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,



**Fig. 3.** Examples of two ITC neurons response modulation. The left sides in both panels show the neuronal responses to the visual stimuli, while the right sides show their responses when the visual and auditory stimuli were presented simultaneously. Peristimulus time histograms show the mean firing rates after ten presentations of the stimuli.

thus, there might be a “pre-wired” pathway for sound stimuli reaching the inferotemporal cortex.

Since the ITC has been regarded as the last unimodal area in the ventral visual stream, we might face some objections. The first might be that our recordings perhaps were not collected from the inferotemporal cortex, since right above the STS there is the auditory cortex and the electrode has to pass on it. However, as stated in the method section, we are quite confident that our cells were recorded in area TE of ITC. Since sounds were presented together with a neutral grey visual background, one has to consider the possibility that the responses we recorded were indeed responses given to the grey background. If there was a response to the grey background, that should be already present from the start of the trial. If the response onset is time locked to the onset of the auditory signal, it could not be a response to the grey background (a possible visual stimulus) unless these cells had unusual long response latency. To be on the safe side, cells responding to this visual stimulus (neutral grey) or the auditory responses that did not have frequency dependence, were excluded from the analysis.



**Fig. 4.** Effects of audio-related modulation in the ITC. The upper panel shows the distribution of the differences in firing rates between the VIS and AUDIOVIS conditions. The lower panel shows the distribution of the latency differences between VIS and AUDIOVIS.

This later is a very strict criterion, since some ITC neurons might just signal the presence of the auditory stimulus, without being frequency selective. Furthermore, neuronal response latencies collected in this study ( $116.5 \text{ ms} \pm 30.9$ ) were practically identical to those collected earlier in our laboratory in different animals (mean latency:  $114 \text{ ms} \pm 21.0$ ). Also, selectivity indices obtained in the present study ( $0.573 \pm 0.227$ ) were similar to those reported earlier (Sary *et al.* 2004).

The possible effects of noises coming from outside the recording chamber, or behavior-related effects, like reward expectancy, etc. should also be considered. Our animals were seated in a soundproof recording room, excluding external effects. Tasks performed correctly ended in a reward, coming at 100–300 ms after stimulus offset. This jitter makes it unlikely that the modulatory effects seen in our sample would be related to the reward; on the other hand it is very unlikely that reward expectancy or any other behavior related effect would cause frequency dependent responses, like

observed in some neurons.

There exist earlier reports 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. Desimone and Gross (1977) failed to find cells in this area. However, we have to note that they used anaesthetized animals. This might explain the differences between this and our study, even if our animals had to perform a simple fixation task only.

In one 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 and 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 and 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 *et al.* (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. We report that a substantial number of ITC neurons in the monkey responded to auditory stimuli, or modulated their responses to visual images when visual and auditory stimuli were presented simultaneously. Excluding other possible sources, we believe that our results represent the first real report on auditory sensitive neurons in the monkey inferotemporal cortex, in area TE. 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.

## Conflict of Interest

There is no conflict of interest.

## Acknowledgements

This work was supported by: OTKA # K68594 and OTKA # K83671. The authors are thankful to G. Dosai, P. Liszli and L RÁCZ for the technical support and for taking care of the animals.

## References

- BAYLIS GC, ROLLS ET, LEONARD CM: Functional subdivisions of the temporal lobe neocortex. *J Neurosci* **7**: 330-342, 1987.
- BUDINGER E, HEIL P, HESS A, SCHEICH H: Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience* **143**: 1065-1083, 2006.
- DESIMONE R, GROSS CG: Visual areas in the temporal cortex of the macaque. *Brain Res* **178**: 363-380, 1979.
- IWAI E, AIHARA T, HIKOSAKA K: Inferotemporal neurons of the monkey responsive to auditory signal. *Brain Res* **410**: 121-124, 1987.
- JUDGE SJ, RICHMOND BJ, CHU FC: Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* **20**: 535-538, 1980.
- KOBATAKE E, WANG G, TANAKA K: Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J Neurophysiol* **80**: 324-330, 1998.

- KOVACS G, SARY G, KOTELES K, CHADAIDE Z, TOMPA T, VOGELS R, BENEDEK G: Effects of surface cues on macaque inferior temporal cortical responses. *Cereb Cortex* **13**: 178-188, 2003.
- LEGENDY CR, SALCMAN M: Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *J Neurophysiol* **53**: 926-939, 1985.
- LIU Y, JAGADEESH B: Neural selectivity in anterior inferotemporal cortex for morphed photographic images during behavioral classification or fixation. *J Neurophysiol* **2**: 966-982, 2008.
- PAXINOS G, HUANG FX, TOGA AW: *The Rhesus Monkey Brain in Stereotaxic Coordinates*. Academic Press, 1999.
- RINGO JL, O'NEILL SG: Indirect inputs to ventral temporal cortex of monkey: the influence of unit activity of alerting auditory input, interhemispheric subcortical visual input, reward, and the behavioral response. *J Neurophysiol* **70**: 2215-2225, 1993.
- SALEEM K S, SUZUKI W, TANAKA K, HASHIKAWA T: Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J Neurosci* **20**: 5083-5101, 2000.
- SARY G, CHADAIDE Z, TOMPA T, KOVACS G, KOTELES K, BODA K, RADULY L, BENEDEK G: Relationship between stimulus complexity and neuronal activity in the inferotemporal cortex of the macaque monkey. *Brain Res Cogn Brain Res* **22**: 1-12, 2004.
- SARY G, KOTELES K, CHADAIDE Z, TOMPA T, BENEDEK G: Task-related modulation in the monkey inferotemporal cortex. *Brain Res* **1121**: 76-82, 2006.
- SÁRY GY, KÖTELES K, KAPOSVÁRY P, LENTI L, CSIFCSÁK G, FRANKÓ E, BENEDEK G, TOMPA T: The representation of Kanizsa illusory contours in the monkey inferior temporal cortex. *Eur J Neurosci* **28**: 2137-2146, 2008.
- SCHILLER PH: On the specificity of neurons and visual areas. *Behav Brain Res* **76**: 21-35, 1996.
- TAMURA H, TANAKA K: Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb Cortex* **11**: 384-399, 2001.
- TANAKA K: Inferotemporal cortex and object vision. *Annu Rev Neurosci* **19**: 109-139, 1996.
- UNGERLEIDER LG, MISHKIN M: Two cortical visual systems. In: *Analysis of Visual Behaviour*. DJ INGLE, MA GOODALE, RJW MANSFIELD (eds), The MIT Press, Cambridge, MA, USA, 1982, pp 549-580.
- WANG Y, FUJITA I, MURAYAMA Y: Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nat Neurosci* **3**: 807-813, 2000.
- WATANABE J, Iwai E: Neuronal activity in visual, auditory and polysensory areas in the monkey temporal cortex during visual fixation task. *Brain Res Bull* **26**: 583-592, 1991.