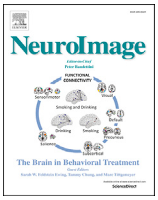
NeuroImage 161 (2017) 67–79

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage



Cortical representation of persistent visual stimuli

Edden M. Gerber[a,*], Tal Golan[a], Robert T. Knight[b,c], Leon Y. Deouell[a,d,**]

[a] *Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, 9190401, Israel*
[b] *Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA*
[c] *Dept. of Psychology, University of California, Berkeley, Berkeley, CA 94720, USA*
[d] *Dept. of Psychology, The Hebrew University of Jerusalem, Jerusalem, 9190501, Israel*

ARTICLE INFO

Keywords:
Electrocorticography
Visual cortex
Sustained perception
High-frequency activity
Early visual cortex
Inferior temporal cortex
Fusiform face area

ABSTRACT

Research into visual neural activity has focused almost exclusively on onset- or change-driven responses and little is known about how information is encoded in the brain during sustained periods of visual perception. We used intracranial recordings in humans to determine the degree to which the presence of a visual stimulus is persistently encoded by neural activity. The correspondence between stimulus duration and neural response duration was strongest in early visual cortex and gradually diminished along the visual hierarchy, such that is was weakest in inferior-temporal category-selective regions. A similar posterior-anterior gradient was found within inferior temporal face-selective regions, with posterior but not anterior sites showing persistent face-selective activity. The results suggest that regions that appear uniform in terms of their category selectivity are dissociated by how they temporally represent a stimulus in support of ongoing visual perception, and delineate a large-scale organizing principle of the ventral visual stream.

1. Introduction

Although visual perception of scenes or objects typically occurs over an extended time window, what we know of human high level visual perception is largely based on onset responses, that is, responses to change (e.g. the EEG/intracranial face-selective N170/N200; Bentin et al., 1996; Allison et al., 1999). While onset responses hold a wealth of information on the functional properties of the respective neural population, they cannot disentangle the transient response to change, from neural activity driven by the ongoing presence of a percept. Thus, how real-time experience persists beyond stimulus onset is unknown, leaving the neural basis for the brunt of the time course of perception unaccounted for. Put simply, our question is how do we know, in real-time and beyond the initial detection, whether a stimulus we gaze at is still present 500 or 1500 ms beyond its onset, and how do we know its content, e.g. whether it is (still) a face or an object? This question is distinct from that of explicit duration estimation tasks (e.g. was a probe stimulus shorter or longer than a reference stimulus; Buhusi and Meck, 2005; Lewis and Miall, 2003) in which judgments are performed post-hoc, that is, after stimulus offset. Here, we are concerned with how the visual system codes the continued presence of an object in real time. Some evidence of sustained responses to visual stimuli can be gleaned from single unit (Ikeda and Wright, 1974; Kulikowski et al., 1979; Petersen et al., 1988) or fMRI studies (Gilaie-Dotan et al., 2008) which used long duration visual stimuli, and from variable duration EEG/MEG responses in an explicit duration estimation task (N'Diaye et al., 2004; Pouthas et al., 2000). However these studies did not directly examine the neural basis of sustained, real-time perception.

We recorded from ten human subjects implanted with subdural electrodes over visual cortices using electrocorticography (ECoG) while subjects were engaged in a novel target detection paradigm, with faces and objects presented for variable durations. Since even a brief stimulus can produce a sustained response (Fisch et al., 2009), using variable stimulus durations was crucial for isolating the part of the response driven specifically by the stimulus ongoing presence. Notably, no duration estimation was required.

We found that in early visual cortex (EVC) the high-frequency broadband response (HFB, >30 Hz) closely tracked the time course of the stimulus, but this precision decreased along the visual hierarchy, i.e. from V1/V2 to V3/V4, from early visual cortex to inferior temporal (IT) cortex, and from posterior to anterior IT. Only the posterior part of IT robustly encoded the presence of the stimulus over time. This posterior-anterior gradient could not be explained by signal-to-noise ratio, nor by differential influences of attention.

---

* Corresponding author. Edmond and Lily Safra Center for Brain Sciences, The Hebrew University, The Edmond J. Safra Campus, Jerusalem, 9190401, Israel.
** Corresponding author. Edmond and Lily Safra Center for Brain Sciences, The Hebrew University, The Edmond J. Safra Campus, Jerusalem, 9190401, Israel.
E-mail addresses: edden.gerber@mail.huji.ac.il (E.M. Gerber), leon.deouell@mail.huji.ac.il (L.Y. Deouell).

http://dx.doi.org/10.1016/j.neuroimage.2017.08.028
Received 23 May 2017; Received in revised form 12 July 2017; Accepted 8 August 2017

2. Materials and methods

2.1. Subjects

Ten patients (3 female, 7 male, mean age 41.5 years, range 19–65, two left-handed) undergoing surgical treatment for intractable epilepsy and implanted with chronic subdural electrodes participated in the experiment (Table 1). Seven patients were recorded in the Stanford School of Medicine, 2 in the California Pacific Medical Center (CPMC), and 1 in the UCSF Medical Center.

Electrode arrays were implanted on the right hemisphere for 8 of the patients, and on the left hemisphere for 2, with electrode location determined solely by clinical needs. All subjects gave informed consent approved by the UC Berkeley Committee on Human Research and corresponding IRBs at the clinical recording sites.

2.2. Stimuli and tasks

Recording was conducted in the epilepsy ICU. Stimuli were presented on a laptop screen and responses captured on the laptop keyboard. Stimuli were grayscale images of either frontal human faces, man-made objects (round watch-faces, tools, vehicles, furniture, clothing items, musical instruments and household objects), or animals, presented on a square uniform gray background in the center of the screen and extending across approximately 5° of the visual field in each direction. Only the face and object categories were used in the main analyses. Additional categories including body parts and houses were excluded from analysis due to the paucity of exemplars. Face and object images did not significantly differ in terms of luminance (206.2 and 207 mean pixel intensity for faces and objects, respectively) or contrast defined as the standard deviation of pixel intensity (41.8 and 41.1 respectively). Each stimulus was presented for either 300, 600, 900, 1200 or 1500 ms for the first 3 subjects (Fig. 1A), or 300, 900 or 1500 ms for the last 7 subjects (Fig. 1B. the number of durations was reduced to allow the additional control task described below without excessively prolonging the experiment). All categories had the same probability of appearing with each duration. Inter-stimulus interval varied from 600 to 1200 ms with 150 ms steps, during which a fixation cross was presented. In the main experimental condition, subjects were instructed to fixate on the center of the screen, and respond with a button press whenever a clothing item was presented (10% of trials). The last 7 subjects performed an additional experimental task (dual attention-control task) where in addition to responding to images of clothing items, they were also required to respond to rare instances when an image of any category became blurry for the last 200 ms of its presentation (clothing targets and blurred image targets each accounted for 5% of the trials). The onset time for each image was registered alongside the ECOG data from a photodiode placed on the laptop screen, recording a white rectangle displayed at the same time as the image at the corner of the screen.

Table 1
ECoG patients.

Patient ID	Gender	Age	Handedness	Electrode Coverage	Number of Electrodes	Responsive Electrodes
S1	F	38	R	RH: Tem, Occ, Fro	64	32
S2	M	46	R	RH: Tem, Occ, Par	112	28
S3	M	41	L	RH: Tem, Occ, Par	118	44
S4	M	19	R	RH: Tem, Occ, Par, Fro	123	15
S5	F	42	R	RH: Tem, Occ, Par	107	25
S6	M	29	R	RH: Tem, Occ, Par, Fro	113	16
S7	M	47	L	LH: Tem, Occ, Fro	64	9
S8	F	65	R	RH: Tem, Par, Fro	116	39
S9	M	47	R	LH: Tem, Par	128	52
S10	F	36	R	RH: Tem, Occ, Par, Fro	122	32

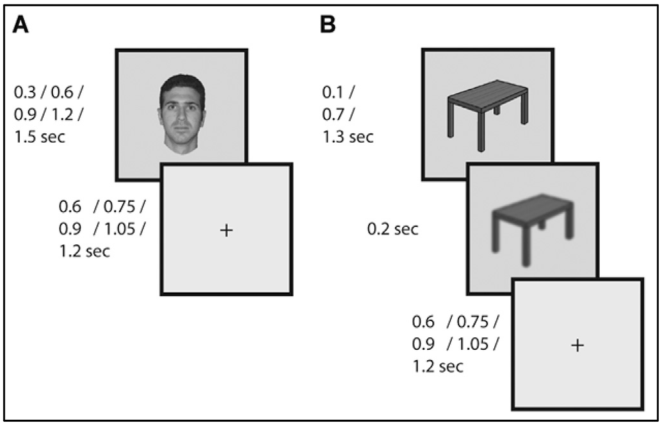


Fig. 1. Experimental paradigm. (**A**) Images were presented for 300, 900, or 1500 ms (for 3 subjects also 600 or 1200 ms) with variable inter-stimulus interval (ISI) during which a fixation cross was presented. Subjects responded with a button press to presentation of targets (clothing images; 10% of trials). (**B**) Dual-task control. This task was identical to the first except that subjects also had to respond to rare blurring of the image in the last 200 ms of its presentation.

2.3. ECoG acquisition and data processing

Each subject was implanted with subdural arrays containing 53–128 contact electrodes (AdTech Inc.). In total, 1067 electrodes were examined. Each electrode was 2.3 mm in diameter, with 5 or 10 mm spacing between electrodes within an array, arranged in 1-dimensional strips or 2-dimensional grids. Recordings were sampled at 1000 Hz (CPMC), 3051.76 Hz (Stanford, UCSF) or 1535.88 Hz (Stanford) and resampled to 1000 Hz offline (with the exception that the original 3051.76 Hz sampling rate was used for the results reported in Fig. 2). A high-pass filter was applied online to the signal at either 0.1 Hz or 0.5 Hz in different subjects. 159 electrodes manifesting ictal spikes or persistent noise were visually identified and removed from analysis, as were time intervals with excessive noise or ictal activity as determined by one of the authors (RTK). All remaining electrodes were re-referenced offline to the average potential of all non-rejected electrodes, separately for each subject. Unless otherwise indicated, all data processing and analysis were done using custom Matlab code (Mathworks, Natick, MA).

2.4. Electrode localization

Electrode locations were identified manually using BioImageSuite (www.bioimagesuite.org) on a post-operative Computed Tomography (CT) scan co-registered to a pre-operative MR scan using the FSL software package (Jenkinson et al., 2002; Jenkinson and Smith, 2001). Individual subjects' brain images were skull-stripped and segmented using FreeSurfer (http://surfer.nmr.mgh.harvard.edu). Localization errors driven by both co-registration error and anatomical mismatch between pre- and post-operative images were reduced using a custom procedure which uses a gradient descent algorithm to jointly minimize the squared distance between all electrodes within a single electrode array/strip and the cortical pial surface (see Dykstra et al., 2012 for a similar procedure). In contrast to methods which only attempt to correct individual electrodes'