NeuroImage 125 (2016) 280–290

Contents lists available at ScienceDirect

NeuroImage

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

Fast ventral stream neural activity enables rapid visual categorization

Maxime Cauchoix [a,b,c,*,1], Sébastien M. Crouzet [b,c,d,1], Denis Fize [b,c], Thomas Serre [d]

[a] Institute for Advanced Studies in Toulouse, France
[b] Centre de Recherche Cerveau et Cognition, Université Paul Sabatier, Université de Toulouse, Toulouse, France
[c] Faculté de Médecine de Purpan, CNRS, UMR 5549, Toulouse, France
[d] Cognitive, Linguistic and Psychological Sciences Department, Institute for Brain Sciences, Brown University, USA

ARTICLE INFO

Article history:
Received 23 April 2015
Accepted 7 October 2015
Available online 20 October 2015

Keywords:
Rapid categorization
Object recognition
Ventral stream
Monkey electrophysiology
Natural scenes

ABSTRACT

Primates can recognize objects embedded in complex natural scenes in a glimpse. Rapid categorization paradigms have been extensively used to study our core perceptual abilities when the visual system is forced to operate under strong time constraints. However, the neural underpinning of rapid categorization remains to be understood, and the incredible speed of sight has yet to be reconciled with modern ventral stream cortical theories of object recognition.

Here we recorded multichannel subdural electrocorticogram (ECoG) signals from intermediate areas (V4/PIT) of the ventral stream of the visual cortex while monkeys were actively engaged in a rapid animal/non-animal categorization task. A traditional event-related potential (ERP) analysis revealed short visual latencies (<50–70 ms) followed by a rapidly developing visual selectivity (within ~20–30 ms) for most electrodes. A multi-variate pattern analysis (MVPA) technique further confirmed that reliable animal/non-animal category information was possible from this initial ventral stream neural activity (within ~90–100 ms). Furthermore, this early category-selective neural activity was (a) unaffected by the presentation of a backward (pattern) mask, (b) generalized to novel (unfamiliar) stimuli and (c) co-varied with behavioral responses (both accuracy and reaction times). Despite the strong prevalence of task-related information on the neural signal, task-irrelevant visual information could still be decoded independently of monkey behavior. Monkey behavioral responses were also found to correlate significantly with human behavioral responses for the same set of stimuli.

Together, the present study establishes that rapid ventral stream neural activity induces a visually selective signal subsequently used to drive rapid visual categorization and that this visual strategy may be shared between human and non-human primates.

Introduction

The robust and accurate categorization of natural object categories is critical to survival, as it allows an animal to generalize many properties of an object from its category membership (Fabre-Thorpe, 2003; Fize et al., 2011; Rosch, 1975; Thompson and Oden, 2000; Zentall et al., 2008). Human and non-human primates excel at visual categorization: They can rapidly and reliably categorize objects embedded in complex natural visual scenes in a glimpse (see Fabre-Thorpe, 2011; Potter, 2012 for recent reviews).

It is well known that object recognition is possible for complex natural scenes viewed in rapid visual presentations that do not allow sufficient time for eye movements or shifts of attention (Biederman, 1972; Potter and Levy, 1969; Thorpe et al., 1996). The underlying visual representation remains relatively coarse as participants frequently fail to localize targets that are correctly detected in an image stream (Evans and Treisman, 2005). Studies using backward-masking (Bacon-Macé et al., 2005) and saccadic responses (Crouzet et al., 2010; Kirchner et al., 2009) have further demonstrated that recognition is possible under severe time constraints. While much is known about the psychological basis of rapid categorization, much less is known about the underlying neural processes and, in particular, the timing of the corresponding perceptual decisions.

Using human scalp electroencephalography (EEG), Thorpe et al. (1996) first demonstrated that a category-selective signal can be isolated from frontal electrodes shortly after a stimulus is flashed (within ~150 ms post stimulus onset). Previous work using intra-cranial recordings has shown that it is possible to decode object category information from the ventral stream of the visual cortex very rapidly (within ~100 ms post stimulus onset) in both humans (Liu et al., 2009) and monkeys (Hung et al., 2005; Kreiman et al., 2006; Vogels, 1999a). However, this work either used a passive viewing paradigm (Hung et al., 2005; Kreiman et al., 2006; Liu et al., 2009) or involved a relatively simple basic-level categorization task, such as trees vs. objects (Vogels, 1999a), and did not establish any link between ventral stream neural

* Corresponding author at: Institute for Advanced Studies in Toulouse, 21 allée de Brienne, 31015 Toulouse Cedex 6, France.
E-mail address: maxime.cauchoix@iast.fr (M. Cauchoix).
[1] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.neuroimage.2015.10.012
1053-8119/

activity and (speeded) behavioral responses during rapid categorization. Indeed, previous monkey electrophysiology work has found little co-variation between reaction time and neural latencies in the inferotemporal cortex (DiCarlo and Maunsell, 2005; Eifuku et al., 2004; but see also Mruczek and Sheinberg, 2007).

Here, we recorded ECoG activity in monkeys actively engaged in a rapid natural scene categorization task with the aim to characterize the time-course of visual processing and establish a link between fast ventral stream neural activity and rapid behavioral responses.

Material and methods

Successful learning of the categorization task

Protocol

Two male rhesus macaques (M1 and M2, both aged 14) performed the experiment. The animals were restrained in a primate chair (Crist Instruments, GA), sitting 30 cm away from a 1024 × 768 tactile screen. Stimuli (Fig. 1a) were flashed centrally for 33 ms covering about 7° of visual angle on a black background, with a 1.5–3 s random inter-trial time interval between successive images.

In masked trials (presented in separate blocks), a pattern image corresponding to visual pink noise was presented 50 ms post stimulus onset for 33 ms (Fig. 1b). Brief presentations, in addition to masking on specific trials, prevented exploratory eye movements and constrained the time available for information uptake.

The two monkeys performed a natural scene rapid categorization task by releasing a button and touching the screen when they saw an animal in the stimulus presented (target) or keeping their hand on the button otherwise (distractor). A drop of fruit juice rewarded correct responses; on incorrect trials, the stimulus was re-displayed for 3 s, thus delaying the next reward and motivating the animal to answer as fast and accurately as possible. All procedures conformed to French and European standards concerning the use of experimental animals. Protocols were approved by the regional ethical committee.

Animal training

Initial training followed the procedure previously reported in Fabre-Thorpe et al. (1998): learning was gradual, starting with 10 images and progressively introducing new scenes everyday over a period of several weeks until both monkeys were performing well on the familiar set of stimuli. While the monkeys' motivation and level of reward were kept stable by randomly interleaving familiar and new stimuli, the recurrent introduction of new stimuli (usually 10–20%) forced the animals to learn to generalize to novel stimuli rather than to rely on stimulus memorization. Both monkeys were trained for intermittent periods on the animal/non-animal task since 2005 (Fabre-Thorpe et al., 1998; Fize et al., 2011; Macé et al., 2005).

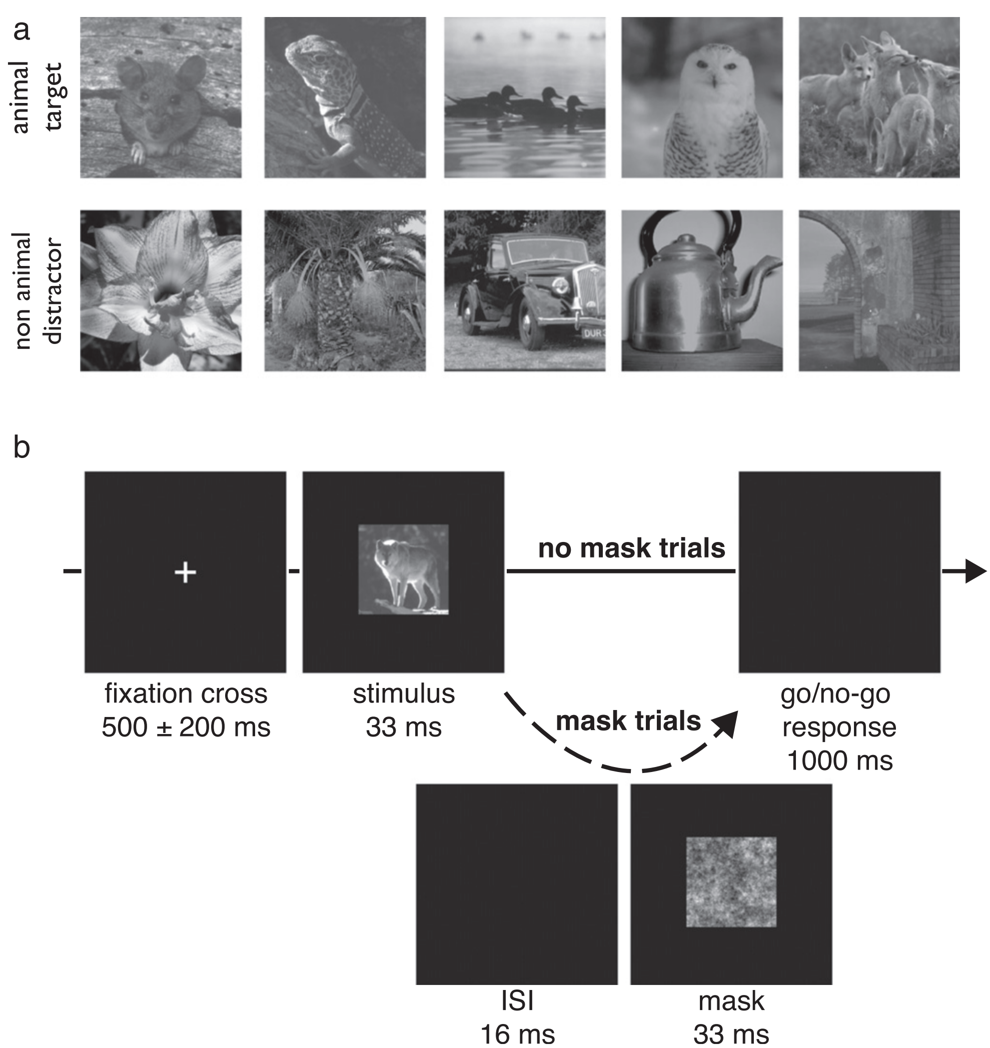

Fig. 1. Stimulus set and visual categorization task. a) The stimulus set consisted of natural gray-scale images with both animal targets (n = 340) and non-animal distractors (n = 340). b) Following the presentation of a fixation cross, an image was flashed for 33 ms. The monkeys indicated whether a target was present by releasing a button within 1 s following image presentation. On half the trials, a backward mask (1/f pink noise) was displayed for an additional 33 ms following a 16-ms blank screen (50 ms SOA).