Brain and Cognition 84 (2014) 34–43

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Incongruent object/context relationships in visual scenes: Where are they processed in the brain?

Florence Rémy [a,b,*], Nathalie Vayssière [a,b], Delphine Pins [c,d], Muriel Boucart [c,d], Michèle Fabre-Thorpe [a,b]

[a] Université de Toulouse, UPS, Centre de Recherche Cerveau et Cognition, France
[b] CNRS, CerCo, Toulouse, France
[c] Université Lille Nord de France, UDSL, Laboratoire Neurosciences Fonctionnelles et Pathologies, CHU Lille, F-59000 Lille, France
[d] CNRS, F-59000 Lille, France

ARTICLE INFO

Article history:
Accepted 28 October 2013
Available online 23 November 2013

Keywords:
Natural scenes
Contextual associations
Congruence
fMRI

ABSTRACT

Rapid object visual categorization in briefly flashed natural scenes is influenced by the surrounding context. The neural correlates underlying reduced categorization performance in response to incongruent object/context associations remain unclear and were investigated in the present study using fMRI. Participants were instructed to categorize objects in briefly presented scenes (exposure duration = 100 ms). Half of the scenes consisted of objects pasted in an expected (congruent) context, whereas for the other half, objects were embedded in incongruent contexts. Object categorization was more accurate and faster in congruent relative to incongruent scenes. Moreover, we found that the two types of scenes elicited different patterns of cerebral activation. In particular, the processing of incongruent scenes induced increased activations in the parahippocampal cortex, as well as in the right frontal cortex. This higher activity may indicate additional neural processing of the novel (non experienced) contextual associations that were inherent to the incongruent scenes. Moreover, our results suggest that the locus of object categorization impairment due to contextual incongruence is in the right anterior parahippocampal cortex. Indeed in this region activity was correlated with the reaction time increase observed with incongruent scenes. Representations for associations between objects and their usual context of appearance might be encoded in the right anterior parahippocampal cortex.

1. Introduction

In our everyday environment, objects most often appear in typical visual settings that contain other related objects arranged in specific spatial configurations. This provides us with a rich set of semantic and spatial contextual associations that are implicitly learned through visual experience. Accordingly, the concept of *schemata* has been introduced to describe a high-level unified representation for contextual associations in a scene (Biederman, Mezzanotte, & Rabinowitz, 1982; Hock, Romanski, Galie, & Williams, 1978; Mandler & Johnson, 1976). During the processing of real-world scenes, *schemata* would be activated in a rapid and automatic manner. Following this latter hypothesis, object search and recognition are facilitated when objects are seen in an expected *congruent* context and at an expected position and size (Bar, 2004; Biederman et al., 1982; Chun, 2000; Ganis & Kutas, 2003; Oliva & Torralba, 2007; Palmer, 1975). Conversely, visual *incongruence* between object and its context of presentation (Davenport, 2007; Davenport & Potter, 2004; Fabre-Thorpe, 2011; Fize, Cauchoix, & Fabre-Thorpe, 2011; Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Kret & de Gelder, 2010; Mudrik, Lamy, & Deouell, 2010) or between pairs of objects (Green & Hummel, 2006; Gronau, Neta, & Bar, 2008; Kim & Biederman, 2010) results in decreased performance when processing object, context or global scene. This consistently reported drop of performance suggests that previously unseen, i.e. incongruent, contextual associations influence visual recognition, possibly because representations for these new associations or *schemata* have not been registered through experience. However, the neural correlates underlying this behavior in response to visual incongruence remain to establish. The present functional magnetic resonance imaging (fMRI) study aimed at determining which neural regions evidence differential activity during processing of semantically-incongruent vs. congruent associations between objects and their surrounding, and how activity in these regions contributes to decrease in object categorization performance.

Medial temporal areas previously involved in contextual binding (Chalfonte & Johnson, 1996; Goh et al., 2004; Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000) are candidate areas that may be

* Corresponding author. Address: CNRS, CerCo, UMR 5549, Pavillon Baudot, CHU Purpan, BP 25202, 31052 Toulouse Cedex, France. Fax: +33 5 62 74 61 17.
E-mail address: florence.remy@cerco.ups-tlse.fr (F. Rémy).

0278-2626/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandc.2013.10.008

sensitive to novel object/context associations in incongruent scenes. Through visual experience, objects are not encoded individually, but rather in relation to typical settings. It has been shown that these binding operations are made in an automatic way, without any instructions to relate the different elements in a scene (Cohen et al., 1999; Goh et al., 2004). Relational encoding of visual picture elements, such as a person in a house context (Henke, Weber, Kneifel, Wieser, & Buck, 1999), induces greater activity in right hippocampal and parahippocampal regions relative to separate encoding of the same elements. Using an fMRI-adaptation paradigm during passive viewing of natural scenes, the anterior parts of the right parahippocampal cortex (PHC) and of the right hippocampus were specifically activated when participants successively processed sequences of novel object/context associations compared to repeated ones (Goh et al., 2004). Consistently, increased activity in the anterior part of the PHC during retrieval of meaningless patterns has been related to previous encoding of these patterns in nonspatial configurations (Aminoff, Gronau, & Bar, 2007), which suggests that nonspatial familiar associations between items, or between an item and its context, may be encoded in this region. The hippocampus and anterior part of PHC may receive converging information from more posterior parahippocampal regions about the item features and the context in which items are encountered (Diana, Yonelinas, & Ranganath, 2007; Goh et al., 2004; Suzuki & Amaral, 2004). Therefore these medial temporal regions may show differential involvement when processing incongruent vs. congruent object/context associations. In particular, encoding of never-seen incongruent associations may reinforce relational binding operations, resulting in greater activity in medial temporal regions (Cohen et al., 1999).

In imaging studies that have looked at neural processing of semantically-incongruent contextual associations, changes in brain activity in response to incongruence have been investigated in high-level regions of the visual ventral pathway, the lateral occipital complex (LOC) and the parahippocampal place area (PPA) (Gronau et al., 2008; Jenkins, Yang, Goh, Hong, & Park, 2010; Kim & Biederman, 2010), as well as in the prefrontal cortex (Ganis & Kutas, 2003; Gronau et al., 2008; Mudrik et al., 2010). The former occipito-temporal regions are thought to process bottom-up perceptual information from early visual areas, while the latter frontal region may modulate this processing through top-down mechanisms (Bar & Ullman, 1996; Biederman, 1972; Gronau et al., 2008; Palmer, 1975).

Using line-drawings of interacting object pairs, Kim and Biederman (2010) have reported increased activity in the LOC when objects were semantically unrelated vs. related (e.g. a bird on an ear vs. a bird on a birdhouse), suggesting that the LOC may process interactions between objects. Consistently, Gronau et al. (2008) suggested that activity in the LOC may depend on contextual surroundings of objects. In their study, LOC activity differentiated target-object recognition when primed with semantically-related vs. –unrelated objects. Besides, the PPA may mediate contextual associations (Aminoff et al., 2007; Bar & Aminoff, 2003; Bar, Aminoff, & Ishai, 2008; Bar, Aminoff, & Schacter, 2008). Activations in the PPA are increased when participants are presented with isolated objects that have strong associations with a specific context (e.g. a cow or an oven) relative to objects with only weak contextual associations (e.g. a fly or an apple) (Bar & Aminoff, 2003). It was proposed that representations for contextual associations may be stored within the PPA and activated during processing of highly contextual objects. The concept of *context frames* (Bar & Ullman, 1996), which is reminiscent of the concept of *schemata*, was proposed to describe this long-term memory trace for contextual associations. Activation of representations stored in the PPA may be modulated through top-down mechanisms, as described in a model for object recognition in environmental scenes proposed by Bar and colleagues (Bar, 2004). With a rapid grasp of a scene, the processing of low spatial frequency information would allow access to a coarse object shape and activate, in the prefrontal cortex, a large set of all object representations fitting this shape. In parallel, the extraction of the context frame would restrict the activated representations only to possible objects that could be predicted in such context. Object recognition would be achieved with detailed bottom-up high spatial frequency visual information. Consistent with this model, differential activity in the prefrontal region for congruent and incongruent associations has been reported (Ganis & Kutas, 2003; Gronau et al., 2008; Mudrik et al., 2010), which may reflect activation of semantic knowledge on contextual associations. Using pictures of objects embedded in contexts, Mudrik et al. (2010) have shown that this contextual congruence effect in the prefrontal region occurs around 270 ms after scene presentation. However, decrease in object categorization performance due to contextual violations has been observed for motor responses as short as 270 ms post-stimulus, suggesting that scene incongruence may influence bottom-up perceptual object processing before top-down modulation on high-level visual areas takes place (Joubert et al., 2008).

In the present study, real-world scenes with semantically congruent and incongruent object/context associations were briefly (100 ms) presented, and participants were asked to categorize objects as fast as possible. Stimulus exposure duration was shorter than the duration used in previous imaging studies (Ganis & Kutas, 2003; Gronau et al., 2008; Jenkins et al., 2010; Kim & Biederman, 2010; Mudrik et al., 2010) in order to prevent eye movements and increase speed of response, which may mostly rely on bottom-up visual analysis (Fize et al., 2011; Joubert et al., 2008). Moreover, object categorization performance was recorded in each participant, for congruent and incongruent trials. We hypothesized that high-level regions of the ventral pathway would show differential activity for congruent vs. incongruent scenes. In particular, anterior ventral regions previously involved in object/context binding and processing of contextual associations (Aminoff et al., 2007; Goh et al., 2004) may specifically respond to scene incongruence. Moreover, a key objective in our study was to examine correlations between activity and behavioral performance in object categorization. Such correlation analysis was crucial in order to specifically point towards the locus of incongruence processing that would induce the behavioral cost consistently reported in previous studies (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007, 2008; Mudrik et al., 2010).

2. Material and methods

2.1. Participants

Fifteen healthy volunteers (8 females, 11 right-handed, mean age = 23.6 years, range 20–35 years) gave informed consent for their participation. All of them had normal or corrected-to-normal vision. The study was approved by the local Ethics Committee in Lille, France (CPP n°05/79 Nord-Ouest IV).

2.2. Stimuli

Stimuli were horizontal color scenes (768×512 pixels), all drawn from a large set of scenes created in our lab and previously used in psychophysical experiments (Joubert et al., 2008). All scenes contained a foreground object pasted in a background context. Contexts and objects were selected from commercial CD-ROM libraries (Corel Stock Photo Libraries for the contexts, Hemera Photo Objects for the objects). In order to avoid sharp edges when