Neuropsychologia 56 (2014) 447–458

Contents lists available at ScienceDirect

Neuropsychologia

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

Synchronous contextual irregularities affect early scene processing: Replication and extension

Liad Mudrik [a,b,*,1], Shani Shalgi [c], Dominique Lamy [a], Leon Y. Deouell [d]

[a] Department of Psychology, Tel Aviv University, PO Box 39040, Tel Aviv 69978, Israel
[b] Division of Biology, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
[c] Department of Cognitive Science, The Hebrew University of Jerusalem, Jerusalem 91905, Israel
[d] Department of Psychology and the Edmond and Lily Safra Center for brain sciences, The Hebrew University of Jerusalem, Jerusalem 91905, Israel

ARTICLE INFO

Article history:
Received 19 August 2013
Received in revised form
19 February 2014
Accepted 20 February 2014
Available online 2 March 2014

Keywords:
Visual scenes
Context effects
N300/N400
Event-related potentials
Incongruency
Matching models

ABSTRACT

Whether contextual regularities facilitate perceptual stages of scene processing is widely debated, and empirical evidence is still inconclusive. Specifically, it was recently suggested that contextual violations affect early processing of a scene only when the incongruent object and the scene are presented a-synchronously, creating expectations. We compared event-related potentials (ERPs) evoked by scenes that depicted a person performing an action using either a congruent or an incongruent object (e.g., a man shaving with a razor or with a fork) when scene and object were presented simultaneously. We also explored the role of attention in contextual processing by using a pre-cue to direct subjects' attention towards or away from the congruent/incongruent object. Subjects' task was to determine how many hands the person in the picture used in order to perform the action. We replicated our previous findings of frontocentral negativity for incongruent scenes that started ~210 ms post stimulus presentation, even earlier than previously found. Surprisingly, this incongruency ERP effect was negatively correlated with the reaction times cost on incongruent scenes. The results did not allow us to draw conclusions about the role of attention in detecting the regularity, due to a weak attention manipulation. By replicating the 200–300 ms incongruity effect with a new group of subjects at even earlier latencies than previously reported, the results strengthen the evidence for contextual processing during this time window even when simultaneous presentation of the scene and object prevent the formation of prior expectations. We discuss possible methodological limitations that may account for previous failures to find this an effect, and conclude that contextual information affects object model selection processes prior to full object identification, with semantic knowledge activation stages unfolding only later on.

1. Introduction

Biological organisms grasp and interpret visual scenes amazingly fast and effortlessly. How they overcome the formidable challenge of processing the enormous amount of details embedded in natural scenes is one of the greatest puzzles in the study of visual perception. Significant help in achieving this feat may come from the existence of contextual regularities: objects tend to co-appear in particular scenes, allowing for prior knowledge and expectations to narrow the range of probable interpretations, thereby rendering scene analysis easier. Indeed, when such expectations are violated (e.g., a whale showing up in the middle of a football stadium), scene processing is impeded (Biederman, Glass, & Stacy, 1973; Biederman, Rabinowitz, Glass, & Stacy, 1974; Friedman, 1979; Palmer, 1975; Rayner & Pollatsek, 1992), in terms of both speed (Bar & Ullman, 1996; Boyce & Pollatsek, 1992; Chun & Jiang, 1998; Davenport & Potter, 2004) and accuracy (e.g., Antes, Penland, & Metzger, 1981; Bar & Ullman, 1996; Boyce, Pollatsek, & Rayner, 1989).

Evaluation of contextual relations during perceptual stages of scene processing, prior to full identification, would allow maximal benefits and facilitate the ongoing processing of both the scene and its constituents. However, whether contextual evaluation indeed facilitates perception remains controversial. Some theoretical models deny any contextual processing prior to scene and objects identification, and claim that it can occur only at later, post-perceptual stages (i.e., Functional isolation models; De Graef, 1992; Hamm, Johnson, & Kirk, 2002; Hollingworth & Henderson, 1998, 1999), at least 300 ms after the scene has been presented (Ganis & Kutas, 2003). Others posit that contextual processing occurs earlier and influences object identification processes. Such

* Corresponding author at: Department of Psychology, Tel Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, ISRAEL.
E-mail address: liadmu@gmail.com (L. Mudrik).
[1] Current address: Division of Biology, California Institute of Technology, 1200 E California Blvd, Pasadena, CA, 91125, USA.

http://dx.doi.org/10.1016/j.neuropsychologia.2014.02.020
0028-3932

influence can take place when object processing commences, during the first 200 ms of scene processing (when initial differences between object catoegories are observed; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001), at the stage of attentional feature selection (i.e., Perceptual schema models; Antes et al., 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce et al., 1989). Alternatively, contextual processing were suggested to facilitate object identification at somewhat later stages (i.e., Object model selection or Matching models; Bar, 2004; Bar & Aminoff, 2003; Bar & Ullman, 1996; Kosslyn, 1994), between 200 ms and 300 ms post stimulus presentation (Schendan & Kutas, 2002, 2003; Schendan & Maher, 2008), when pre-activated scene-congruent object representations are being matched with upcoming visual information about the scene's constituents.

Relevant empirical evidence has been inconclusive. In particular, a recent series of ERP studies yielded conflicting results. Effects of contextual processing of congruent and incongruent scenes in the 200–300 ms time window, prior to full object identification, were found in three previous studies (Mudrik et al., 2010; Sun, Simon-Dack, Gordon, & Teder, 2011; Võ & Wolfe, 2013). For instance, we (Mudrik et al., 2010) reported an anterior negativity related to incongruent scenes that started around 270 ms post scene presentation, and lasted about 330 ms. This negativity was followed by a later broadly distributed negativity between 650 ms and 850 ms, possibly related to late processes of semantic evaluation and response preparation. The earlier negativity we found was interpreted as a combination of the N300 (McPherson & Holcomb, 1999; Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008) and N400 (Kutas & Hillyard, 1980a, 1980b) components. N300, which occurs 200–300 after stimulus onset, was previously suggested to reflect processes that lead to object identification (Ganis & Kutas, 2003). Accordingly, its amplitude is higher for unidentified than for identified objects (Folstein, Van Petten, & Rose, 2008; Holcomb & McPherson, 1994; Schendan & Kutas, 2002), and is modulated by identification difficulty (Doniger et al. 2000; Henson, Rylands, Ross, Vuilleumeir, & Rugg, 2004; Holcomb & McPherson, 1994). Thus, this finding was taken as evidence supporting matching models of contextual processing, which postulate that scenes activate schemas that reduce the amount of perceptual evidence needed to match a particular schema-congruent object with its representation.

However, in a previous study Ganis and Kutas (2003) failed to observe such an N300 effect and reported only a later negativity, namely the "N390 congruency effect", that emerged in the 300–500 ms time window, similarly to the N400 component, albeit with a more frontal distribution. The absence of any earlier differences in either the 200–300 or the 0–200 time-windows was interpreted as ruling out contextual influences on perceptual stages of scene processing, thereby supporting functional isolation models.

The discrepancy between these two findings is especially surprising because it should have been easier to observe earlier differences using Ganis and Kutas' a-synchronous paradigm than using ours. Ganis and Kutas first presented a pre-cue, followed by the scene, and only then added the critical object at the cued location. Thus, subjects had time to form expectations regarding probable objects that matched the scene. By contrast, we presented the scene and object simultaneously in order to prevent subjects from forming prior expectations (see Mudrik et al., 2010 for a detailed argumentation). Nevertheless, we found the early N300 described above.

This discrepancy widens when considering a more recent ERP study (Demiral et al., 2012), which manipulated objects' spatial congruency (e.g., a bus was presented in the sky vs. on the road), rather than their semantic congruency (we use the term "semantic congruency" following Biederman (1981), to denote contextual violations in which the probability of an object to occur in a scene is manipulated. Accordingly, such contextual violations rest on previous knowledge about the co-occurrence of objects and scenes). Demiral et al. conducted two experiments: the first followed Ganis and Kutas' (2003) sequential design, that is, a pre-cue was presented first, followed by the scene, and only then the spatially congruent/incongruent object was presented. Conversely, the second experiment followed our simultaneous design (Mudrik et al., 2010). N300 effects arose in the sequential condition but not in the simultaneous condition, and the N400 component was smaller in the simultaneous than in the sequential condition. The authors concluded that earlier contextual influences are contingent on previously formed expectations about the forthcoming object, in sharp contrast to Mudrik et al.'s (2010) conclusions. Thus, under the premise that direct replications are the best way to establish the reliability of results (Cumming, 2014; Pashler & Harris, 2012), the first aim of our study was to provide a replication of the N300 congruity effects in a new group of subjects, and using more trials to obtain sensitivity to even earlier effects.

The second aim of this study was to examine the role of attention in contextual processing: is focused attention on the critical object necessary for detecting that it is incongruent with its context, or can such detection be performed without focused attention, possibly leading to attention being drawn to the critical object? Loftus and Mackworth's (1978) model of scene perception (see also Underwood, Templeman, Lamming, & Foulsham, 2008) proposed that low-level preattentive extraction of a scene's gist occurs before complete identification of the objects that compose it. Then, partial recognition of an unattended or non-fixated object may be sufficient to determine that it violates the gist of the scene and requires further inspection. Only at that stage does attention come into play, and it triggers an eye movement to the location of the incongruent object. In other words, the incongruent object is labeled as such before it is attended (Underwood et al., 2008). In line with this suggestion, several studies reported object categorization (Evans & Treisman, 2005; Kirchner & Thorpe, 2006; Li, VanRullen, Koch, & Perona, 2002; Potter, Staub, & O'Connor, 2004; Thorpe et al., 1996) as well as contextual processing (Brockmole & Henderson, 2006; Chun & Jiang, 1998, 1999; Hidalgo-Sotelo, Oliva, & Torralba, 2005; Oliva, Wolfe, & Arsenio, 2004), during dual tasks or with very short stimuli exposures, that seem to take place outside the focus of attention, or with very little attentional resources.

However, whether the *semantic relationship* that links an object to its context can also be processed in the absence of attention remains under debate. While several eye fixation studies reported earlier fixations on incongruent than on congruent objects (Friedman, 1979; Loftus & Mackworth, 1978; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006), others observed only *prolonged* but not *earlier* fixations on incongruent objects (De Graef, Christiaens, & Dydewalle, 1990; Henderson, Pollatsek, & Rayner, 1989; Henderson, Weeks, & Hollingworth, 1999; Võ & Henderson, 2009, 2011), suggesting that attention is engaged by incongruent objects, but not drawn to them. Using binocular rivalry (for review, see Logothetis, Leopold, & Sheinberg, 1996), we found support for the latter view (Mudrik, Deouell, & Lamy, 2011).

To examine the role of spatial attention, in the current study we used exogenous cues (Posner, 1980) to direct subjects' attention towards or away from the location of a critical congruent/incongruent object and measured the effects of this manipulation on the electrophysiological markers of congruency processing (i.e., the N300/N400 component). We reasoned that if attention is needed for congruency processing, N300/N400 should be found for attended but not for unattended objects, and larger behavioral incongruency effects should be observed with attended than with unattended objects.

In summary, the aim of the current study was twofold: (a) to replicate the N300 effects found in our previous ERP study using a simultaneous object-scene presentation (Mudrik et al., 2010) in a new group of subjects, and thereby to provide critical support for contextual effects prior to full object identification in the face of conflicting data (Demiral et al., 2012) and (b) to directly manipulate attention in order to examine its influence on the