Neuropsychologia 50 (2012) 1271–1285

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing

Şükrü Barış Demiral[a,*], George L. Malcolm[b], John M. Henderson[c]

[a] Department of Psychiatry, School of Medicine, Washington University, Saint Louis, MO, USA
[b] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK
[c] Department of Psychology and McCausland Center for Brain Imaging, University of South Carolina, Columbia, SC, USA

ARTICLE INFO

Article history:
Received 17 August 2011
Received in revised form 25 January 2012
Accepted 18 February 2012
Available online 27 February 2012

Keywords:
Scene–object identification
Spatial congruency
N300
N400

ABSTRACT

Object processing is affected by the gist of the scene within which it is embedded. Previous ERP research has suggested that manipulating the semantic congruency between an object and the surrounding scene affects the high level (semantic) representation of that object emerging after the presentation of the scene (Ganis & Kutas, 2003). In two ERP experiments, we investigated whether there would be a similar electrophysiological response when *spatial congruency* of an object in a scene was manipulated while the semantic congruency remained the same. Apart from the location of the object, all other object features were congruent with the scene (e.g., in a bedroom scene, either a painting or a cat appeared on the wall). In the first experiment, participants were shown a location cue and then a scene image for 300 ms, after which an object image appeared on the cued location for 300 ms. Spatially incongruent objects elicited a stronger centro-frontal N300–N400 effect in the 275–500 ms window relative to the spatially congruent objects. We also found early ERP effects, dominant on the left hemisphere electrodes. Strikingly, LORETA analysis revealed that these activations were mainly located in the superior and middle temporal gyrus of the right hemisphere. In the second experiment, we used a paradigm similar to Mudrik, Lamy, and Deouell (2010). The scene and the object were presented together for 300 ms after the location cue. This time, we did not observe either an early or the pronounced N300–N400 effect. In contrast to Experiment 1, LORETA analysis on the N400 time-window revealed that the generators of these weak ERP effects were mainly located in the temporal lobe of the left hemisphere. Our results suggest that, when the scene is presented before the object, top-down spatial encoding processes are initiated and the right superior temporal gyrus is activated, as previously suggested (Ellison, Schindler, Pattison, & Milner, 2004). Mismatch between the actual object features and the spatially driven top-down structural and functional features could lead to the early effect, and then to the N300–N400 effect. In contrast, when the scene is not presented before the object, the spatial encoding could not happen early and strong enough to initiate spatial object-integration effects. Our results indicate that spatial information is an early and essential part in scene–object integration, and it primes structural as well as semantic features of an object.

1. Introduction

Humans constantly need to identify objects in order to properly interact with their environment. For example, when standing at the bathroom sink to brush your teeth, you need to recognise particular objects in order to facilitate the completion of your goal (e.g., toothbrush, toothpaste, faucet, etc.).

While the majority of research into spatial and semantic object processing has tended to use object stimuli on blank backgrounds (Gronau, Neta, & Bar, 2008), an important issue in object recognition research is that objects are typically embedded in visually complex scenes (Biederman, Mezzanotte, & Rabinowitz, 1982; Davenport & Potter, 2004; Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998; Mudrik, Lamy, & Deouell, 2010). Research on this topic has mainly focused on the effects of semantic congruency between an object and its environment. Semantic congruency refers to the probability of an object occurring in a particular scene. Prior exposure to a scene increases our expectations concerning objects likely to appear. This has been demonstrated in a variety of paradigms in which processing is facilitated when the target object is semantically congruent with the scene, while processing is inhibited when the object is semantically incongruent with the scene (Biederman et al., 1982; Castelhano & Henderson, 2008; Ganis & Kutas, 2003; Henderson, Weeks, & Hollingworth,

---

* Corresponding author at: Department of Psychiatry, School of Medicine, Washington University, 660 S. Euclid Avenue, Box: 8134, Saint Louis, MO 63110, USA. Tel.: +1 314 7471603.
E-mail address: demiralb@psychiatry.wustl.edu (Ş.B. Demiral).

0028-3932/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2012.02.011

1999; Hollingworth & Henderson, 1998; Loftus & Mackworth, 1978; Palmer, 1977; Võ & Henderson, 2009). For example, in a farm-yard scene, a tractor would have high-semantic congruency while an octopus would have low-semantic congruency; as such, the octopus would need a longer time to process.

An important electrophysiological study has suggested that the effect of semantic congruency occurs after the perceptual analysis and structural matching phases of object recognition, during the epoch in which semantic knowledge of the object is activated (Ganis & Kutas, 2003). This appears to be manifested in a negative wave beginning around 300 ms and peaking around what has been termed the N390 effect with a slight central distribution. This N390 effect appears to be similar in nature to the N400 effect, which has been found in participants viewing film clips in which actors used either environmentally congruent or incongruent tools (e.g., a man in a bathroom shaving with a razor or a rolling pin) (Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008; Sitnikova, Kuperberg, & Holcomb, 2003). Pictures that are semantically unrelated to a previous picture elicited a similar effect (Barrett & Rugg, 1990; Hamm, Johnson, & Kirk, 2002; McPherson & Holcomb, 1999). Semantically unrelated primes for the line drawings were also found to elicit an N400 effect, but topographically shifted towards the left hemisphere electrodes (Holcomb & McPherson, 1994).

Detection of incongruent items in visual–semantic paradigms was found to elicit not only an N400 effect, but in some cases, also elicit another negative component starting to peak a bit earlier, with a more frontal distribution, namely the N300 component (Barrett & Rugg, 1990; Hamm et al., 2002; McPherson & Holcomb, 1999; Sitnikova et al., 2008; West & Holcomb, 2002). This component is thought to be the index of structural and categorical mismatches generated by the visual system before the semantic and associative integration takes place (see pp. 1340 and 1347, Hamm et al., 2002, for the structural category violations indexed by the N300 component). Thus the earlier N300-like components emerging before the semantic N400 effect can be taken as indexes of structural and categorical mismatches.

In their recent study, Mudrik et al. (2010) used a semantic congruency paradigm where the object and the context were presented together. The authors reported an early effect of congruency starting around 270 ms in a small number of electrodes in the frontal regions of the left-hemisphere, spreading to a larger number of electrodes in the centro-frontal regions before 400 ms. This study is remarkable in that it has shown that the N300/N400 component can emerge even when the context and the object are presented together, and that the congruency effect can emerge as early as 180 ms.

The N400 component has also been extensively investigated in the psycholinguistic community. Similar to the visual context-target semantic mismatches, during sentence reading tasks, when semantically driven expectations of an attended word within a sentence do not hold, the N400 effect emerged (Ganis & Kutas, 2003; Ganis, Kutas, & Sereno, 1996; Kutas & Federmeier, 2000; Kutas & Hillyard, 1980, 1984). Since the psycholinguistic N400 and visual N300 and N400 effects have peak latencies closer to each other, and have the same polarities (but show some slight differences in their topographies), it is helpful to consider the functional explanation of the psycholinguistic N400 to understand the functional aspects of the visual N300 and N400.

There are two main views about the N400 component in psycholinguistics. The integrationist view of N400 suggests that the N400 effect is the reflection of the integration difficulty of the critical word within the working context (Brown & Hagoort, 1993; Hagoort, 2008; Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Hauk & Pulvermuller, 1888; Kutas & Hillyard, 1980; Osterhout & Holcomb, 1992; Sereno, Rayner, & Posner, 1998). In contrast, the lexicalist view claims that the N400 effect reflects facilitated activation of features of the long-term memory representations that are associated with a lexical item. N400 is shown to be sensitive to the lexical properties of the stimulus (e.g., frequency) where it can be manipulated by lexical priming effects as well as other lexically oriented effects (i.e., lexical neighbourhood) (Allen, Badecker, & Osterhout, 2003; Federmeier & Kutas, 1999; Grossi, 2006; Kiefer, 2002; Rugg, 1985; Van Petten & Kutas, 1990; Van Petten & Luka, 2006). Both the lexicalist and the integrationist accounts can exist together, where the context can make lexical access and the integration of the lexical item easier (Lau, Phillips, & Poeppel, 2008). Thus, observing a N300–N400 component indicates the cost of activating and integrating the unexpected visual target, with the previous context in both structural and semantic dimensions. Due to the low frequency (i.e., cloze-probability) of the target item in the context, and also due to competing activation of other contextually relevant and more expected items, the activation and the integration of the unexpected target becomes harder.

However, the Ganis and Kutas (2003) study has only examined one aspect of a scene's influence: semantic congruency. This ignores a second aspect of scenes: that scenes are "arranged in a spatially licensed manner" (p. 244, Henderson & Hollingworth, 1999) with objects occurring in highly predictable regions. Objects fit into locations where their functions should be active, appropriate and useful. Thus, it is currently not very well known whether spatial congruency and semantic congruency would be treated similarly in natural scenes.

Behavioral research suggests that manipulating the spatial congruency[1] of the object, while maintaining its semantic congruency with the scene, inhibits object processing (Biederman et al., 1982; Castelhano & Heaven, 2011; Davenport & Potter, 2004; Malcolm & Henderson, 2010; Võ & Henderson, 2009). Using the previous example, the tractor is more semantically related to the farm yard scene than the octopus. However, a tractor in a farm yard scene located on the roof of a barn will demonstrate spatial incongruence. In this case, it will take longer to fully process the tractor. In fact, Võ and Henderson (2009) found in an eye-tracking study that during scene viewing, the first pass gaze duration on a congruency-manipulated object increased to similar levels regardless of whether the incongruence was semantic or spatial in nature.

Given the similarly inhibitive effect of semantic and spatial congruency manipulations, it is tempting to assume that spatial congruency affects object processing in the same manner as semantic congruency. However, as Võ and Henderson (2009) used only behavioral measures (eye movements and gaze duration), it is difficult to infer the time course and nature of spatial congruency's effect on object processing. Therefore, further research is required to verify the spatial and temporal pattern of processing in the visual system responsible for the spatial incongruity effect.

A recent fMRI study by Gronau et al. (2008) investigated the interaction between the spatial and semantic congruency between the prime and the target objects. In this study, authors used objects on a blank background, where the prime object was followed by the target object. In the initial presentation of the prime-target pairs, reaction times were slower for the semantically unrelated pairs. The effect of spatial congruency appeared only in the fMRI response (BOLD, blood-oxygen-level-dependent response). Depending on this finding the authors argued that the spatial congruency effects might have emerged later in the processing, after the semantic processing took effect. Since the target location was not cued, and the prime objects were presented on a blank background, it is possible that the participants initially depended on the semantic processing more than spatial processing

---

[1] Sometimes known as syntactic congruency; see Võ and Henderson (2009).

in this paradigm. In line with studies reported in this paper, we cued the target object locations in scene contexts.

In order to study the time course of spatial congruency's effect on object identification in the present study, we used a brain measure with a high-temporal resolution: event-related brain potentials, or ERP. ERP allows real-time recording with high temporal resolution of the underlying processes. Similarly, the recorded spatial patterns of processing across the scalp will allow for insight into the systems involved.

We conducted two experiments. Our paradigm in the first experiment was similar to that of Ganis and Kutas (2003) in that participants were given a preview of the scene in order to generate expectancies about the possible class and attributes of an object. In this experiment, the target objects were always semantically congruent with the scene, however, half of the trials (incongruent trials) occurred in a spatially incongruent location. In the second experiment, we used a paradigm similar to Mudrik et al. (2010); the scene and the object were presented together.

In the first experiment, if the effect of spatial congruency on object processing occurs at the level of semantic analysis, we should find, as did Ganis and Kutas (2003), a modulation of a N400-like response, beginning at 300 ms and peaking around 390 ms. If this involves activation in similar topographical regions, we should see evidence of activation around this time point on the centro-frontal electrodes.

Alternatively, if object processing is affected by spatial congruency, not as a result of effects at the level of semantic analysis, but in the earlier phases of processing due to early perceptual selection, then there are conceivably two other epochs during which spatial congruency could influence object processing.

The first epoch is the earliest epoch, where *perceptual encoding* takes place. That is, the preview of the object's location within a scene might facilitate the perceptual analysis of the object (Biederman et al., 1982). The ERP correlate for this effect on object processing would be found in early components indexing perceptual processes, onsetting within the first 150 ms of the object's appearance. Such early ERP components have been shown to be affected by (i) perceptual attributes of the stimulus, such as orientation (Kenemans, Kok, & Smulders, 1993), spatial frequency (Kenemans, Baas, Mangun, Lijffijt, & Verbaten, 2000), spatial location (Clark, Fan, & Hillyard, 1994), and non-spatial features such as color (Anllo-Vento, Luck, & Hillyard, 1998), (ii) task manipulations that mainly affect perceptual encoding of the stimulus by modulated attention, specially reflecting on the P1 and N1 components onsetting between 80 and 130 ms post stimulus (Heinze et al., 1994). Thus, in Experiment 1, we tested whether scene context with the location cue could modulate the low level perceptual encoding and the expectation of a set of object attributes. If the expected perceptual attributes conflict with the target attributes, this may create a perceptual conflict and lead to early integration problems.

The second possible epoch falls somewhere between the early perceptual stage and the semantic processing stage. The spatial congruency of an object with its background may affect the process of matching a structural description of an unidentified object with object representations stored in long-term memory (Henderson et al., 1999; Hollingworth & Henderson, 1998). If an effect of spatial congruency was found here, it would suggest that prior knowledge of the location of the object within a scene influences the process of comparing structural object descriptions with expected object types (Bar & Ullman, 1996; Friedman, 1979; Friedman & Liebelt, 1981; Kosslyn, 1996; Palmer, 1975; Ullman, 1996). The estimated ERP correlate of structural description matching can be found in previous object processing research, where unidentified objects resulted in a frontal negative component starting around 200 ms and peaking around 350 ms, also named as N300 (see above), or structural description negativity, Nsd (Doninger et al., 2000; Holcomb & McPherson, 1994; Schendan & Kutas, 2002). Larger negativity here would reflect unsuccessful attempts to match the structural, categorical and spatial attributes to the target object. As structural description matching is divorced from the perceptual stages of object processing – that is, spatial congruency would have no effect on the perceptual analysis of the object – we would not find any activations relating to this process until after the initial perceptual processes are complete (~200 ms), but there would be a centro-frontal negativity around 300 ms.

Overall, our hypotheses can be summarized as follows. In Experiment 1, when the scene and the location cue are presented before the target object for a brief period of time (in our case this is 300 ms) (i) if spatial incongruency affects low level perceptual cueing then we should see activation around 100–200ms, whereas (ii) if the spatial incongruent object causes a structural mismatch then we should find a N300/Nsd response, and (iii) if the spatial congruency of the object affects its semantic analysis then we expect to find an N400 response.

If we do find evidence of N400 activation, we expect this to be caused by the cued location within the scene, rather than the global scene image, suggesting local Congruity. In this interpretation, N400 might be related to the difficulty of activating the semantic meaning of the target object, which is not one of the locally primed set of candidate objects expected for that particular location. If the integration account of the N400 is correct (see above), we expect N400 to emerge due to the integration difficulty of the target object with the local context.

We expect that the quality of the early low level perceptual processes will be influenced in Experiment 2. This will be due the fact that, in simultaneous processing of the object and the scene, the early mental representation of the scene has to be constructed either in parallel to or after the construction of the object (bottom-up). For instance, the successful evaluation of the spatial frequency of the object in the scene would need more time or more computations. Thus, the quality of the perceptual integration of the object with the scene could potentially be altered in the early time windows. These early processes, that we expect to see in Experiment 1, might be replaced by the semi-independent and parallel computations of the scene and the object in Experiment 2, and thus disappear in the ERP analysis.

While the task we used in our study is different from the Mudrik et al. (2010) study, the ERP correlates of the structural and semantic integration problems may potentially still be present in Experiment 2. On the other hand, if the very brief simultaneous presentation of the scene and object is not providing strong structural expectancies about the possible set of object features, or the task we use does not facilitate the quality of such expectancies about the set of objects fast enough, such structural mismatch effects will disappear in the Nsd/N300 time-window. Similarly, if the parallel or the bottom-up processing of the object and the scene has a dominant effect in the scene–object semantic integration, we expect that the N400 effect would attenuate in Experiment 2 as well. This will be mainly due to removing the scene preview and thereby removing the chance for the scene context to provide pro-active memory representations, which are used in the object identification process.

2. Experiment 1

As summarized above, we expected that early perceptual-spatial ERP effects would emerge due to top-down spatial-encoding (Heinze et al., 1994); In addition, if there is a post-perceptual scene–object integration difficulty for the spatially incongruent objects, we expect to find the N300–N400 complex to emerge for those incongruent objects, indicating structural and semantic integration cost. As suggested by the previous studies, an N300–N400 complex should be centro-frontally oriented (Sitnikova et al., 2003).

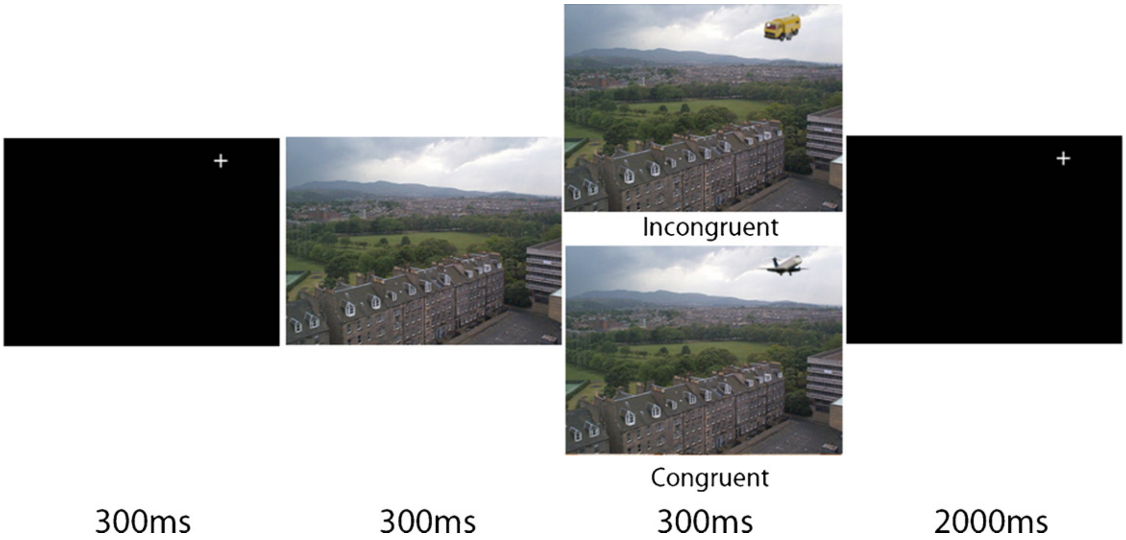


Fig. 1. Experimental procedure: Initial 300 ms fixation cross (location cue) was shown before the scene. After the presentation of the scene for 300 ms, the object popped up in the scene on the cued location for 300 ms. Following, the cross re-appeared on the same location for 2 s for the subjects to make a congruency judgment. Note that the target objects in the figure appear a bit larger than the original version for visualization purposes.

2.1. *Materials and methods*

2.1.1. *Participants*

Twenty-seven undergraduate and graduate students, 19 females and 8 males, aged 18–25, from the University of Edinburgh participated in the experiment. They were paid £6 for their participation. All the participants were native speakers of English. An experiment lasted around 1 h, including cap and electrode preparation. All subjects were reported to have normal or corrected to normal vision, and no neurological or psychiatric problems.

2.1.2. *Apparatus*

Forty color photographs of scenes were taken from Google Images and were scaled to 800 × 600 pixel resolution. A further 80 objects were taken from either the Hemera Images database (Hemera Technologies, Gatineau, Canada) or Google Images. Objects were modified and placed into scenes with Adobe Photoshop CS (Adobe, San Jose, CA).

In each scene the identity of the two target objects selected were constrained only by the rules that they should both be semantically congruent with a particular scene, and that the regions of the scene in which they have a high probability of being located should differ. For example, in a bedroom, a cat and a painting are both semantically congruent, but would have different high probability areas associated with them (the floor and wall, respectively; see Fig. 2) such that, if the objects' positions were rotated they would appear in less probable regions. This created forty base scenes, each with two potential target objects.

In a post hoc analysis, 10 participants scored the scenes to determine whether the targets were placed in high and low-probability regions as intended. None of these participants took part in Experiments 1 or 2. Participants were given a 5-point Likert-like scale and were asked to evaluate whether the target was positioned where they would expect to find it in the given scene; 5 for yes, definitely; 1 for no, not at all. Targets positioned in high-probability regions were judged to be in more expected regions of the scene than objects in a low probability regions $t(9) = 3.803$, $p = .004$. Participants also rated how likely target objects would occur in their respective scene, ignoring position, again using a 5 point Likert-like scale. A One-Sample Wilcoxon Signed Rank Test found mean rating for all the objects – whether they appeared in a spatially congruent or incongruent region – was significantly above the middle ranking score (3; $p = .005$). Together, this suggests that selected objects were considered semantically congruent with the scenes, but only spatially congruent when intended.

Only one object appeared in a trial, meaning that there was a total of 160 potential scene stimuli (2 objects × 2 locations × 40 scenes). The 160 scenes were divided into two lists of 80: participants only saw one of the two lists of scenes. In each list the two objects appeared in only one location of their respective scene (e.g., the painting and the cat appearing on the wall of the bedroom; the truck and the plane appearing in the sky as in Fig. 1) meaning that one object would be spatially congruent and the other incongruent. Participants viewing the scenes in the other list would see the same two objects appear over the course of the experiment, but this time in the other location (e.g., the cat and painting on the floor; the truck and the plane in the parking lot).

The locations of the objects were further counterbalanced by flipping scenes around a vertical axis, so that each participant saw half of the stimuli in the right-half and the other half in the left-half of the screen. Each participant saw 80 trials, split into two blocks of 40 trials.

2.1.3. *Procedure*

Participants were seated on a chair in front of the computer screen (90 cm away from the participant). The stimuli were presented on a screen by using E-prime software (Psychology Software Tools Inc., Pittsburgh, PA; http://www.pstnet.com/eprime). At the beginning of each trial a fixation cross (location cue) appeared on the screen for 300 ms. After 300 ms, the fixation cross disappeared and the scene was shown. The scene stayed on the screen for 300 ms, then an object appeared at the same location where the cross had been. The scene with the object remained on screen for another 300 ms. The scene and the object then disappeared, and the cross re-appeared for the duration of 2 s. Participants were instructed to press either the right or the left button to indicate whether the object was in an appropriate location or not. Participants were informed not to rush and to make their decisions properly. Buttons used for yes and no responses were counterbalanced among the subjects. Participants were told to try not to blink or move their eyes during the trial.

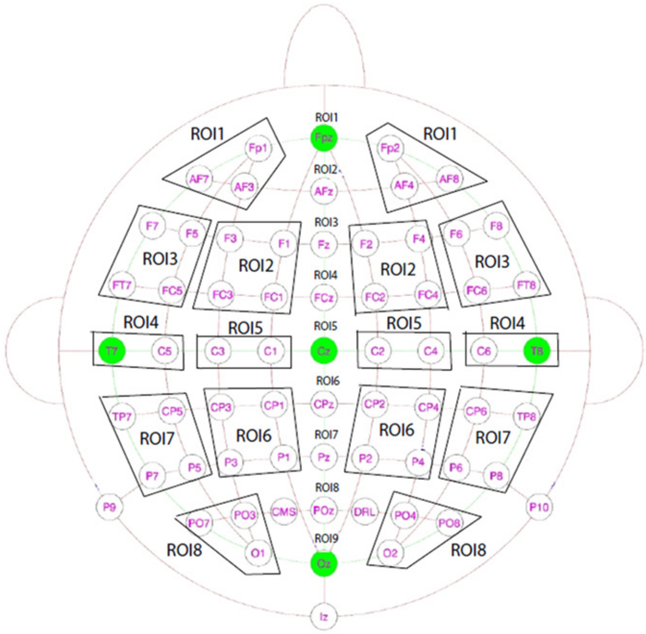


Fig. 2. Segmentation of the scalp electrodes into regions of interests (ROIs): right and left hemisphere electrodes were pooled into eight ROIs each, and the midline electrodes constituted nine ROIs.

2.1.4. EEG recording

EEG activity was recorded by a BioSemi ActiveTwo system (http://www.biosemi.com). We used an international 10–20 electrode cap configuration with 64 EEG channels. In addition to two mastoid electrodes (as a linked-reference), four EOG electrodes were placed on the horizontal cantus of the right (ROC) and the left eyes (HEOG), and vertically on the top (VEOG) and the bottom of the right eye (IOC) of the subjects in order to track eye blinks and horizontal eye movements for further artefact correction procedures. Sampling rate was 512 Hz.

2.1.5. EEG signal processing

We used a combination of algorithms provided by Brain Vision Analyzer2 software (http://www.brainproducts.com) and EEGLAB software (Delorme & Makeig, 2004). After the DC trends in the data were taken out, a Butterworth zero phase filter was used for high- and low-pass filtering the data with the half-amplitude cut off values of 0.1 Hz and 80 Hz respectively (12 dB/oct.) After this, we re-referenced the data to the mean of the mastoid electrodes. Epochs were selected between −450 ms and 1200 ms in which 0 ms was the time the object image was shown. Automatic ocular artefact correction with independent component analysis (ICA) was applied by using eye electrodes VEOG and HEOG (Jung et al., 2000; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997). The length of the baseline was 150 ms between −250 ms and −100 ms before the object image onset. This window optimally coincided with the temporal interval in which the subjects were viewing the scene image.

2.1.6. Performance assessment

In order to assess whether participants approached the experiment correctly, we conducted a behavioral assessment. We eliminated three subjects who showed incorrect responses (no- or false-responses) in more than 40% of the trials in any condition. The behavioral and ERP data for the remaining 24 subjects will be presented below.

2.1.7. EEG electrode regions of interest

We grouped electrodes as shown in Fig. 2 and constructed regions of interests (ROIs) in three different topographic fields (left, right and central). This led to eight Left- and eight Right-Hemisphere ROIs, and nine Central ROIs. ERP analysis was conducted within each topographic field for the constructed ROIs. Peak latency values of the selected electrodes are reported when informative.

2.1.8. Statistical analysis and the source localization of the ERP data

We conducted repeated-measures analysis of variance (ANOVA) as well as linear mixed effects modelling (LME) of the behavioral data and the ERPs. We will report repeated measures ANOVAs with the subjects ($F1$) and items ($F2$) selected as random variables. To protect against excessive Type I error due to violations of the assumption of equal variances of differences between the conditions of within-subject factors in the repeated-measures ANOVA, the Huynh–Feldt (Huynh & Feldt, 1976) correction was applied when evaluating effects with more than one degree of freedom. Bonferroni correction was used for multiple comparisons. We used linear mixed effects with the *lmer* function in R (Baayen, Davidson, & Bates, 2008). In LME, in order to minimize Type I error due to the non-random item variability, we also constructed models where fixed effects were introduced into the equations related to the random effects which accounted for non-homoscedasticity. In the statistical analysis of the accuracy data, we used family *binomial* in the lmer function.

We further analysed the data with Low Resolution Brain Electromagnetic Tomography Analysis, LORETA (Pascual-Marqui, Mitchel, & Lehmann, 1994). LORETA analysis is a functional neuroimaging method that detects the 3D distribution of the brain sources used under particular tasks from the scalp activity. Our goal was to examine the brain sources of the visual–spatial congruency effects, and then contrast such visuo-spatial mismatch effects with the previously reported lexico-semantic mismatch effects. This approach should provide fruitful outcomes for understanding any possible cortical dissociation between visuo-spatial and lexical stimuli generating N300 and N400 components. LORETA is a simple and straightforward analytical technique for approximating the underlying sources.

3. Results

3.1. Behavioral results

Rating speed in congruous trials, was numerically slower, $M = 932.4$ ms (SD = 266.6), than the incongruous trials, $M = 924.9$ ms (SD = 253.7), but this did not reach significance either in the repeated measures ANOVA by-subjects, $F1(1,23) = 1.6$, $p > .1$, or by-items, $F < 1$. Linear mixed effect modelling also confirmed this result, the log-odds of the incongruent conditions being insignificantly lower, $t = -0.6$. The absence of any effect in RTs is probably due to the fact that the task of the subject was not to respond as quickly as possible, but as correctly as possible. Accuracy rate was numerically smaller for the congruous conditions, 87.7% (SD = 32.9) than the incongruous conditions, 89.2% (SD = 31.1), but this small difference did not reach significance either in ANOVAs, $F1(1,23) = 1.3$, $p > .1$, $F2 < 1$, or LME with family chosen as binomial, $t = 0.85$. Subjects performed with equal accuracy in both conditions.

3.2. ERP results

Scalp activity of the correct trials of the 24 participants indicated that there were early low level perceptual effects starting around 50–250 ms, which were then followed by an N300/N400 effect. In the later time windows, there was a small negative going component (Figs. 3 and 4).

For the ERP analysis, we first prepared a series of small time-bins (25 ms time-bins) from 0 ms up to 1200 ms. We then conducted statistical analyses (repeated-measures by-subjects ANOVA) on each of these time-bins for each topographical location separately. The results are presented in Appendix A.1. By using this time-bin analysis, we constructed five time-windows to analyse the possible effects of Congruity: 50–150 ms, 150–225 ms, 275–500 ms, 500–850 ms, and 1025–1200 ms. The summary of ANOVA analyses is presented in Appendix A.2. We also present the lmer models tested, and the best fitting model for each ROI, in Appendix A.3.

3.2.1. Early effects

We found that there is an early positive peak (latency around 75 ms on the C3 electrode) for the incongruous condition, which was followed by a negative peak (latency around 166 ms on the C3 electrode) for the congruent conditions. These effects were dominant on the left hemisphere electrodes.

3.2.1.1. 50–150 ms time-window. ANOVAs: A repeated-measures ANOVA revealed that, in the 50–150 ms time-window, there was a significant main effect of spatial congruency only on the Left-ROIs for both by-subject, $F1(1,23) = 4.4$, $p < .05$, and by-item analysis, $F2(1,39) = 5.4$, $p < .05$. In the Midline-ROIs, the effect was significant for the by-subject analysis, $F1(1,23) = 4.5$, $p < .05$, but marginal for the by-item analysis, $F2(1,39) = 3.8$, $p = .058$.

LME analysis: A significant effect of Congruity emerged on ROI2, ROI5, ROI6, ROI7 of the Left-ROIs, and ROI7 of the Midline-ROIs.

3.2.1.2. 150–225 ms time-window. ANOVAs: A repeated-measures ANOVA revealed that, in the 150–225 ms time-window, there was a significant main effect of spatial congruency on the Left-ROIs, $F1(1,23) = 6.27$, $p < .05$, $F2(1,39) = 7.4$, $p < .01$. In the Right-ROIs, the effect was marginal, $F1(1,23) = 3.2$, $p = .09$, $F2(1, 39) = 3.5$, $p = .065$. In the Midline ROIs, the effect was significant, $F1(1,23) = 7.4$, $p < .01$, $F2(1,39) = 12.9$, $p = .01$. We also found a marginal interaction between ROI and Congruity in the Midline-ROIs, $F1(1,23) = 2.37$, $p = .056$, $F2(1,39) = 2.7$, $p = .035$, which revealed the main effect of Congruity on ROI5, ROI6, and ROI7, (corrected $p < .05$).

LME analysis: A significant effect of Congruity emerged on ROI2, ROI5, and ROI6 of the left-ROIs, and ROI5, ROI6, ROI7 and ROI8 of the Midline-ROIs.

LORETA analysis: This analysis revealed that the activity was mainly confined in the right hemisphere, superior temporal gyrus, around Brodmann area 22 ($x = 49$, $y = 3$, $z = 0$) (Fig. 5).

3.2.2. N300/N400 effects: 275–500 ms time window

ANOVAs: The incongruous conditions were more negative than the congruous conditions. An ANOVA revealed the main effect of spatial congruency for Left-ROIs, $F1(1,23) = 12.5$, $p < .01$, $F2(1,39) = 6.4$, $p < .05$, Right-ROIs, $F1(1,23) = 14.1$, $p < .01$, $F2(1,39) = 17.9$, $p < .001$, and Middle-ROIs $F1(1,23) = 12.4$, $p < .01$, $F2(1,39) = 9.75$, $p < .01$. We also found an interaction between ROI and Congruity for the Left-ROIs, $F1(1,23) = 4.6$, $p < .001$, $F2(1,39) = 3.5$, $p < .01$, and for the Middle ROIs, $F1(1,23) = 4.3$, $p < .01$,

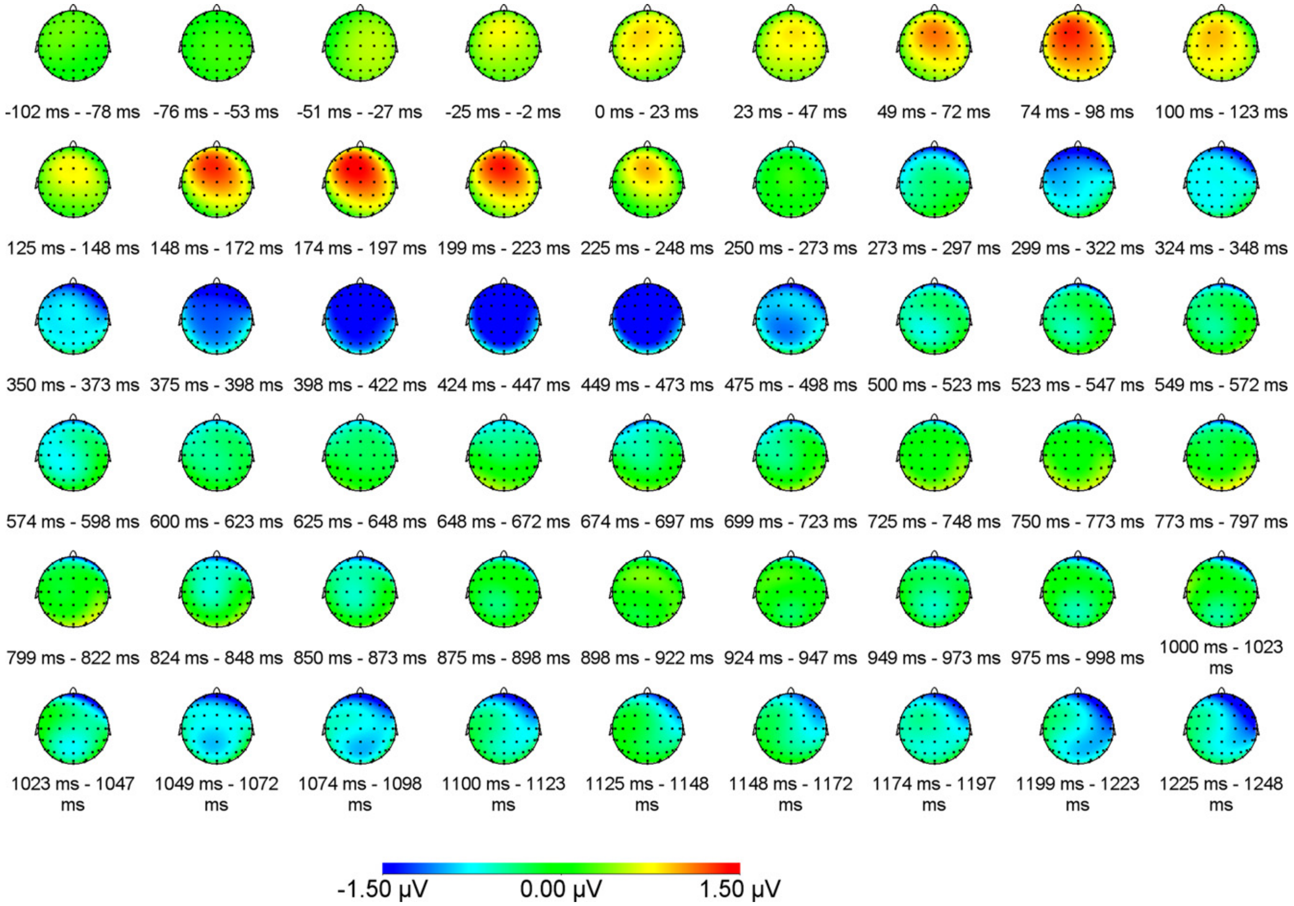


Fig. 3. Scalp maps of the difference between the Incongruent and Congruent conditions, shown for each 25 ms time-bin in Experiment 1. Blue color indicates the difference is negative. Red color indicates the difference is positive. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$F2(1,39) = 3.2$, $p < .05$. The effect of Congruity was significant on the centro-frontal Midline-ROIs, ROI1, ROI2, ROI3, ROI4 (corrected $p < .05$), and ROI5 of the Left-ROIs for both by-item and by-subject ANOVAS (corrected $p < .05$).

LME analysis: A significant effect of Congruity was found on ROI1, ROI2, ROI4, ROI5, and ROI6 of the Left-ROIs; ROI1, ROI2, ROI5, and ROI6 of the Right-ROIs, and ROI1, ROI2, ROI3, ROI4, ROI5, and ROI6 of the Midline-ROIs.

Peak latencies: Close inspection of the N300/N400 complex on the FCz and CPz electrodes indicated that the there was an early frontal N300 component overlapping with the later posterior N400 component. The earlier peak was more frontally oriented and reached its maximum at 313 ms on the FCz electrode (fronto-central electrode, see Fig. 3), and the second posterior peak reached its maximum on Pz (parieto-central electrode, see Fig. 3) at 410 ms.

LORETA analysis: LORETA analysis revealed that the activity was mainly confined again in the right hemisphere, similar to the early effects, but this time it was a bit inferior, extending to the middle temporal gyrus, Brodmann area 21, ($x = 53$, $y = 7$, $z = -18$) (Fig. 6).

3.2.3. *Late effects*

3.2.3.1. 500–850 ms time-window. ANOVAs: In the 500–850 ms time-window, there was an interaction between ROI and Congruity in the Left- and Midline-ROIs, observed significantly for the by-subject analysis, Left-ROIs; $F1(1,23) = 3.35$, $p = <.01$, and Middle-ROIs, $F1(1,23) = 3.1$, $p = <.05$, and marginally in the by-items analysis, Left-ROIs; $F2(1,39) = 1.83$, $p = .1$, and Middle-ROIs, $F2(1,39) = 2.4$, $p = .062$. None of the ROIS was found to be significant as revealed by the p-values falling above the Bonferroni adjusted alpha levels (corrected $p > .05$).

LME analysis: A main effect of Congruity was only observed on the ROI2 and ROI4 of the Midline-ROIs.

3.2.3.2. 1025–1200 ms time-window. ANOVAs: In the later time-window, the effect shifted towards the right hemisphere electrodes. There was a significant effect of Congruity on the Right-ROIs, $F1(1,23) = 6.37$, $p = <.05$, $F2(1,39) = 6.1$, $p < .05$.

LME analysis: In contrast to the ANOVA results, only one ROI showed a main effect of Congruity on the right hemisphere, namely ROI6, $t = -1.7$. This might be due to high level of variation between the individuals and items, which might have been accounted by the mixed model, but have been missed in the ANOVA analyses.

4. Discussion

When Ganis and Kutas (2003) manipulated an object's semantic Congruity with the surrounding scene's gist, they found an N400-like response, suggesting that a scene's gist does not affect object identification processes before associated semantic information is activated. The results of Experiment 1 similarly found an N400 component for the target objects which were spatially incongruent with their scene backgrounds. There was an early centro-frontal N300 component overlapping with the posterior N400 component. The earlier N300 peak reached its maximum at 313 ms on the FCz electrode (fronto-central electrode, see Fig. 3), and the second posterior

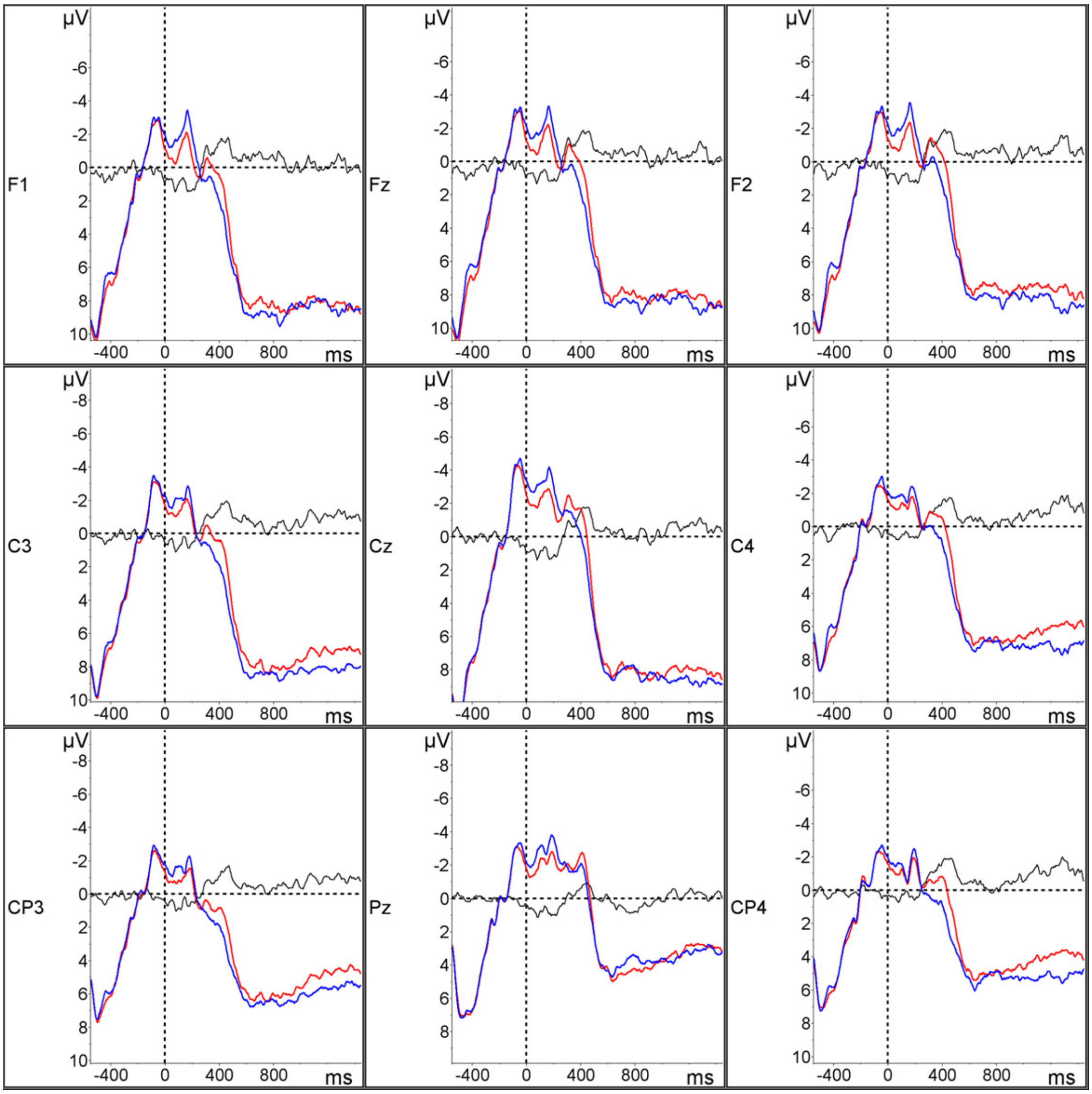


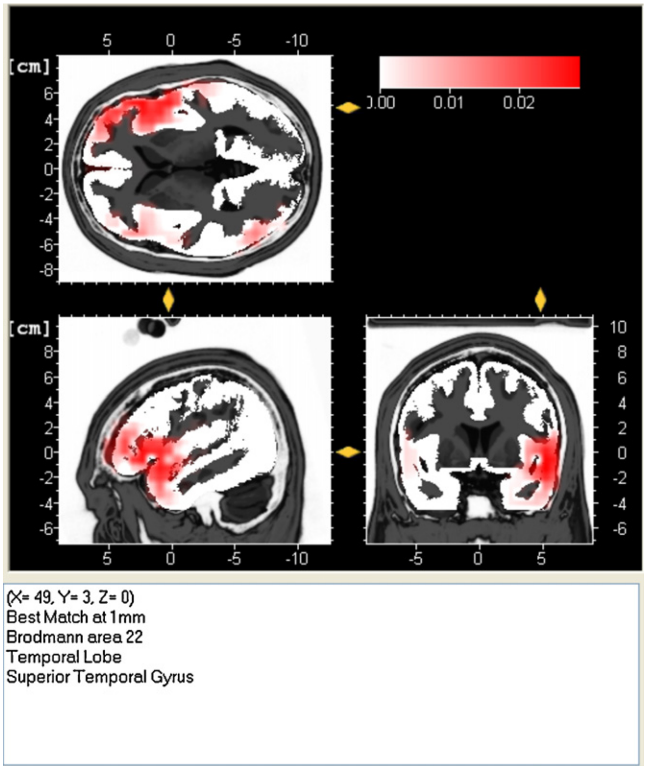
Fig. 4. ERP amplitude-time plots and the difference maps of Experiment 1: red line shows the incongruent conditions, and the blue line shows the congruent conditions. Black line shows the difference between the two. The mean of the mastoid electrodes is the reference. Base-line is selected as 150 ms duration between −250 ms and −100 ms before the object onset, overlapping with the viewing of the scene. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

N400 peak reached its maximum on Pz (parieto-central electrode, see Fig. 3) at 410 ms.

At first glance, these findings suggest that the effect of spatial congruency on object processing occurs first at the structural level (Nsd/N300) around 300 ms, and later at the semantic (N400) level, around 400 ms. In other words: the results suggest that once the object is recognised, its context – including its spatial position relative to the background environment – is analysed at structural/schematic and semantic levels. Note that the N300–N400 found in Experiment 1 is centro-frontally distributed, similar to Ganis and Kutas' (2003) findings.

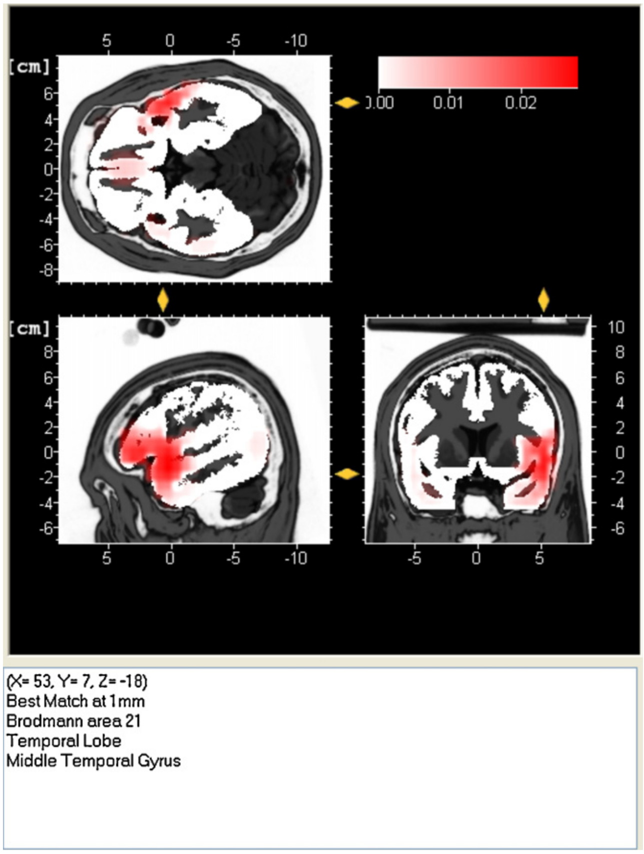
Interpreting the Ganis and Kutas (2003) study, Mudrik et al. (2010) noted that the context of the critical object was known prior to the object appearance in that study, so participants most likely formed a priori contextual expectations of a set of objects. Using an example from our experiment, when participants were cued to the sky above the buildings, there are only a handful of objects that could realistically fit there (e.g., a plane, a bird). When the truck appeared in the sky, rather than a plane or a bird, this might have violated the previously formed contextual expectations. Compared to other domains such as psycholinguistics, studies that investigated the N400 effect (Kutas, 1993, 1997; Kutas & Federmeier, 2000; Kutas & Hillyard, 1984) used linguistic context *preceding* the target item. Again, a set of expectations was generated prior to the target variable. In other instances, a picture or a non-word item has been used after the preceding context or word (Federmeier & Kutas, 1999, 2001, 2002). These studies identified N400 as the index of the difficulty of mapping the target item onto the activated set of semantic representations or expectations derived by cloze-probability of the target appearing in particular contexts. Thus, N400 may not be taken as a component of a semantic integration cost per se, but as a component of reflecting the cost of semantic "reorganization": when the unexpected target asks for an update and a modification of the declarative (and maybe also episodic) memory representations of the scene with an object appearing in its non-canonical position.

However, the N300–N400 complex was not the only statistically significant response that we found. The results of Experiment 1 also showed early effects. The early negative component emerged for the congruent objects as opposed to the incongruent objects,



(X= 49, Y= 3, Z= 0)
Best Match at 1mm
Brodmann area 22
Temporal Lobe
Superior Temporal Gyrus

Fig. 5. LORETA analysis of the early time windows, 50–225 ms in Experiment 1. LORETA is conducted on the difference wave between the incongruous and the congruous conditions.



(X= 53, Y= 7, Z= -18)
Best Match at 1mm
Brodmann area 21
Temporal Lobe
Middle Temporal Gyrus

Fig. 6. LORETA analysis for the N300/N400 time window, 275–500 ms in Experiment 1. LORETA is conducted on the difference wave between the incongruous and the congruous conditions.

peaking at 166 ms on the left centro-frontal electrode C3 (see Fig. 2). Other studies looking at object–scene semantic congruency did not find an early response that would have suggested an effect of perceptual analysis (Ganis & Kutas, 2003; Mudrik et al., 2010). Thus, at this point, we need to specify an important factor which led to the early effects in our experiment. First of all, the most prominent difference in our experiment was the task requirement. In our experiment, participants had to pay attention to the details of the local context *in relation to* the rest of the scene (i.e., spatial relation of the cued location with the other locations should be computed), rather than processing the overall context as in Ganis and Kutas study (i.e., understanding the overall context from a simple gist would be sufficient.). In contrast, Ganis and Kutas used the location cue just to inform the target object's location to prevent possible eye movements. Participants in that study could gain an idea about the context of the scene without paying attention to the spatial relations between the fixated and not-fixated scene parts.

The early negative effect found in Experiment 1 could support the idea that participants constructed local spatial features of the candidate objects, which can potentially fit into that location. Thus, combining this result with the LORETA analysis, the earlier negativity might index the activation of the spatially encoded low-level object features, probably due to the need for constructing the spatial relations between the object and the scene.

Our findings indicate that spatial congruency is a very important factor in scene–object integration. The cued location within the scene is fed back to the object recognition process, activating a set of features prominent for the objects fitting in those locations, leading to the activation of brain areas around the superior temporal gyrus (Ellison, Schindler, Pattison, & Milner, 2004; Karnath, Ferber, & Himmelbach, 2001).

The found spatial congruency effect further suggests that prior knowledge of the location of the object within a scene influences the process of comparing structural object description with expected object types at relatively early stages of processing (Bar & Ullman, 1996; Friedman, 1979; Heinze et al., 1994; Kosslyn, 1996; Palmer, 1975; Ullman, 1996). This effect may indicate that the *perceptual analysis* of spatially incongruent objects might have been processed automatically even before the deep semantic analysis takes place (Biederman et al., 1982). As mentioned above, such components have been shown to be affected by task manipulations that influence perceptual encoding, with spatial (Heinze et al., 1994) or non-spatial features such as orientation and frequency (Kenemans et al., 1993).

One interesting outcome of Experiment 1 is that the earlier perceptual/spatial match/mismatch process did not eliminate the structural and the semantic analysis revealed by the Nsd/N300 and N400 effects. That is, the spatially mismatching object was still being evaluated at the structural and semantic levels, suggesting some dissociation and independence between these two processes.

Compared to the late positivities found in the Ganis and Kutas (2003) study, we found an attenuated late negative shift for the incongruent conditions confined to a few electrodes in the lme analysis, but not in the ANOVAs. In the Ganis and Kutas study, the authors found a strong centro-parietal positive shift, and interpreted this component as part of the P3b family. The authors noted that "... late positivity might be related to the engagement of processes required to integrate the incongruous object and scene into a mental model, perhaps to enable efficient episodic encoding" (p. 140, citing Donchin & Coles, 1988). In both Ganis and Kutas and the present experiment, the scene preceded the object and this might have created expectations. Later, these expectations might have turned into a surprise effect when they were not satisfied, leading to more difficult episodic integration processes (Neville, Synder, Woods, & Galambos, 1982).

The question arises why we did not find a positive shift. There are two possible ways to interpret this finding: (i) this might be due to the weakness of the surprise effect (see above) preceding the late integration processes or (ii) this might be due to the "impossible" episodic encoding of some of our scene–object pairs (i.e., the cat on a wall vs. the painting on the floor both of which are spatially unexpected but, in the latter case, entirely possible; similarly, seeing a plane in the parking lot is spatially unexpected but nevertheless possible, while seeing a truck hanging in the air is almost impossible), which cannot trigger a possible contextual update. Also note that, the Congruity effect was not observed in the by-item ANOVAS in the 500–850 ms time window, indicating that the effect cannot be generalized over all the items. We propose that the object–scene association and episodic encoding might be more eligible (plausible) in paradigms used by Ganis and Kutas (2003) (i.e., a toilet paper appearing in the soccer field) than in many of the stimuli used in the current study. Since the episodic update is very difficult in some stimuli in our experiment, this might have blocked further episodic evaluation/encoding processes leading to the attenuation of the Late Positive Shift.

Strikingly, comparing the late positive effect with previous psycholinguistic studies, it has been shown that for syntactically un-repairable (ungrammatical) sentences, late negative shifts could emerge on the right hemisphere electrodes, while such negative shifts disappeared for the syntactically repairable but difficult to process (garden-path) sentences (Hopf, Bader, Meng, & Bayer, 2003). In the much later time-window, 1025–1200 ms, we observed a negative-going waveform, supporting the hypothesis that late negativities or the absence of the late positivities in our experiment might be a by-product of the impossibility of the episodic and spatial update for the stimuli we used.

In Experiment 1, we followed Ganis and Kutas (2003) and presented the scene prior to the target object. Given the speed with which scene gist can be determined, an important theoretical question is whether the interaction between object and scene representations can take place when they are concurrently presented. In the second experiment we provided the scene and the object together without the scene preceding the object. In this case, would we still be able to observe the earlier influences of context as well as the later integration effects?

Thus, in the second experiment we asked the question of (i) whether the early perceptual processes would still be apparent when the scene and object are simultaneously presented and (ii) whether the N300–N400 effect would emerge when the scene and object are presented together without any preceding context. We also tested the generalizability of Mudrik et al.'s (2010) findings to the *spatial* domain. Note that, in contrast to our design, in their experiment the images were presented for a second, and the experimental task was very different: images of the objects used in that experiment were always held by human beings (which probably led to sensory-motor cortical activation regarding to the hand and the fingers), and the participants were asked to decide whether the object was held by one or two hands.

5. Experiment 2

5.1. Materials and methods

5.1.1. Participants

Twenty-four undergraduate and graduate students, 16 females and 8 males, aged 19–27, from the University of Edinburgh participated in the experiment. They were paid £6 for their participation. All the participants were native speakers of English. An experiment lasted around 1 h, including cap and electrode preparation. All subjects were reported to have normal or corrected to normal vision, and no neurological or psychiatric problems.

5.1.2. Apparatus

The same apparatus and stimuli were used as in Experiment 1.

5.1.3. Procedure

The procedure was identical to Experiment 1, besides the following differences. At the beginning of each trial, a location cue (cross) appeared on the screen for a variable time-interval of 725–1075 ms (mean 900 ms). Then, the cue disappeared and the image (object and scene together) was shown for 300 ms. We used this longer duration cue with a variable time-interval in the trial onset in this experiment in order to increase the performance of the subjects at comparable levels with Experiment 1. In Experiment 1, subjects had a 300 ms cue followed by the 300 m scene, before they saw the target image. We tested two people, who were not included in the ERP experiments, in a behavioral version of Experiment 2, and realized that the task was too fast with the 300 ms duration cue. In this way we aimed to give subjects enough time to be ready for the detection of the scene–object pair.

5.1.4. EEG recording and signal processing

Recording was similar to Experiment 1, except the length of the baseline was 150 ms between −150 ms and 0 ms before the image onset.

5.1.5. Performance assessment

Three subjects were found to show an extremely high number of incorrect or no responses (40% of the trials per condition), and one subject was found to show excessive artefacts. They were eliminated from the subject pool before EEG analysis. The behavioral and ERP data for the remaining 20 subject is presented below.

6. Results

6.1. Behavioral results

Rating speed in congruous trials, $M = 1118.2$ ms (SD = 256.1) was numerically faster than the incongruous trials, $M = 1120.1$ ms (SD = 263.1), but we did not find any effect of Congruity on the reaction times either in the repeated measures ANOVAs, Fs < 1, or linear mixed effect modelling, $t = 0.27$. Similar to Experiment 1, the absence of any effect in RTs is probably due to the fact that the task of the subject was not to respond as quickly as possible, but as correctly as possible. In contrast, there was a significant effect of condition on the accuracy results; more errors were committed and the accuracy rate was smaller for the incongruous conditions, 82.8% (SD = 37.7) than the congruous conditions, 88.4%, (SD = 31.9); $F1(1,19) = 9.72$, $p < .01$, $F2(1,39) = 9.24$, $p < .01$, LME with family chosen as binomial; $t = -2.25$.

6.2. ERP results

In the ERP analysis, similar to the first experiment, we first prepared a series of small time-bins (25 ms time-bins) from 0 ms up to 1200 ms. Then, we conducted statistical analyses (repeated-measures by-subjects ANOVA) on each of these time-bins. The results are presented in Appendix B.1. ANOVA analysis in the early time-bins did not yield any significant differences up to 450 ms. By using this time-bin analysis, we constructed two larger time-windows to analyse the possible effects of Congruity: 450–700 ms and 925–1125 ms. The following reports repeated measure ANOVAs by items ($F1$) and by subjects ($F2$) for these time windows, as well as linear mixed effects by using lmer function. The summary of ANOVA analyses is presented in Appendix B.2. We also presented lmer models tested and the best fitting model for each ROI in Appendix B.3.

6.2.1. N400 effect: 450–700 ms

ANOVAs: A repeated measure ANOVA revealed that, in the 450–700 ms time-window, there was a significant interaction between ROI and Congruity only on the Left Hemisphere ROIs, $F1(1,19) = 3.8$, $p < .01$, $F2(1,39) = 3.5$, $p < .01$. In the Midline-ROIs, the interaction was marginal only for the by-subject analysis, $F1(1,19) = 2.1$, $p = .072$. Resolving the interactions showed that none of the ROIs revealed any effect of Congruity in by-item as well as by-subject analyses (corrected $p > .05$) (Figs. 7 and 8).

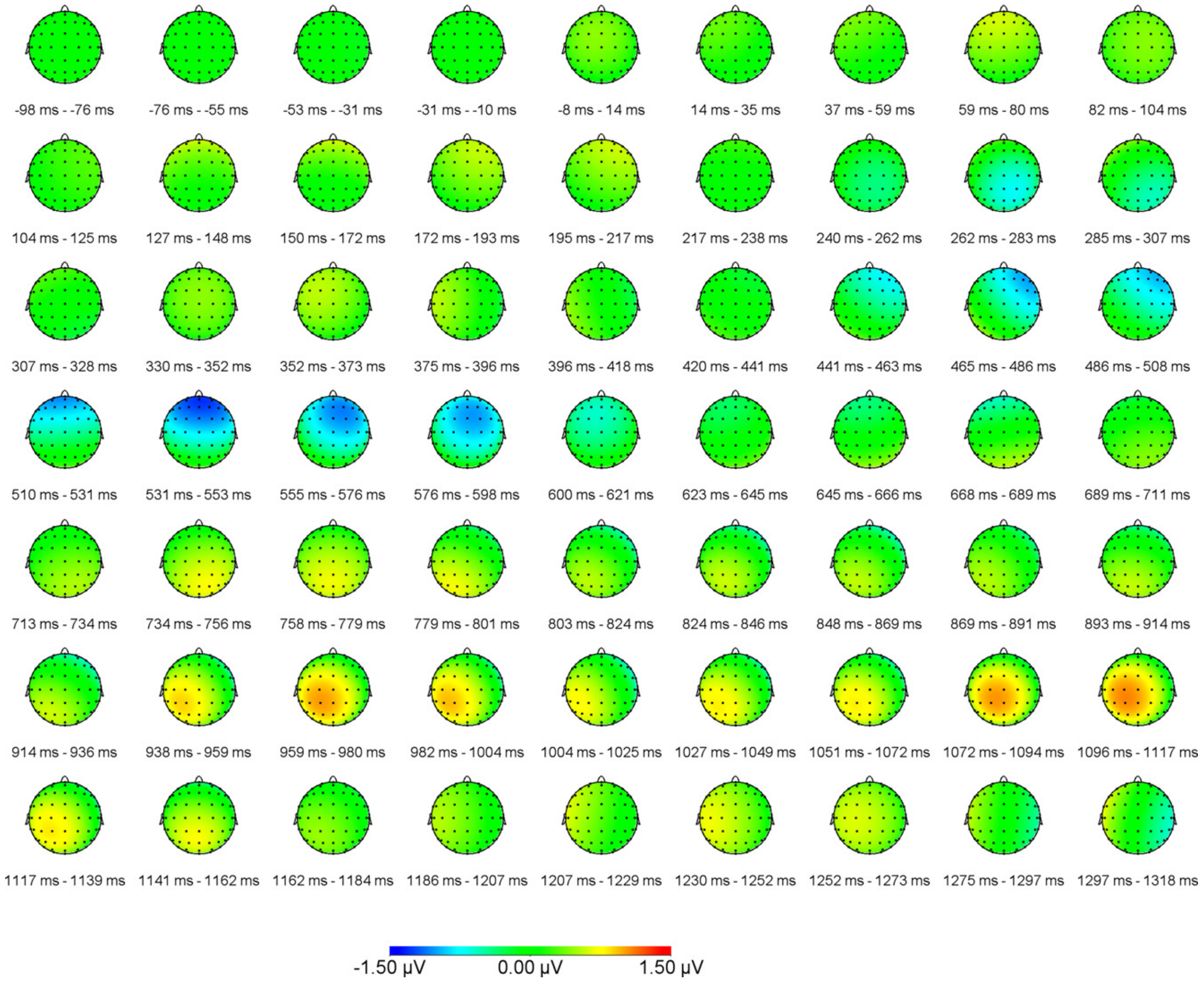


Fig. 7. Scalp maps of the difference between the Incongruent and Congruent conditions, shown for each 25 ms time-bin in Experiment 2. Blue color indicates the difference is negative. Red color indicates the difference is positive. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

LME analysis: Linear mixed effect modelling showed that a marginal effect of Congruity emerged on the ROI3, $t = -1.86$, and a significant effect emerged on the ROI8, $t = 2.45$ of the Left-ROIs. The effect on the Midline-ROIs was marginal on ROI5, $t = -1.86$, and ROI9, $t = 1.90$.

These effects revealed that, unlike Experiment 1, the N300/N400 effect was weaker and confined to a smaller region. The findings revealed that the expected N400 effect might have a deeper source, which has a weak negative polarity reflecting mainly on the small set of left centro-frontal electrodes, and a positive polarity reflecting on the posterior left electrodes. In order to assess the localization of the source of the effect, LORETA analysis was conducted.

LORETA analysis: LORETA analysis showed that the effect in the N400 time window was mainly located in the left hemisphere middle temporal lobe, Brodmann Area 21 ($x = -52$, $y = -3$, $z = -24$). This topography is very similar to the N300/N400 effect observed in Experiment 1, except the fact that it is now on the opposite side of the brain (Fig. 9).

6.2.2. Late effect: 925–1125 ms

ANOVAs: In the 925–1125 ms time-window, there was a significant effect of Congruity on the Left-ROIs in the by-subjects analysis, $F1(1,19) = 5.63$, $p < .05$, but the effect was marginal in the by-items analysis, $F2(1,39) = 3.46$, $p = .07$.

LME analysis: Linear mixed effect modelling showed that the effect of Congruity was marginal on ROI6, $t = -1.58$, and a significant on ROI7, $t = 2.76$, and ROI8, $t = 3.48$, of the left hemisphere. The effect on the Midline-ROIs was marginal on ROI8, $t = -1.8$, and significant on ROI9, $t = 2.9$.

These results showed that the incongruent conditions elicited a slow positive-going component relative to the congruent conditions on the left posterior electrodes.

LORETA analysis: LORETA showed a distributed network of activity in the late time windows. The activity was dominant on the medial frontal gyrus, Brodmann area 11 ($x = -4$, $y = 53$, $z = -10$) (Fig. 10).

7. Discussion

Experiment 2 did not produce early perceptual effects of spatial congruency. There was an attenuated N400 effect, which shifted to the left hemisphere electrodes. LORETA analysis showed that the N400 effect was mainly confined to the left medial temporal gyrus, in contrast to the findings of Experiment 1. This small N400 might indicate that negative components in this time interval are related

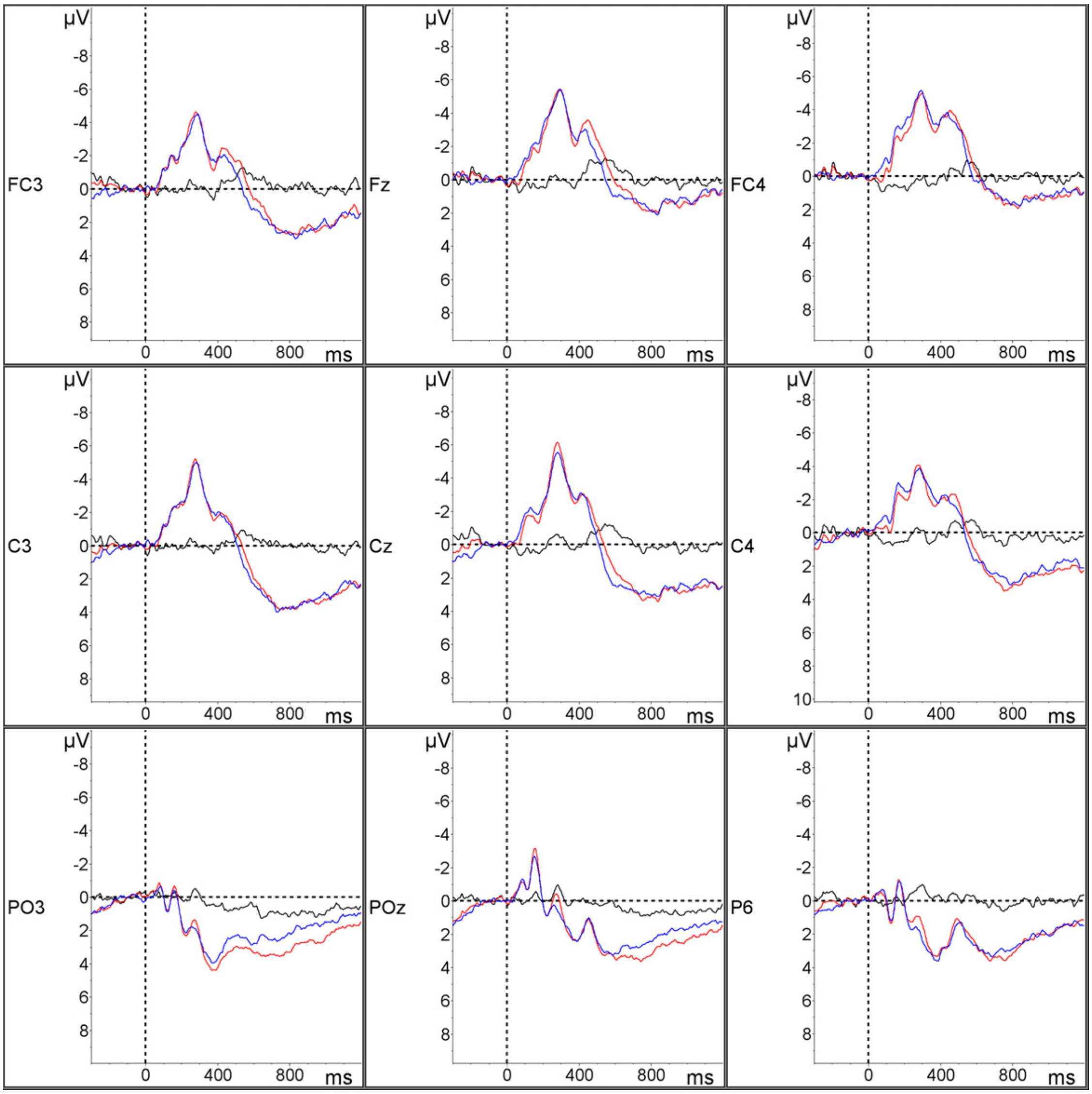


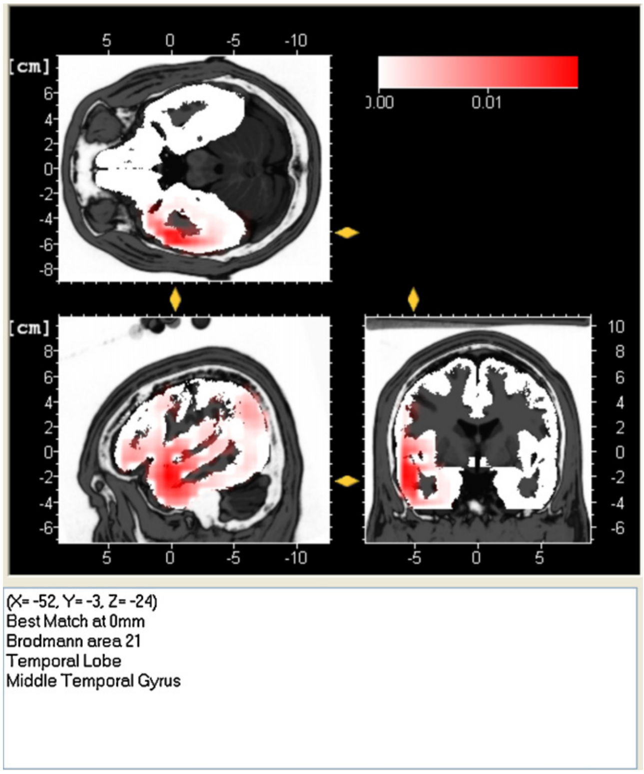
Fig. 8. ERP amplitude-time plots and the difference maps of Experiment 2. Red line shows the incongruent conditions, and the blue line shows the congruent conditions. Black line is the difference between the two. The mean of the mastoid electrodes is the reference. Base-line is selected as 150 ms duration between −150 ms and 0 ms before the object onset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to the semantic integration difficulty of the objects, which normally should appear in different locations in the scene. Since this effect is much smaller than the N400 effect observed in Experiment 1, we propose that the N400 effect in Experiment 1 might not have been an outcome of a single process (i.e., difficulty in integration) but also a combination of a number of different processes (i.e., violating the semantic expectancy; difficulty in selecting and activating the target object representations while the representations of the expected items were active in memory). Most of those effects might have been triggered by contextual expectation. Similarly, Nsd/N300 was absent, indicating that such structural processes might have also followed an expectancy-based selection and integration processes. So, in Experiment 2, only a subset of those effects might have played a role.

For the late effects, we found a left-posterior positive shift for the incongruent conditions significant only in the by-subject analysis. This is in contrast to Mudrik et al. (2010) study, in which late negativities over most of the electrodes were being reported. This difference might be due to the difference in the task requirements between studies as well as the experimental paradigms used. In our study, participants were not given much time (only 300 ms) to perceive both the object and the scene. In Mudrik et al. experiment participants had a second to process the image, and they were asked to perform a different task (i.e., decide whether the action was executed with one or two hands). In psycholinguistics, it has been well documented that the amplitude changes in the late ERP components, particularly P600 and Syntactic Positive Shift (SPS), are associated with non-automatic, task-related processes such as conflict monitoring between a number of syntactic and semantic features (Kuperberg, 2007). Thus, late ERP components should be interpreted with caution. The positive shift in Experiment 2 is very small, and confined to a very narrow topography which made it very hard for us to interpret. Below in Section 8, we will further evaluate the differences between the Mudrik et al. study and our findings.

8. General discussion

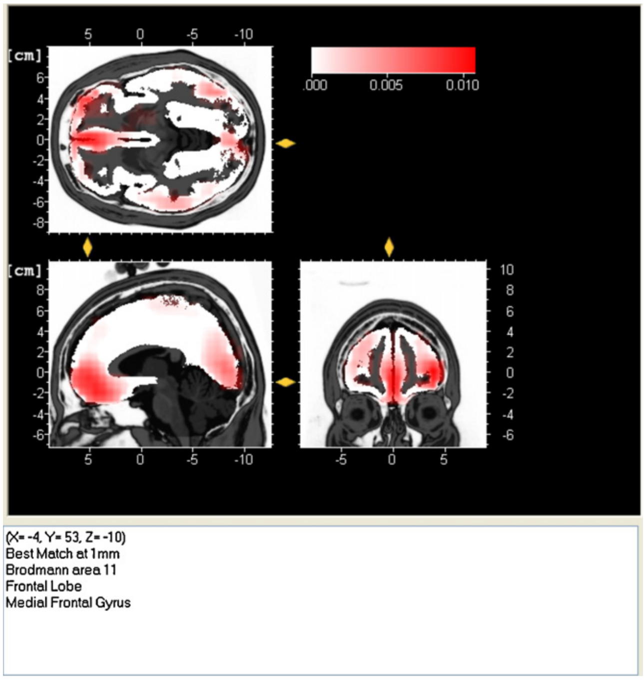
In Experiment 1, we showed that when the scene context precedes the target object, an early negativity is observed for the congruent conditions relative to the incongruent conditions around 160 ms in left hemisphere electrodes. This component was then



(X= -52, Y= -3, Z= -24)
Best Match at 0mm
Brodmann area 21
Temporal Lobe
Middle Temporal Gyrus

Fig. 9. LORETA analysis of the N400 effect in Experiment 2 on the difference wave between the incongruous and the congruous conditions, time window of 450–700 ms.

followed by an N300–N400 complex for the incongruent conditions relative to the congruent conditions; N300 peaking centro-frontally around 300 ms, and N400 peaking centro-parietally around 400 ms. LORETA analysis showed that these activations are taking place in the right superior and right middle temporal gyrus. In Experiment 2, we showed that most of these components found in Experiment 1 were actually by-products of the scene and location preview, which created expectancies towards a set of spatial, categorical, and semantic object features. Only a small N400 effect was observed on the left hemisphere electrodes in Experiment 2. LORETA analysis revealed that this activity was located in the left middle temporal gyrus.



(X= -4, Y= 53, Z= -10)
Best Match at 1mm
Brodmann area 11
Frontal Lobe
Medial Frontal Gyrus

Fig. 10. LORETA analysis of late effects in Experiment 2 on the difference wave between the incongruous and the congruous conditions, time window of 925–1125 ms.

Our findings suggest three important aspects of the spatial relations between an object and the scene within which it is embedded during object processing. First, early spatial-perceptual encoding can only be possible if a context and a location cue are provided before the target image, and the participants are asked to conduct spatial encoding of that part of the scene relative to the other parts of the scene in advance. Second, structural description (Nsd/N300) and semantic N400 effects are also mainly context driven, reflecting processes related to contextual expectations. Third, scene–object integration related processes, and the incongruity effect, can still be observed when the scene and object are presented together, but may be attenuated.

Our findings indicate that a spatial relationship between the object and the scene is a fundamental part of the scene–object integration. One important point to note about the current findings is that participants might have followed different processing routines depending on the availability of sources: in Experiment 1, when they had the location cue and the context prior to the object's appearance, they may have followed a spatial-encoding pathway to initiate the scene–object integration. But when they had to process the scene and object simultaneously in Experiment 2, they might have instead utilized a different pathway. One possibility in this case is that subjects first attempted to identify the object and then follow this up with scene–object integration.

Another possibility is that they processed the object and scene in parallel as suggested by a number of recent studies on scene processing (Bar, 2004; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Oliva, Torralba, S. Martinez-Conde, & Tse, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006). Thus, in Experiment 2, the process would consist of processing of the target object followed by the integration of the object with the surrounding scene, or else the scene and object processing unfolds in parallel, minimizing feedback. Either routine might not have allowed successful scene–object structural and spatial integration, because processing was initially allocated on the target object features, rather than their spatial relation to the scene.

LORETA analysis confirms that two different cortical regions were involved in Experiments 1 and 2. When the early perceptual processes are in play, this did not prevent further structural and semantic analysis, as the Nsd/N300–N400 effects were still observed in Experiment 1. Strikingly, even though the spatial–semantic integration processes followed the initial perceptual processes, and such effects were located in the right medial temporal gyrus in Experiment 1, similar regions in the contralateral side of the brain was found to be activated in Experiment 2 in the N400 time window. This may indicate that, either top-down (Experiment 1) or bottom-up or parallel (Experiment 2) processes can finally yield spatio-semantic integration, but in different topographies on the scalp, in different locations in the brain, and with very different strengths.

One of the interesting overlaps we realized between our LORETA analysis and the Gronau et al. (2008) fMRI findings (compared to the first presentation of the prime-target pairs in their experiment) is that, some of the temporal and frontal activations in our experiments looked very similar to their fMRI results. For example, Gronau et al. reported middle and superior temporal gyrus activity along with the frontal gyrus activity for the spatial × semantic interactions. While our source location method is not, of course, spatially as strong as the high-resolution fMRI method, we were able to show that these regions were also active in our tasks. In contrast to our findings, Gronau et al. fMRI findings are mainly bilateral. Thus, LORETA analysis might have shown a transient short-duration activity, such that the initiation of the relevant cortical regions would be mainly unilateral and hemisphere dependent. On the other hand, bilateral findings of Gronau et al. might be due to the spreading of the cortical network activity in the later stages of processing, over a long period of time.

One question we need to answer is why we did not observe early as well as strong N400 effects of congruency in Experiment 2, as Mudrik et al. (2010) did in their study. Comparison of our findings with Mudrik et al.'s findings requires some points to be made clear. First, in contrast to the current study, Mudrik et al. did not assess spatial congruency, they mainly targeted semantic congruency. Second, as mentioned above, their task was very different than ours; they asked participants to detect whether the object was being held with one or two hands. Such a task might have influenced some of the brain mechanisms which might have facilitated semantic scene–object integration. First, participants might have been more prone to activating their sensory-motor cortical regions relevant for grasping and hand movements, which in turn, might have facilitated the match-mismatch computations. Second, participants might have put themselves in the position of the individuals in the image, as if they were doing the action. Third, participants might have paid attention to the surroundings of the object or the human entity holding the object more than the object itself, and may require a different semantic interpretation of the task rather than the semantic congruency per se.

Related to the first possibility, we think that the task requirement in Mudrik et al. (2010) study narrowed down the possible set of candidate object features, or made some features of the object more accessible than the others. This in turn, might have had an effect on allocating a subset of mental search space, instead of using all possible search space in the memory for the scene–object matches. This could have speeded up, or at least increased the quality of the perceptual and semantic processes. In our experiment, there was no narrowing-down of the mental search space via selecting a task which was favoring some of the sensory-motor attributes. Instead, our task was simply to indicate if the target object was appropriately positioned within the scene. In this regard, our task in Experiment 2 made object–scene integration in short time-windows very difficult. The second possibility is related to the first one, theoretically closer to recent approaches in mirror neuron research (Hauk & Pulvermuller, 2004) and multisensory information use of the functional and the goal-directed action affordances induced by the context (Bach, Knoblich, Gunter, Friederici, & Prinz, 2005). Overall, these possibilities are not within the scope of our paper, and should be examined by new studies.

Absence of early effects was not the only difference between our findings and the Mudrik et al. (2010) findings. While the weak N300/N400 effect we found in Experiment 2 starts with a similar topography – left centro-frontal hemisphere – we failed to show a strong negative going slow wave for the incongruent conditions. We think that the lack of strong late effects (towards positive or negative polarities) in Experiment 1 and Experiment 2 might be due to the lack of similar cognitive processes. Thus, we will evaluate and compare late ERP component in both experiments in the following. The Ganis and Kutas (2003) study found a P600 response to exist in visual object recognition preceded by a scene. We did not find P600 in Experiment 1. We observed a negative going slow wave mainly on the small portion of the right hemisphere electrodes. Before we interpret the late negativity, we should ask why P600 is missing in Experiment 1. In psycholinguistic research, Late Positive Shift (LPS) and P600 components were reported to index syntactic re-analysis of the linguistic input (Holcomb, 1993; Osterhout & Holcomb, 1992, 1993; West & Holcomb, 2000). P600 is generally taken as a neurological response to an effort to generate alternative parses of the linguistic input. To put it another way, if the combinatory properties of words do not completely match in semantic or syntactic dimensions, but we can detect an alternative route to consolidate and combine such affordances, then the P600 emerges. Applying similar logic to object–scene processing, it could be that if objects are not semantically congruent, but the participant can detect an alternative way to encode the object–scene relationship, then this may lead to P600.

As noted above, late negative shifts were reported for ungrammatical (and non-recoverable) syntactic constructions in psycholinguist research (Hopf et al., 2003; Hopf, Bayer, Bader, & Meng, 1998). In our experiments, spatially manipulated objects were not always in positions where their affordances could be stretched to incorporate the surrounding structural properties (i.e., a cat cannot be walking on a surface of the side-wall, a truck cannot fly). Thus, our findings might indicate that the participants could not consistently reconsolidate incongruent objects with the scene. This inability to re-analyse the object–scene relationship may be what led to the lack of a P600 effect for the incongruent objects in Experiment 1. The slow continuation of the negativity after N400 might be related to the "open" episodic computations which could not be resolved for such object–scene pairs, while positive shifts might indicate "closed" episodic representations. We also attribute the lack of a strong P600 effect in Experiment 2 to the very same reason. In addition, very brief and simultaneous presentation of the scene-object pairs was possibly not sufficient in providing strong feedback loop activity in memory. This might in turn have attenuated the quality of episodic evaluation and contextual update.

9. Conclusion

In two experiments, we showed that (i) when the spatial location of an object is known prior to its appearance, participants can trigger relevant object representations that can lead to expectations about its identity, forming possible sets of objects and their structural and spatial features, but (ii) when the spatial position is not known in advance, early perceptual and later structural and semantic effects mainly disappear.

Objects and scenes have spatially defined intricate functional and structural relationship which may be used to create expectations not only about such sets of objects, but also about common features and attributes shared among the objects. Spatial encoding of the parts of the scene may lead to early left hemisphere negativity which is relevant for rapid perceptual analysis. Spatially relevant structural and semantic attributes of the object then can be accessed, and the violations of expectations between the sets of objects and the actual target object can lead to an N300–N400 effect. When the scene and object are presented together, the object might be analysed first, or in parallel to the scene. Only later, the semantic congruency evaluation between the scene and the object can be made.

Acknowledgments

This study has been conducted in the University of Edinburgh, Department of Psychology, when the first author was working there as a postdoctoral associate. We would like to thank Sandra Andraszewicz for her assistance in ERP data collection. We also thank two anonymous reviewers for their invaluable comments and suggestions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuropsychologia.2012.02.011.

References

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes*, *18*, 405–430.

Anllo-Vento, L., Luck, S. J., & Hillyard, S. A. (1998). Spatio-temporal dynamics of attention to color: Evidence from human electrophysiology. *Human Brain Mapping*, *6*, 216–238.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005). Action comprehension: Deriving spatial and functional relations. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(3), 465–479. doi:10.1037/0096-1523.31.3.465, pii: 2005-06735-006

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629. doi:10.1038/nrn1476

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception*, *25*, 343–352.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201–212.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34–44.

Castelhano, M. S., & Heaven, C. (2011). The relative contribution of scene context and target features to visual search in real-world scenes. *Attention, Perception, & Psychophysics*, *72*(5), 1283–1297.

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(3), 660–675.

Clark, V. P., Fan, S., & Hillyard, S. A. (1994). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, *2*(3), 170–187.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*, 559–564.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–427.

Doninger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., Schroeder, C. E., et al. (2000). Activation timecourse of ventral visual stream object-recognition areas: High density electrical mapping of perceptual closure processes. *Journal of Cognitive Neuroscience*, *12*, 615–621.

Ellison, A., Schindler, I., Pattison, L. L., & Milner, A. D. (2004). An exploration of the role of the superior temporal gyrus in visual search and spatial perception using TMS. *Brain*, *127*(10), 2307–2315. doi:10.1093/brain/awh244

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469–495.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology Learning, Memory, and Cognition*, *27*, 202–224.

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730–747.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.

Friedman, A., & Liebelt, L. S. (1981). On the time course of viewing pictures with a view towards remembering. In D. F. Fischer, R. a. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 137–155). Hillsdale, NJ: Erlbaum.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*, *16*(2), 123–144.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for common sense: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*, 89–106.

Gronau, N., Neta, M., & Bar, M. (2008). Integrated contextual representation for objects' identities and their locations. *Journal of Cognitive Neuroscience*, *20*(3), 371–388. doi:10.1162/jocn.2008.20027

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, *43*, 21–30.

Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical transactions of the Royal Society of London. Series B: Biological sciences*, *363*, 1055–1069.

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339–1350.

Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, *30*, 1383–1400.

Hauk, O., & Pulvermuller, F. (1888). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*, 1090–1103.

Hauk, O., & Pulvermuller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping*, *21*(3), 191–201. doi:10.1002/hbm.10157

Heinze, H. J., Mangun, G. R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T. F., et al. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature*, *372*, 543–546.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.

Henderson, J. M., Weeks, P. A. J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210–228.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*, 47–61.

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, *24*(259–276).

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, *127*, 398–415.

Hopf, J.-M., Bader, M., Meng, M., & Bayer, J. (2003). Is human sentence parsing serial or parallel? Evidence from event-related brain potentials. *Brain Research. Cognitive Brain Research*, *15*(2), 165–177.

Hopf, J.-M., Bayer, J., Bader, M., & Meng, M. (1998). Event-related potentials and case information in syntactic ambiguities. *Journal of Cognitive Neuroscience*, *10*(2), 264–280.

Huynh, H., & Feldt, L. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics*, *1*(1), 69–82.

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*(26), 3286–3297. doi:10.1016/j.visres.2007.09.013

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163–178.

Karnath, H.-O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature*, *411*(6840), 950–953.

Kenemans, J. L., Baas, J. M. P., Mangun, G. R., Lijffijt, M., & Verbaten, M. N. (2000). On the processing of spatial frequencies as revealed by evoked-potential source modeling. *Clinical Neurophysiology*, *111*(6), 1113–1123.

Kenemans, J. L., Kok, A., & Smulders, F. T. Y. (1993). Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements. *Electroencephalography and Clinical Neurophysiology*, *88*, 51–63.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*, 27–39.

Kosslyn, S. M. (1996). *Image and brain*. Boston, MA: MIT Press.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, *8*, 533–572.

Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology*, *34*(4), 383–398.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(565–572).

Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences*, *94*(20), 10979–10984. doi:10.1073/pnas. 94. 20.10979

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2), 1–11.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(01), 53–65.

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object–scene processing. *Neuropsychologia*, *48*(2), 507–517.

Neville, H. J., Synder, E., Woods, D., & Galambos, R. (1982). Recognition and surprise alter the human visual evoked response. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(6), 2121–2123.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*(Part 2), 23–36.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly—Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413–437.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*, 519–526.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*(441–474).

Pascual-Marqui, R. D., Mitchel, C. M., & Lehmann, D. (1994). Low reolution electromgneti tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, *18*, 49–65.

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, *22*, 642–647.

Schendan, H. E., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia*, *40*, 931–945.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *Neuroreport*, *9*, 2195–2200.

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, *20*, 2037–2057.

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, *40*, 160–164.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. doi:10.1037/0033-295x.113.4.766

Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge, MA: MIT Press.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*, 380–393.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, *97*, 279–293.

Võ, M. L.-H., & Henderson, J. M. (2009). Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, *9*(3), 1–15.

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience*, *12*, 1024–1037.

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363–375.