Spatiotemporal analyses of the N170 for human faces, animal faces and objects in natural scenes

Guillaume A. Rousselet,[1,CA] Marc J-M. Macé and Michèle Fabre-Thorpe

Centre de Recherche Cerveau & Cognition, CNRS-UPS UMR 5549, Faculté de Médecine de Rangueil, 133 route de Narbonne, 31062 Toulouse, France
[1]Present address: McMaster University, Department of Psychology, Hamilton L8S4K1 ON, Canada

[CA]Corresponding Author: rousseg@mcmaster.ca

Received 2 September 2004; accepted 12 October 2004

We assessed the specificity to human faces of the N170 ERP component in the context of natural scenes. Subjects categorized photographs containing human faces, animal faces and various objects. Spatiotemporal topography analyses were performed on the individual ERP data. ERPs elicited by animal faces were similar to human faces ERPs but with a delayed face activity. In the N170 time window, ERPs to human and animal faces had a different topography compared with object ERPs. Such data suggest that N170 generators might process various stimuli with a coarse facial organization and show the care that must be taken in comparing scalp signal to faces and other objects as they are probably generated, at least partially, by different cortical sources. *NeuroReport* 15:2607–2611 © 2004 Lippincott Williams & Wilkins.

Key words: Animal faces; ERP; Human faces; Natural scenes; N1; N170; Objects; Visual categorization

INTRODUCTION

The N170 is a posterior lateral event-related potential (ERP) characterized by a larger amplitude in response to human faces than to many object categories, including animal faces [1–3]. However, recent experiments have showed a clear N170 elicited by ape and monkey faces [4,5], suggesting that the N170 might not be specific to human faces *per se* and might extend to other species or objects sharing with the human face a similar spatial organization [4,6].

We tested this idea in an experiment comparing human and varied animal faces [7]. Contrary to many experiments that relied on homogenous stimuli centered on a uniform background we used close-ups of faces with different sizes and positions in natural scenes. Scenes without faces and containing various objects were used as distractors. Surprisingly, both types of faces were associated with a larger N170 compared with control objects. The N170 for human faces differed from the N170 for animal faces only by a shorter peak latency that might be explained by the broad heterogeneity of animal faces. A much coarser facial organization than previously thought might thus be sufficient to trigger a N170. This pattern of results could be explained by a dual system in which both faces and objects would be processed by the same ventral cortical areas [8], with the additional recruitment of lateral posterior areas by faces and face-like objects, which would generate in large part the N170 [3,9]. However, even if the N170 for animal and human faces had the same amplitude in our experiment, they could still have different scalp topographies, implying the involvement of different cortical sources [3,4,10]. In such case, the comparison between the N170 for human and animal faces might not be valid, as argued regarding the comparison between the face N170 and the object N1 [3,4].

To test this hypothesis we applied the same topographical analyses used previously [3] to our data set. Similar topographies suggest the involvement of the same sources, although no such conclusion can be considered as definitive since one topography can potentially be produced by different sources. In contrast, different topographies imply the involvement of at least partially different sources [11]. Following this logic, we expected (1) no difference between the human face N170 and the animal face N170 if they were generated by the same sources and (2) different topographies between object and face ERPs if they originated from different sources.

MATERIALS AND METHODS

Subjects (24, 12 women, mean age ($\pm$ s.d.) age 30 ± 10 years) had normal or corrected to normal vision and gave their written informed consent.

The stimuli were horizontal pictures of natural scenes containing various objects or human and animal faces with different sizes and positions (Fig. 1). Subjects started a block of trials by placing a finger on a response pad for 1 s, then a small fixation point (FP) appeared for 300–900 ms followed immediately by a central stimulus ($20 \times 13.5^{\circ}$) for 26 ms. Subjects had to lift their finger within 1000 ms as accurately as possible each time a target was presented. Subjects had to keep their finger on the pad for at least 1000 ms for non-targets. This delay was followed by a 300 ms black screen, before the FP was presented again. There was a training session with 48 images before the start of the experiment.

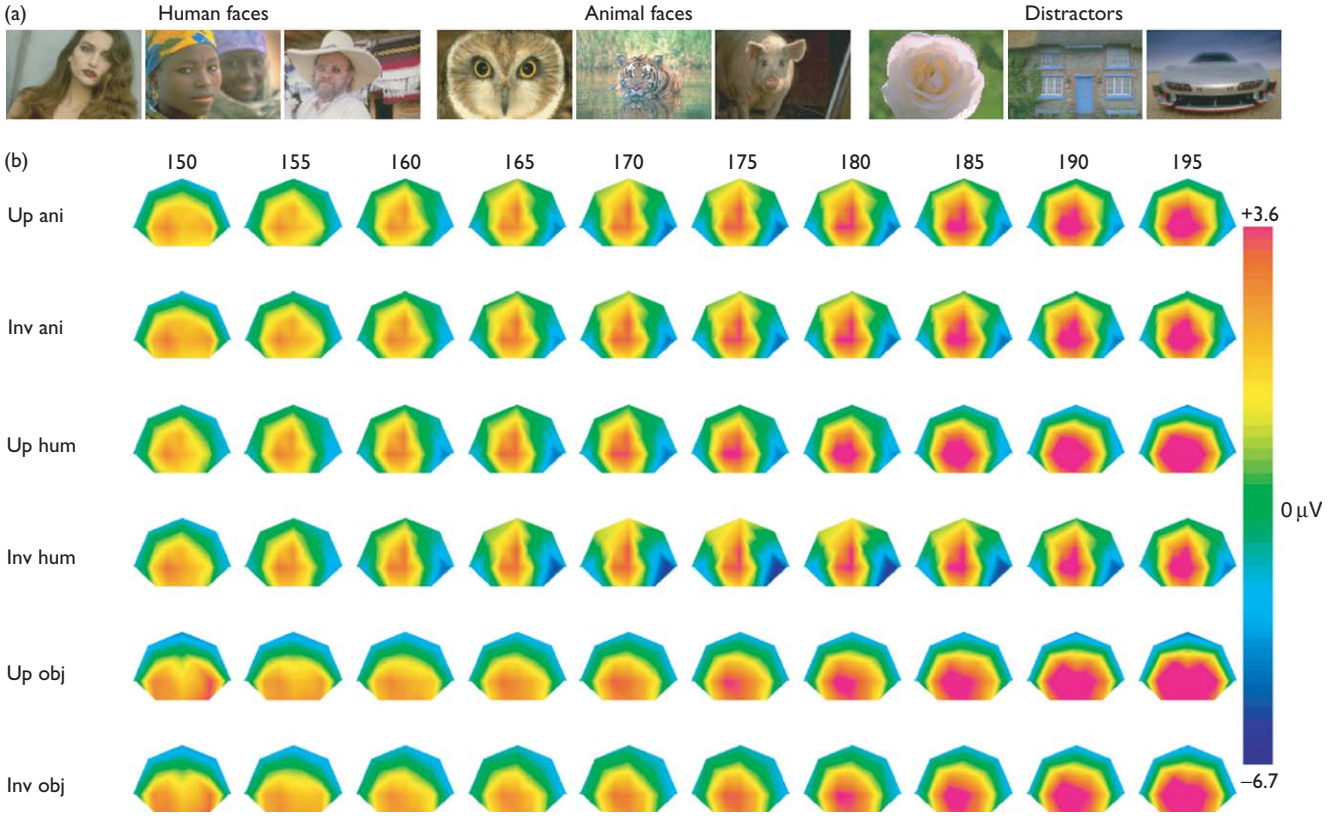

Fig. 1. (**a**) Examples of pictures used in the experiment. (**b**) Back-view maps of the interpolated grand-averaged ERPs between 150 and 195 ms (the 7 most frontal channels are not visible, see Fig. 2). Results are presented for human (hum) and animal (ani) faces seen as targets, the topographies being similar when they were seen as non-targets. Objects (obj) were seen as non-targets in the human face task, the topographies being the same when they were seen in the animal face task. Up=upright, Inv=inverted.

The experiment included 8 blocks of 96 trials. In 4 consecutive blocks, targets were animal faces and in the other 4 blocks, targets were human faces. Among the 48 non-targets, 24 were distractor objects and 24 were targets of the other categorization task (i.e. animal faces in the human face task and vice versa). Half of the images were presented upright and the other half were presented inverted (rotation 180°). There were thus 12 categories in this experiment (8 face categories: human/animal shown upright/inverted and processed as target/non-target; 4 object categories: upright/inverted and processed in the human/animal task). Each subject saw each image once and all conditions were counterbalanced across subjects.

ERP signals were recorded using a SynAmps amplifier system (Neuroscan Inc.) with 32 electrodes (Oxford Instruments) according to the 10-20 system (FP1/2, F3/4, F7/8, Fz, C3/4, Cz, T7/8, Pz, P3/4, PO3/4, POz, TP7/8, T5/T6, PO7/8, O1/2, Oz, Iz, PO9/10, O9/10). Impedances were kept below 5 kΩ. Signals were digitized at 1000 Hz and low-pass filtered at 40 Hz before analysis. Potentials were on-line referenced on electrode Cz and averaged referenced off-line. Baseline correction was performed using 100 ms of pre-stimulus activity. Trials with ocular artifacts over $\pm 80\,\mu V$ and alpha bursts greater than $\pm 40\,\mu V$ were rejected. Only correct trials were averaged.

To evaluate topography differences, mean amplitudes were measured at all 32 electrodes in three time windows, one centered on the mean N170 peak latency (175 ms, range 155–195 ms) and the two others preceding and following this N170 window (115–155 ms and 195–235 ms), for each subject and each condition. Mean amplitudes were normalized to correct for absolute amplitude variations [12] and entered into MANOVAs with category (human faces/animal faces/objects) and time-period as two fixed factors and the 32 electrodes as dependent variables.

We then determined whether topography differences were due to real configuration changes or a latency shift of the same topography between conditions. First, the grand-averaged ERPs for the 12 conditions were segmented into series of stable scalp configurations or template maps [13–15]. This method is independent of the reference electrode and relies on normalized ERPs. Using cross-validation criteria, seven template maps were found to explain optimally the data set over a 50–300 ms post-stimulus interval. Second, the occurrence of the template maps in each of the individual ERP was evaluated over time by a spatial correlation fitting procedure [13,14]. This analysis was performed on the interval 130–210 ms post-stimulus (that contained 5 maps), spanning the entire range of N170 latencies observed across subjects. This procedure revealed how well and how often a given template map explains a given condition for each subject. The results, expressed in terms of global explained variance, were then entered into an ANOVA with Greenhouse-Geisser correction with categories (12) and maps (5) as within-subject factors. *Post-hoc* paired *t*-tests included a Bonferroni correction.

RESULTS

A detailed peak analysis of the N170 recorded in this experiment is described elsewhere [7]. Human and animal faces were characterized by the same N170 topography that differed from the one elicited by objects (Fig. 1). Face ERPs presented a transient posterior lateral negativity (in blue) associated with a median positivity (in red). This topography appeared earlier for human than for animal faces and lasted longer for inverted than for upright human faces. ERPs for objects presented a very weak posterior lateral negativity associated with a much broader median positivity compared to faces.

The global MANOVA analysis on the normalized amplitudes revealed a category × time-period interaction ($F(128,3308)=1.541$, $p<0.0002$). MANOVAs performed separately on each of the three time-periods defined above revealed a significant difference between human faces, animal faces and objects (all $F(64,510)>2.4$, all $p<0.0001$). Figure 2 illustrates the results of the planned comparisons between these three categories at each electrode and each time window. Differences between human faces and objects were found in all time-periods, confirming that the two categories have different topographies. Animal faces differed also from objects in the second and third time-periods but not in the first one. Finally, small differences were found between animal and human faces in the first and third time-periods but not in the second one centered on the N170. The difference in the first time-period appeared at electrodes T6 and PO10 where the N170 had the largest amplitude in the second time-period [7]. It thus probably reflects the earlier onset of the N170 topography for human faces (Fig. 1). The difference in the third time-period was only significant between upright human faces and upright animal faces, not with inverted faces. Thus in this time-period animal faces

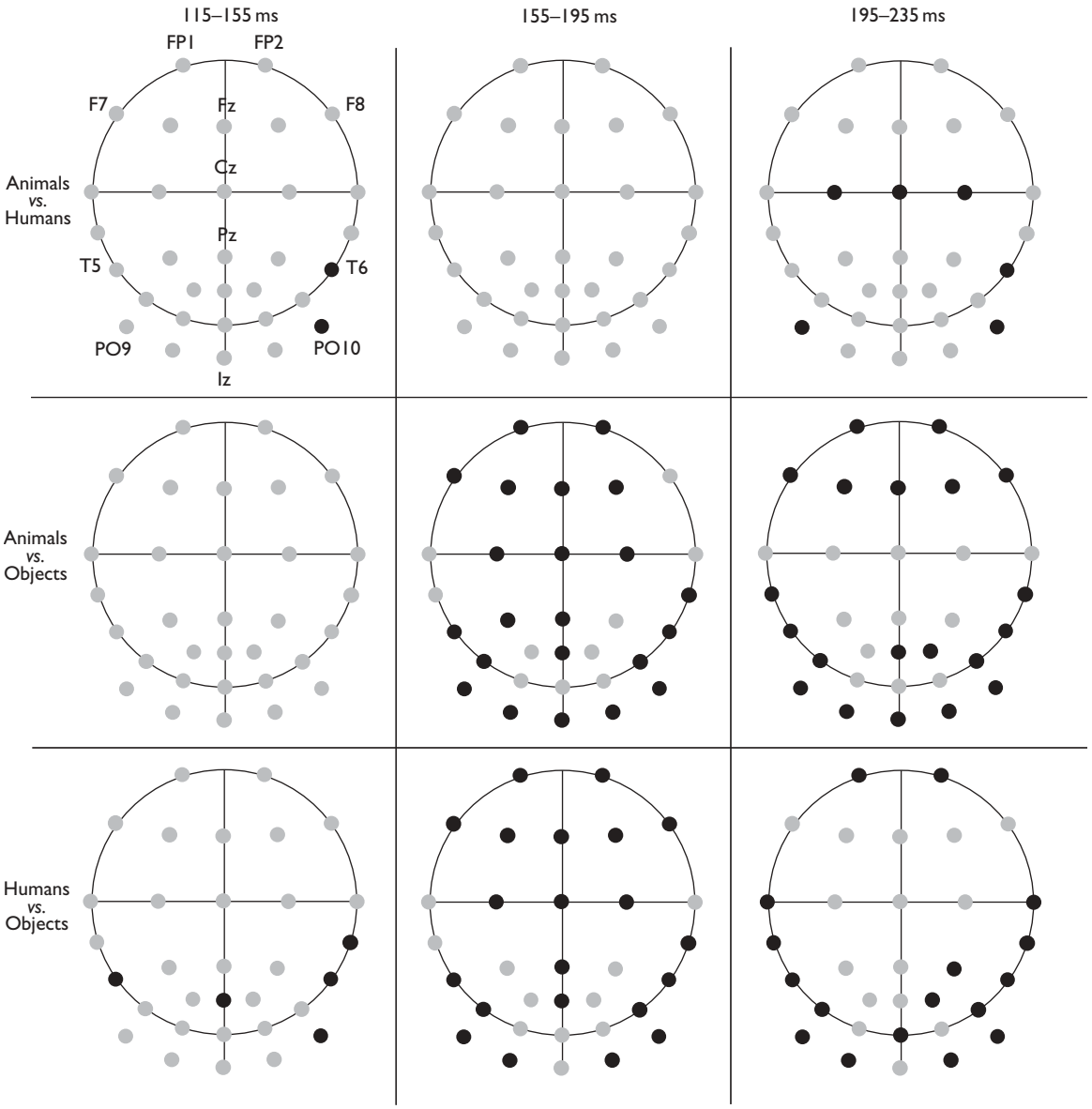

Fig. 2. Topography analysis. The normalized mean amplitudes at each scalp electrode in the different conditions were compared in three time windows. For each comparison the 32 electrodes are presented at their relative position on the head, nose pointing upward. Black electrodes presented a significant difference between 2 conditions ($p<0.05$) while grey electrodes did not.

regardless of their orientation and inverted human faces shared a common topography that differed from the one associated with upright human faces.

A segmentation procedure was applied to the grand-averaged ERPs to determine periods of stable scalp topographies. With this procedure, map numbers are arbitrarily assigned but identical map numbers stands for the same topography [15]. In the time range 150–200 ms after stimulus onset, one stable map was found for all categories (map 4, Fig. 3). However, an extra map centered at the N170 latency was necessary to explain the ERP signal recorded for human and animal faces, but not for objects (map 1, Fig. 3). The statistical reliability of this pattern was assessed on individual data. The ANOVA performed on the global explained variance showed a category × map interaction ($F(10.5,242.5)=4.04$, $p<0.0001$). To clarify this point, category effects were tested in each map. Maps 3, 4, 6 and 7 showed no significant difference (Fig. 3). In contrast, map 1 presented a significant category effect ($F(4.9,112.5)=12.8$, $p<0.0001$). Paired *t*-tests between the 12 categories compared two by two revealed no difference between human and animal faces. However, all comparisons between the face categories and the object categories were significant (all $p<0.05$). Thus, at the N170 latency, map 1 reflected an additional activity associated with the processing of human and animal faces but not objects. Although very similar, the time course of this extra activity diverged between the different face stimuli. It started earlier for human faces

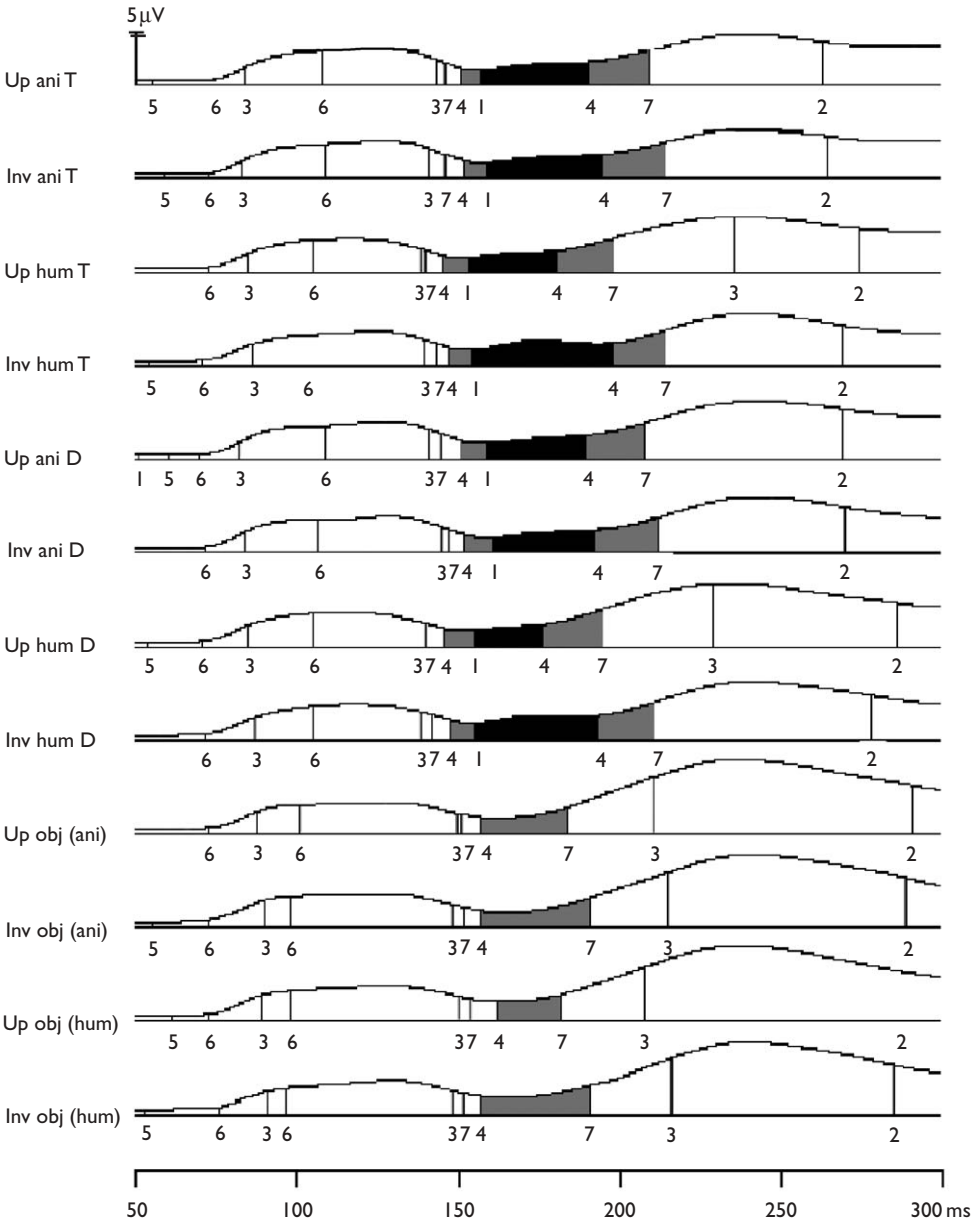

Fig. 3. Segmentation analysis. Segmentation maps are represented on the global field power [22] between 50 and 300 ms after stimulus onset. Human and animal faces, independently of their orientation or task status, were associated with an extra map centered on the N170 latency (map 1, in black) compared with objects (map 4, in grey). T=target, D=distractor, obj=objects seen as distractors in the human face task (hum) or in the animal face task (ani). See also Fig. 1 caption.

compared with animal faces and had a shorter duration for upright human faces compared with inverted human faces or animal faces.

DISCUSSION

We performed a spatiotemporal analysis of the N170 recorded during the presentation of human faces, animal faces and objects in the context of natural scenes. Compared with objects, human and animal faces were associated with a larger N170 [7], a difference that has often been interpreted in terms of the N170 face specificity [1]. Here we demonstrate that independently of these amplitude differences, face and object ERPs also have different scalp topographies at the N170 latency. These different topographies imply that the combination of underlying neuronal sources are also different [11]. The face N170 might thus be qualitatively distinct from the object N1, replicating a conclusion reached in a recent study using isolated objects [3]. Here we show that, in the context of natural scenes, the N170 scalp topographies for animal faces and for human faces were very similar. Following the rational according to which identical brain sources lead to the same scalp topographies [11], this suggests that the varied animal faces used in the task might activate the same neural generators than those activated by human faces. Although one must keep aware that combination of different sources could potentially lead to similar topographies.

The strong similarity between the topographies recorded for human and animal faces might be due to the strong sensitivity of the N170 to the region of the eyes [7,16]. Indeed, lateral cortical areas around the superior temporal sulcus (STS) have been suggested as likely N170 generators [3,9,17], and the STS is part of a social perception network involved in the processing of such attributes as eye gaze and facial expressions [18].

Whatever the precise origin of the N170, our results, along with others [3], strongly suggest the limitations of comparing the signals recorded on the scalp to faces and to other objects since they are at least partly generated by different cortical sources. However, the conditions in which the face-like generators are recruited still need to be understood. From the present data it is clear that the N170 topography is not triggered in an all-or-nothing fashion, as it started later for animal than for human faces. It also lasted longer for inverted than for upright human faces. Furthermore, some of the face-like generators might be involved in the processing of non-face object categories, maybe depending on the task performed. After all, the different topographies recorded for faces and for objects do not necessarily imply the involvement of totally different sources but might also reflect the activity of the same sources whose strength depends on the category and the task at hand. To determine how specifically the cortical network underlying the N170 is recruited by faces, it will be essential to combine the kind of analyses used here and the approach used by others to study the effects of expertise and levels of categorization [19–21].

REFERENCES

1. Bentin S, Allison T, Puce A, Perez E and McCarthy G. Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 1996; **8**:551–565.
2. Rossion B, Gauthier I, Tarr MJ, Despland P, Bruyer R, Linotte S and Crommelinck M. The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 2000; **11**:69–74.
3. Itier RJ and Taylor MJ. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebr Cortex* 2004; **14**:132–142.
4. Carmel D and Bentin S. Domain specificity *vs* expertise: factors influencing distinct processing of faces. *Cognition* 2002; **83**:1–29.
5. de Haan M, Pascalis O and Johnson M. Specialization of neural mechanisms underlying face recognition in human infants. *J Cogn Neurosci* 2002; **14**:199–209.
6. Jeffreys D. Evoked potential studies of face and object processing. *Vis Cogn* 1996; **3**:1–38.
7. Rousselet GA, Mace MJ-M and Fabre-Thorpe M. Animal and human faces in natural scenes: how specific to human faces is the N170 ERP component? *J Vis* 2004; **4**:13–21. http://journalofvision.org/4/1/2/, doi:10.1167/4.1.2.
8. Rousselet GA, Macé MJ-M and Fabre-Thorpe M. Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis* 2003; **3**:440–455. http://journalofvision.org/3/6/5/, doi:10.1167/3.6.5.
9. Watanabe S, Kakigi R and Puce A. The spatiotemporal dynamics of the face inversion effect: a magneto- and electro-encephalographic study. *Neuroscience* 2003; **116**:879–895.
10. Caldara R, Thut G, Servoir P, Michel CM, Bovet P and Renault B. Face *vs* non-face object perception and the 'other-race' effect: a spatio-temporal event-related potential study. *Clin Neurophysiol* 2003; **114**: 515–528.
11. Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson R Jr *et al.* Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 2000; **37**:127–152.
12. McCarthy G and Wood CC. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalogr Clin Neurophysiol* 1985; **62**:203–208.
13. Michel CM, Seeck M and Landis T. Spatiotemporal dynamics of human cognition. *News Physiol Sci* 1999; **14**:206–214.
14. Michel CM, Thut G, Morand S, Khateb A, Pegna AJ, Grave de Peralta R *et al*. Electric source imaging of human brain functions. *Brain Research Reviews* 2001; **36**:108–118.
15. Pascual-Marqui RD, Michel CM and Lehmann D. Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans Biomed Eng* 1995; **42**:658–665.
16. Schyns PG, Jentzsch I, Johnson M, Schweinberger SR and Gosselin F. A principled method for determining the functionality of brain responses. *Neuroreport* 2003; **14**:1665–1669.
17. Itier RJ and Taylor MJ. Source analysis of the N170 to faces and objects. *Neuroreport* 2004; **15**:1261–1265.
18. Allison T, Puce A and McCarthy G. Social perception from visual cues: role of the STS region. *Trends Cogn Sci* 2000; **4**:267–278.
19. Rossion B, Gauthier I, Goffaux V, Tarr MJ and Crommelinck M. Expertise training with novel objects leads to left-lateralized face like electrophysiological responses. *Psychol Sci* 2002; **13**:250–257.
20. Tanaka J, Luu P, Weisbrod M and Kiefer M. Tracking the time course of object categorization using event-related potentials. *Neuroreport* 1999; **10**:829–835.
21. Tanaka JW and Curran T. A neural basis for expert object recognition. *Psychol Sci* 2001; **12**:43–47.
22. Lehmann D and Skrandies W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr Clin Neurophysiol* 1980; **48**:609–621.

Acknowledgements: We thank Denis Brunet for his software CarTool and Roxane J. Itier for help with data analysis and comments on the manuscript. Financial support was given by the Cognitique and the Integrative and Computational Neurosciences programs. The two first authors had PhD grants from the French government and G.A.R. is currently supported by a Tier I Canada Research Chair grant to P.J. Bennett. The CNRS ethical committee approved the experimental procedure.