Dynamics of contour, object and face processing in the human visual cortex

Topi Tanskanen

Brain Research Unit, Low Temperature Laboratory,
Helsinki University of Technology, Finland

Finnish Graduate School of Neuroscience

Academic dissertation to be publicly discussed,
by the permission of the Faculty of Behavioural Sciences of the University of Helsinki,
in Auditorium KE2 at the Helsinki University of Technology,
on the 14th of May, 2008, at 12 noon.

UNIVERSITY OF HELSINKI
Department of Psychology
Studies 51: 2008

Supervisor:	Professor Riitta Hari Brain Research Unit Low Temperature Laboratory Helsinki University of Technology Finland
Reviewers:	Associate Professor Moshe Bar Visual NeuroCognition Laboratory Massachusetts General Hospital Harvard Medical School USA
	Professor Jari Hietanen Department of Psychology University of Tampere Finland
Opponent:	Professor Rafael Malach Department of Neurobiology Weizmann Institute of Science Rehovot Israel

ISSN 0781-8254
ISBN 978-952-10-4651-3 (pbk.)
ISBN 978-952-10-4652-0 (PDF)
http://ethesis.helsinki.fi/

Picaset Oy
Helsinki 2008

CONTENTS

ABSTRACT

The neural basis of visual perception can be understood only when the sequence of cortical activity underlying successful recognition is known. The early steps in this processing chain, from retina to the primary visual cortex, are highly local, and the perception of more complex shapes requires integration of the local information. In Study I of this thesis, the progression from local to global visual analysis was assessed by recording cortical magnetoencephalographic (MEG) responses to arrays of elements that either did or did not form global contours. The results demonstrated two spatially and temporally distinct stages of processing: The first, emerging 70 ms after stimulus onset around the calcarine sulcus, was sensitive to local features only, whereas the second, starting at 130 ms across the occipital and posterior parietal cortices, reflected the global configuration.

To explore the links between cortical activity and visual recognition, Studies II–III presented subjects with recognition tasks of varying levels of difficulty. The occipito-temporal responses from 150 ms onwards were closely linked to recognition performance, in contrast to the 100-ms mid-occipital responses. The averaged responses increased gradually as a function of recognition performance, and further analysis (Study III) showed the single response strengths to be graded as well.

Study IV addressed the attention dependence of the different processing stages: Occipito-temporal responses peaking around 150 ms depended on the content of the visual field (faces vs. houses), whereas the later and more sustained activity was strongly modulated by the observers' attention. Hemodynamic responses paralleled the pattern of the more sustained electrophysiological responses.

Study V assessed the temporal processing capacity of the human object recognition system. Above sufficient luminance, contrast and size of the object, the processing speed was not limited by such low-level factors. Taken together, these studies demonstrate several distinct stages in the cortical activation sequence underlying the object recognition chain, reflecting the level of feature integration, difficulty of recognition, and direction of attention.

TIIVISTELMÄ

Nähöhavaintojen hermostollisen perustan ymmärtäminen edellyttää näönvaraiseen tunnistamiseen liittyvien aivoalueiden ja niiden ajallisten aktivaatioketjujen selvittämistä. Ketjun varhaisissa vaiheissa, verkkokalvolta ensimmäiselle näköaivokuorelle, kukin hermosolu käsittelee vain pientä osaa näkökentästä, ja laaja-alaisten hahmojen havaitseminen edellyttää näiden osatietojen yhdistämistä. Väitöskirjan osatyössä I selvitettiin, miten näkötiedon käsittely aivoissa etenee paikalliselta tasolta laaja-alaiselle tasolle. Koehenkilöiden aivotoimintaa seurattiin magnetoenkefalogragialla (MEG) heidän katsoessaan ärsykkeitä, joiden osat oli sijoiteltu joko satunnaisesti tai yhtenäiseksi kuvioksi. Ärsykkeiden käsittely alkoi ensimmäisen näköaivokuoren alueella noin 70 ms niiden esittämisestä. Tässä vaiheessa kuvio- ja satunnaisärsykkeitä käsiteltiin samalla tavalla eli vain paikallisten piirteiden tasolla. Noin 50 ms myöhemmin takaraivo- ja päälaenlohkojen takaosat reagoivat voimakkaammin kuvio- kuin satunnaisärsykkeisiin heijastaen paikallisten piirteiden yhdistelyä kokonaisuuksiksi.

Osatöissä II ja III tutkittiin monimutkaisempien hahmojen – kasvojen – tunnistamisen hermostollista perustaa muuntelemalla tunnistustehtävän vaikeutta. Kasvokuvia käsiteltiin takaraivolohkoissa 100 ms:iin saakka riippumatta tehtävän vaikeudesta, mutta noin 50 ms myöhemmin takaraivo- ja ohimolohkojen raja-alueet aktivoituivat sitä voimakkaammin, mitä paremmin koehenkilö onnistui tunnistamisessa. Tiedonkäsittelyn myöhempi vaihe liittyi siis läheisesti kasvojen tunnistamiseen.

Osatyössä IV tarkasteltiin näkötiedon käsittelyvaiheiden riippuvuutta koehenkilön tehtävästä, jolla säädeltiin tarkkaavuuden suuntautumista. Ensimmäisten 150 ms aikana näkötiedon käsittely takaraivo- ja ohimolohkoissa ohjautui esitetyn ärsykkeen mukaan tehtävästä riippumatta, mutta tätä myöhempi käsittely muovautui olennaisesti tehtävän mukaan.

Osatyössä V todettiin, ettei ihmisen näönvarainen tunnistamiskyky tietyn kynnystason yläpuolella riipu havaintokohteen kirkkaudesta, kontrastista tai koosta. Tulos vastaa tiettyjen ohimolohkon alueiden ominaisuuksia ja tukee käsitystä näiden alueiden tärkeästä roolista näönvaraisessa tunnistamisessa. Kaiken kaikkiaan väitöskirjatyö osoittaa näkötiedon käsittelyssä useita erillisiä vaiheita, jotka heijastavat ärsykepiirteiden yhdistelyä, hahmojen tunnistamista ja tarkkaavuuden suuntautumista.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I *Tanskanen T*, Saarinen J, Parkkonen L and Hari R: From local to global: Cortical dynamics of contour integration. Journal of Vision 2008, in press.

II *Tanskanen T*, Näsänen R, Montez T, Päällysaho J and Hari R: Face recognition and cortical responses show similar sensitivity to noise spatial frequency. Cerebral Cortex 2005, 15: 526–534.

III *Tanskanen T*, Näsänen R, Ojanpää H and Hari R: Face recognition and cortical responses: Effect of stimulus duration. Neuroimage 2007, 35: 1636–1644.

IV Furey ML, *Tanskanen T*, Beauchamp MS, Avikainen S, Uutela K, Hari R and Haxby JV: Dissociation of face-selective cortical responses by attention. Proceedings of the National Academy of Sciences of the United States of America 2006, 103: 1065–1070.

V Näsänen R, Ojanpää H, *Tanskanen T* and Päällysaho J: Estimation of temporal resolution of object identification in human vision. Experimental Brain Research 2006, 172: 464–471.

The publications are referred to in the text by their roman numerals.

ABBREVIATIONS

ANOVA	analysis of variance
BEM	boundary element model
BOLD	blood oxygenation level dependent
dSPM	dynamic statistical parametric map
EEG	electroencephalography
EOG	electro-oculogram
ERP	event-related potential
fMRI	functional magnetic resonance imaging
GLM	general linear model
IT	infero-temporal
LGN	lateral geniculate nucleus
LOC	lateral occipital complex
MEG	magnetoencephalography
MNE	minimum norm estimate
MRI	magnetic resonance imaging
NSF	noise spatial frequency
PET	positron emission tomography
PO	parieto-occipital
RF	receptive field
RMS	root mean square
RSVP	rapid serial visual presentation
RT	reaction time
SD	standard deviation
SEM	standard error of mean
SQUID	superconducting quantum interference device
VEF	visual evoked field
VEP	visual evoked potential
V1	visual area 1, primary visual cortex
V*n*	visual area *n* (2–6)

INTRODUCTION

To survive, an organism needs information about its surroundings. Even the simplest creatures seek nutrition, and the more elaborate ones analyze complex social settings to adapt their behavior. Through the course of evolution, organisms have developed sophisticated tools, senses, to obtain relevant information. Different senses react to changes in physical quantities such as temperature, pressure, or concentration of various chemicals. Moreover, the ability to detect the intensity and wavelength of electromagnetic radiation in a certain range has turned out to be extremely useful: The radiation emitted by the sun is reflected and absorbed in characteristic ways by different materials and objects. Therefore, changes in visible light can inform the organism about both what there is in the surroundings and where those things are located. In daily life, such analyses appear easy: For example, we can recognize a wide variety of constructions as suitable for sitting, and should we decide to sit down, the process almost invariably results in positioning us in the correct location. Similarly, we can recognize familiar persons in a crowd and, should we want to meet them, we can navigate through the crowd with limited collisions. However, the required visual information needs to be derived solely from the pattern of photons hitting the photoreceptors inside our eyes. Obviously, the pattern reflected e.g. from the faces of two different persons can be rather similar, whereas the patterns reflected from one person under two different conditions can be highly different. Although humans usually make the correct inferences without effort, for computers this remains a challenge. The goal of visual neuroscience is to understand how humans succeed in such analyses and, more generally, how light reflected from the surroundings is utilized to guide behavior.

The visual system comprises numerous levels. Processing of information starts in the complex neural network of the retina and then proceeds via the deep brain structures to the cerebral cortex. This thesis focuses on characterizing visual processing at the cortical level.

Knowledge on how the cortex is organized to process visual information first came from observing the consequences of brain injuries. For example, damage to the most posterior part of the brain typically leads to loss of sight in some parts of the visual field, or in the worst case, to blindness. Limited lesions in the temporo-occipital cortex, in turn, can have a very different effect: The patient might lose the ability to recognize other people’s faces while retaining most other visual skills. Nevertheless, the extent and sites of such lesions can not be controlled for, which sets limits on how much can be gained from clinical studies. Controlled lesions in experimental animals have been informative, but when the primary goal is to understand the functioning of the human visual system, the potential differences across species pose additional questions.

Within the past decades, rapid developments in brain imaging have made it possible to study the functioning of the healthy human brain under experimental control. Some of these new methods, such as functional magnetic resonance imaging (fMRI), can give detailed information about the brain areas activated under specific conditions. Some other methods, such as electroencephalography (EEG) and magnetoencephalography (MEG), can characterize temporal sequences of brain activity. The core of this thesis comprises a set of studies that utilized MEG to unravel the temporal dynamics of cortical activity underlying various levels of visual information processing. The MEG measurements were complemented by fMRI and behavioral techniques.

Each neuron at the early stages of the visual system processes only a minor fraction of the visual field. Therefore, output from the early local processes needs to be integrated for perception of global patterns. Where and when this process occurs in the human visual cortex was approached in Study I.

After the first steps of visual cortical processing, i.e. analyzing contours, edges and elementary shapes, processing advances to more abstract levels where the neurons and cortical areas are sensitive to complex shapes, for example faces and other object categories. However, the linkage between different temporal processing stages and visual recognition has not been characterized earlier (Studies II and III).

Although an answer to the basic recognition question “who is this person” is either correct or incorrect, the averaged cortical responses linked with recognition increase gradually as the visibility of a face improves. Whether the graded averaged responses reflect truly graded or rather on-off –type processing at the level of single responses was tackled in Study III.

Our visual field is seldom occupied only by an isolated face or object against a neutral background. It is thus necessary to select information for further processing, which, however can be done only after the information has been analyzed to some degree. The time course of such selection is not known, and in particular, it remains unclear whether the first processing stages are automatic, i.e. independent of the subject's attention, or whether attention can modulate visual information processing at even the earliest stages of processing (Study IV).

Although the early stages in visual processing are highly sensitive to basic stimulus attributes, such as luminance and contrast, we can identify persons under very different lighting conditions and from different distances. To what degree our capacity to recognize faces and other objects is dependent on such low-level parameters as brightness, contrast and size was characterized in Study V.

The following presentation will start by a brief overview of the human visual system. Next, the methods applied for studying cortical functions will be introduced. After this background information, the methods and results specific to Studies I–V will be reported, followed by brief discussions. Finally, all results will be discussed in a more general context.

BACKGROUND

Overview of the visual system

Precortical processing

Electromagnetic radiation at wavelengths of 400–700 nm can be detected by the human visual system and is therefore defined as visible light. The light entering the eye is refracted by the cornea and lens to form a sharp picture on the retina. There, the photoreceptors capture photons and through a complex chemical cascade, phototransduction, convert and amplify them into neural signals. Among the photoreceptors, three types of cones are sensitive to different wavelengths and thus form the basis of color vision. Rods, in turn, are sensitive to low light intensities and therefore enable vision under dim light. All experiments in this thesis were performed with grayscale images under lighting levels clearly sufficient for cone vision.

From photoreceptors, the signal is transmitted via bipolar cells to ganglion cells. Horizontal and amacrine cells modulate the retinal transmission. Ganglion cell axons, forming the optic nerve and optic tract, convey the signal from the retina to the lateral geniculate nucleus (LGN) of the thalamus. Different types of ganglion cells synapse in LGN: The responses of the large magnocellular cells are fast and sensitive to low contrasts, whereas the small parvocellular cells convey the signal more slowly, but carry more information about wavelength and spatial detail (De Monasterio et al., 1975; Kaplan et al., 1986; Merigan et al., 1993). The role of a third, koniocellular pathway, remains less well understood (Hendry et al., 1994; Hendry et al., 2000). From LGN, the signal is conveyed to the primary visual cortex (V1).

Besides the geniculo-cortical pathway described above, another important visual pathway is formed by retinal neurons that project to the pulvinar, a nucleus of the thalamus, and then further to the cortex. Part of these connections run via the superior colliculus in the midbrain. These pathways are important for orienting towards salient stimuli (Robinson et al., 1992; Kaas et al., 2007).

Cortical processing

The primary visual cortex, V1, is located around the calcarine sulcus in the occipital lobe. The V1 neurons show specificity across multiple dimensions, the most global organizing principle being retinotopy (Tootell et al., 1982), i.e. the preservation of spatial relations of the environment (neighboring cells respond to neighboring parts of the visual field). Signals from the two eyes first project to distinct cortical ocular dominance columns at V1 (Hubel et al., 1962, 1968), and other V1 neurons thereafter combine signals from both eyes, responding to binocular disparity that forms the basis of stereoscopic vision. In their seminal work, Hubel and Wiesel labeled V1 neurons by their response properties as simple and complex cells. Simple cells respond best to contrast bars (Hubel et al., 1959), i.e. transitions from dark to bright or vice versa, with a specific spatial phase. Complex cells, in turn, are phase-independent (Hubel et al., 1962). The cells show selectivity for orientation, spatial frequency (Schiller et al., 1976c, 1976b), motion direction (Schiller et al., 1976a) and wavelength (Thorell et al., 1984). Such response properties constitute an efficient representation, a sparse linear code, for natural images (Olshausen et al., 1996). A recent development in the characterization of V1 has been the observation that more spatial integration occurs at the level of V1 neurons than was thought when the spatially limited classical receptive fields were discovered (Angelucci et al., 2002; Cavanaugh et al., 2002b, 2002a; Schwabe et al., 2006).

Beyond V1, visual processing involves a large proportion of the human cortex, possibly up to one fourth of the cortical surface (Van Essen, 2003; Wandell et al., 2007). Distinct visual areas, illustrated in Figure 1, have been identified by such criteria as retinotopy, other functional properties, histology, and intercortical connections (Felleman et al., 1991; Tootell et al., 2003; Grill-Spector & Malach, 2004). Some areas appear to be important for processing of e.g. visual motion (Zeki, 1974; Watson et al., 1993; Tootell et al., 1995) or possibly color (Zeki, 1973; Lueck et al., 1989; Tootell et al., 2004). Nevertheless, the exact roles of most areas, as well as how they work together, remain poorly understood.

A major organizing principle proposed by Mishkin et al. (1983) divides the visual cortical system into two components: the ventral pathway for object vision and the dorsal pathway for spatial vision. The division was originally based on monkey studies and subsequently demonstrated in humans as well (Haxby et al., 1991). A modified version labels the pathways by behavioral significance, proposing that the ventral pathway serves perception and the dorsal pathway (motor) action (Goodale et al., 1992).

The organizing principles of the human ventral visual stream have been one of the most intensively studied topics in visual neuroscience for the past ten years, but the question remains open. However, some basic findings and discrepancies regarding the ventral shape, object and face processing areas will be reviewed below, with focus on human studies.

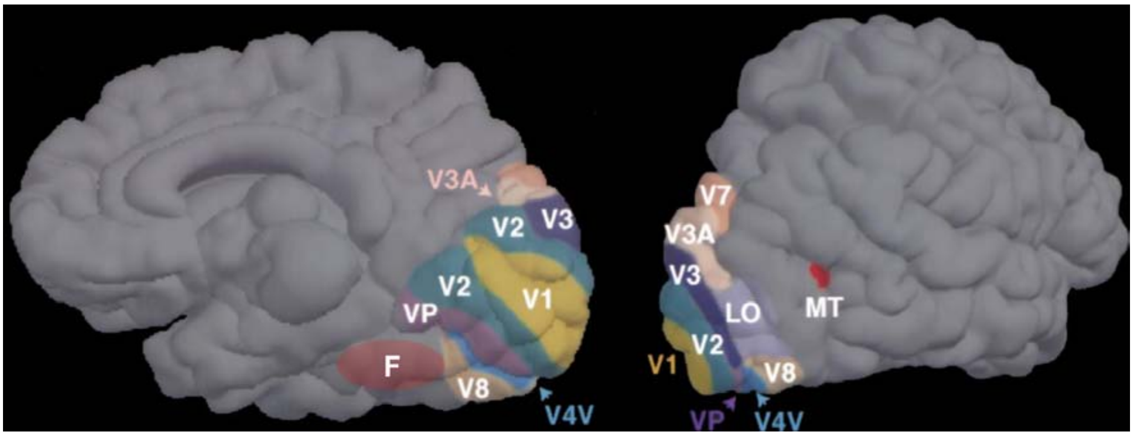

Figure 1. Medial and postero-lateral views of the human brain. V1–V8, VP: retinotopic visual areas; MT – middle temporal area (visual motion processing); LO – lateral occipital area (objects); F – fusiform gyrus (faces). Adapted from Tootell et al. (1998) with permission from Elsevier.

Perception of shapes and objects

Edges and contours

Although the V1 neurons have spatially limited classical receptive fields (RFs), processing of spatially extended contours and edges can occur already at this stage of cortical analysis: For example, orientations of stimuli that surround the classical RF modulate responses of monkey V1 cells to stimuli consisting of small line segments (Knierim et al., 1992; Kapadia et al., 1995; Kapadia et al., 1999, 2000). The difference between classical and extraclassical RF effects is reflected in the time course of processing: The initial 50-ms response to oriented line segments in monkey V1 is affected only by the portion of the stimulus that falls the neuron’s classical RF, whereas a later sustained response, starting around 80–100 ms, emerges when the receptive field is on the edge between two surfaces, defined by a difference in texture orientation (Lamme, 1995; Zipser et al., 1996; Rossi et al., 2001). Besides V1, processing of even simple contours involves higher-order visual areas as well (Altmann et al., 2003; Kourtzi et al., 2003).

Intermediate shapes

Within the ventral pathway, the complexity or abstractness of shape selectivity of the neurons gradually increases (reviewed in e.g. Ungerleider et al., 2003). Although some neurons may show selectivity for complex shapes even at the level of V1, such neurons are more prominent in V2 and, in particular, V4 (Hegde et al., 2006, 2007). The V4 neurons respond best to for example concentric and radial gratings (Gallant et al., 1993; Gallant et al., 1996; Wilkinson et al., 2000). Neurons in area TE show a further increase in the complexity of critical features, but the critical features are typically less complex than what is required to define a specific object (Tanaka, 1997).

Objects

Visual object-processing cortex can be distinguished by identifying regions that respond more strongly to objects than to textures or scrambled objects (Malach et al., 1995). An area in the lateral occipital cortex, showing such selectivity, has been labeled as the lateral occipital complex (LOC; Malach et al., 1995; Grill-Spector, Kourtzi et al., 2001). Sensitivity to different object categories, in turn, can be found in the ventral occipito-temporal cortex. Distinct cortical patches respond best to such stimuli as faces (Sergent et al., 1992; Kanwisher et al., 1997; McCarthy et al., 1997), places (Aguirre et al., 1998; Epstein et al., 1998), body parts (Downing, Jiang et al., 2001) or tools (Martin et al., 1996; Beauchamp et al., 2002). The functional principles underlying these observations remain open: According to Haxby et al. (2001), object categories are represented as distributed patterns of activity in the ventral occipito-temporal cortex, as opposed to areas specialized in the processing of single object categories (Kanwisher et al., 1997).

Nevertheless, even if some regions are specialized in the processing of a single visual category, such regions are likely to exist for only a limited number of categories (Downing et al., 2006). It is plausible that the spatial organization of neurons sensitive to different visual categories could reflect the spatial scales and retinal eccentrities typical for each category (Levy et al., 2001; Hasson et al., 2002; Malach et al., 2002); examples are presented below in the Faces section.

As opposed to early visual areas, the ventral object selective cortices appear relatively insensitive to such low-level features as contrast (Avidan et al., 2002), and they seem to represent object shape rather than contours (Hasson et al., 2001; Kourtzi et al., 2001; Andrews et al., 2002; Lerner et al., 2002). Correspondingly, these areas even respond to shapes defined by illusory contours (Mendola et al., 1999; Kourtzi et al., 2000). Instead of low-level features, activity in the ventral object areas correlates with subjects’ recognition performance (Grill-Spector et al., 2000; James et al., 2000; Bar et al., 2001; Kleinschmidt et al., 2002; Grill-Spector, 2003; Grill-Spector, Knouf et al., 2004). Furthermore, activity in these regions does not need to be stimulus-driven; it can be elicited by imagery (Ishai et al., 2000; O'Craven et al., 2000; Ishai et al., 2002), or reflect the awareness of a face even when the face is occluded (Hulme et al., 2007).

Faces

Among the different categories of visual objects, faces have been most extensively studied. Before the era of brain imaging, neuropsychological studies described a condition in which the ability to recognize facial identity is disrupted, despite normal ability to recognize visual objects (Hecaen et al., 1962). This condition, prosopagnosia, is typically caused by bilateral lesions in the ventral occipito-temporal cortex (Damasio et al., 1982). Correspondingly, face-sensitive activity in the temporo-occipital region has been found in single-unit recordings in monkeys (Bruce et al., 1981; Perrett et al., 1982; Desimone et al., 1984), intracranial recordings in humans (Allison, Ginter et al., 1994), PET and fMRI studies (Sergent et al., 1992; Clark et al., 1996; Kanwisher et al., 1997; McCarthy et al., 1997), and EEG and MEG recordings (Bentin et al., 1996; Sams et al., 1997; Halgren et al., 2000).

Besides the ventral occipito-temporal cortex, visually presented faces activate a number of other cortical regions that presumably serve distinct functions. A model based on neuropsychological studies suggested a major distinction between processes important for the recognition of face identity vs. recognition of facial expressions and speech-related movements of the mouth (Bruce et al., 1986). Building on this distinction, Haxby and coworkers (Haxby et al., 2000; Haxby et al., 2004) proposed a model of the human neural system for face perception (Figure 2). Areas in the occipitotemporal cortex are assumed to form a core system for face perception, with separate modules for invariant vs. variant aspects of faces, i.e. identity vs. eye gaze and expression. This distinction is supported by both monkey (Perrett et al., 1985; Hasselmo et al., 1989) and human studies (Hoffman et al., 2000). The identity module comprises areas with different levels of abstraction: Areas in the inferior occipital gyrus are affected by physical changes in the faces, whereas the fusiform gyrus shows selectivity for facial identity irrespective of physical attributes (Rotshtein et al., 2005). The extended system comprises a number of cortical and subcortical structures related, but not limited, to processing the various aspects of information that can be derived from faces. For example, faces carry information about emotions and intentions, direction of attention (gaze), and speech (lip movements).

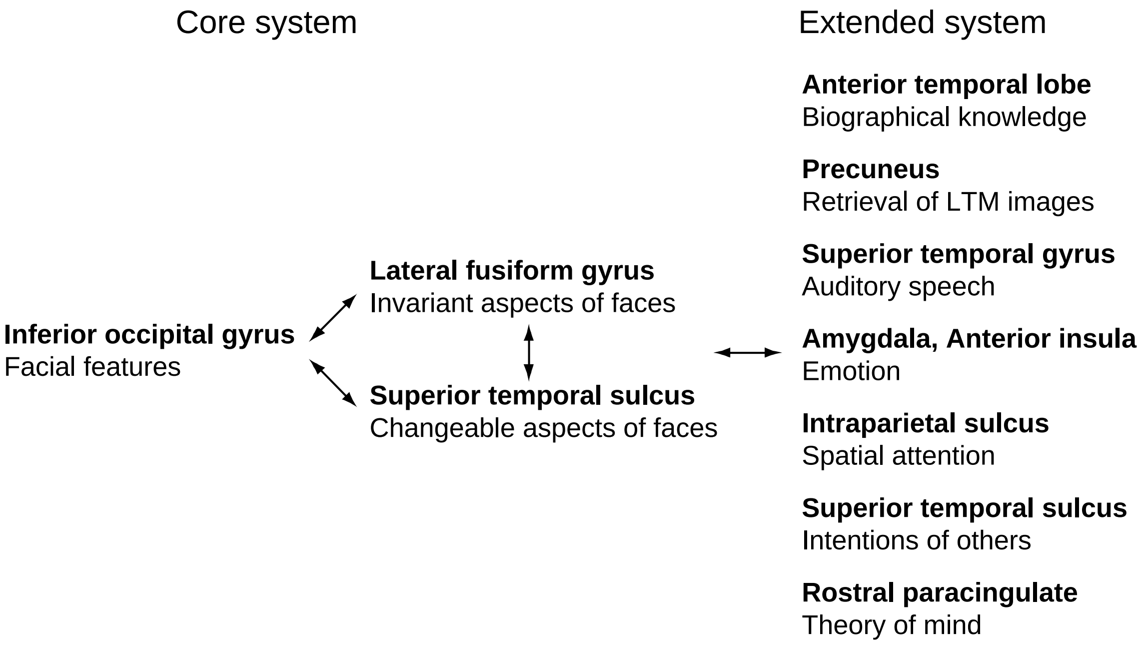

Figure 2. Model of the human neural system for face perception according to Haxby et al. (2000; 2004). LTM –long term memory.

The debate about the organizational principles of the ventral occipito-temporal cortex is particularly relevant for the processing of faces. The data by Haxby et al. (2001) suggest that, like other visual categories, faces are represented in a distributed cortical network, whereas Kanwisher et al. (1997) claim that the fusiform gyrus contains a unit specialized in the processing of faces only. At columnar level, the claimed face-specific region could consist of columns that all respond maximally to faces and sub-maximally to other categories, or, alternatively, strictly face-specific columns with some intermittent columns selective to other categories or features. A recent high-resolution fMRI study (voxel size 1mm x 1 mm x 1mm) favored the latter view (Grill-Spector et al., 2006), but has raised some methodological concerns regarding how selectivity of responses was tested (Baker et al., 2007; Simmons et al., 2007). The results also seem to conflict with another recent study that assessed the same issue with fMRI-guided single-unit electrophysiology in monkeys (Tsao et al., 2006). The question thus remains open for intensive study.

Besides object category, two other principles have been suggested to underlie the functional organization within the ventral occipito-temporal cortex. The eccentricity model attempts to apply the principle of retinotopic organization, prominent in the early visual areas, to the organization of the object-sensitive cortex. Since faces are typically observed from such a distance that they occupy only a small fraction of the visual field, and humans typically focus their gaze on faces, it is conceivable that an area that processes faces samples mainly the center of the visual field. On the other hand, buildings, for instance, might typically occupy a relatively large part of the visual field (Levy et al., 2001; Hasson et al., 2002; Malach et al., 2002). The eccentricity model implies that the typical retinal size should be taken into account in experimental setups; most of the present literature is based on experiments where everything from flowers to buildings has been presented in equal size.

The expertise model, in turn, stresses another difference between faces and other visual categories: It is usually not sufficient to just categorize a percept as a face, as is the case with many other objects; one needs to go further and identify the face as the face of a particular individual. In this sense, most humans are experts in face recognition, and the cortical areas showing strong responses to faces could in fact be areas specialized in within-category identification. In support of this model, the same areas that respond strongly to faces in the general population are activated in experts when they perform within-category discrimination in their field of expertise (Gauthier et al., 1999; Gauthier et al., 2000).

Regarding the interpretations of the infero-temporal responses to faces, the eccentricity and expertise models need not exclude each other, whereas the generic expertise-module vs. true face-module views seem to be in conflict (Grill-Spector, Knouf et al., 2004; Kanwisher et al., 2006; Gauthier et al., 2007; McKone et al., 2007). Nevertheless, whether or not some cortical region or regions are specialized solely in the processing of faces, a network of multiple cortical regions is required for the analysis of various aspects of facial information (Fairhall et al., 2007; Barbeau et al., 2008; Ishai, 2008). This is emphasized by the finding that many congenitally prosopagnosic subjects show normal fMRI responses in the cortical regions most strongly associated with face recognition (Hasson et al., 2003; Rossion et al., 2003; Avidan et al., 2005; Sorger et al., 2007).

Dynamics of visual processing

The latencies of the neural responses to visual stimuli depend on such properties as luminance, size, and contrast, and the effects might differ across areas. Precise generalizations across studies are therefore difficult to make, and differences (e.g. in lengths of neural connections) across species further complicate the situation. In awake macaque, the first LGN cells respond to light flashes 15–18 ms after stimulus onset, and the activity reaches its maximum around 25 ms in the magnocellular and around 35 ms in the parvocellular layers. In V1, activation starts at 25–30 ms and peaks around 45 ms (Schroeder et al., 1998). Importantly, neurons within single areas respond with highly variable latencies, and different visual areas, on the other hand, are activated with overlapping latencies (Figure 3; Bullier et al., 1995; Nowak et al., 1997; Schmolesky et al., 1998; Schroeder et al., 1998).

In humans, the exact starting time of visual cortical processing is hard to measure non-invasively, and for reasons mentioned above, dependent on the exact stimulus conditions. Rough estimates can be obtained by scaling the macaque latencies by a factor of 3/5 (Saint-Amour et al., 2005). In visual evoked potentials (VEPs) recorded from human scalp, early posterior responses to onsets or reversals of simple patterns typically peak around 75 ms (Jeffreys et al., 1972) and are followed by a complex spatiotemporal sequence of activity. It is commonly agreed that responses showing selectivity for complex feature combinations, for example object categories, peak at 150–200 ms (Jeffreys, 1989; Allison, McCarthy et al., 1994; Thorpe et al., 1996). Even earlier responses have been claimed to show category specificity (Linkenkaer-Hansen et al., 1998; Braeutigam et al., 2001; Liu et al., 2002), but it is difficult to distinguish effects of low-level visual properties from true category effects at such latencies. Responses to single images can continue at least up to one second after stimulus onset (Puce et al., 1999; Henson et al., 2003). The sequence of visual cortical activity will be approached in greater detail in the General Discussion of this thesis.

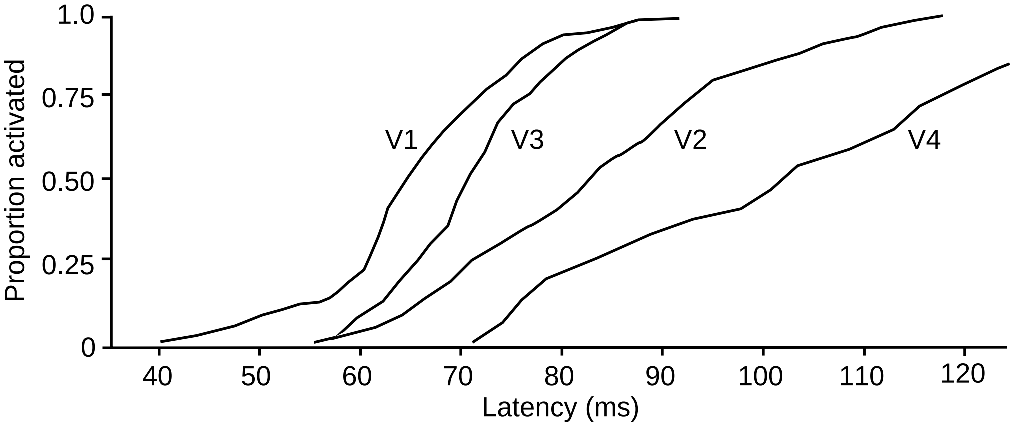

Figure 3. Cumulative distributions of response onset latencies in visual areas V1–V4 of anesthetized macaque monkeys. Adapted from Schmolesky et al. (1998) with permission from the American Physiological Society.

Top-down influences

The preceding paragraphs have described a feed forward sequence of stages in visual information processing, starting from early stages sensitive to simple, local features to later stages sensitive to increasingly abstract dimensions. The true picture is naturally more complex: Most cortical connections are reciprocal, and all visual areas thus receive signals from higher-order areas as well (Felleman et al., 1991). For example, inactivation of visual area MT disturbs processing in V1, V2 and V3 (Hupe et al., 1998), and even the most distant frontal areas are connected to the early visual areas (Catani et al., 2002). The top-down signals from frontal regions could modulate processing in the object-sensitive cortices via attentional selection (Corbetta et al., 2002) or contextual facilitation (Bar, 2004). A recent model proposes that the fast signals conveyed via the magnocellular pathway might initiate in the prefrontal cortex a "rapid guess" about the most likely interpretations of the image, facilitating processing of the bottom-up information in the temporo-occipital object areas (Bar, 2003). Initial data in support of this model has been obtained (Bar et al., 2006).

Methods for studying human cortical processing

Since the 1990's, the techniques for non-invasive study of brain function have developed dramatically, leading to a boom in human systems neuroscience. The studies comprising this thesis are based on MEG measurements of the electric activity of neurons, and on fMRI measurements of its metabolic consequences (hemodynamics). These methods will be reviewed below.

Magnetoencephalography

Electric currents in the neurons are accompanied by magnetic fields. Although these fields are extremely weak compared with e.g. the static magnetic field of the Earth (different by eight orders of magnitude), a cluster of synchronously active neurons can generate a magnetic field strong enough to be detected outside the head. The bulk of the extracranial fields most likely reflects post-synaptic currents in the apical dendrites of pyramidal cells in the cortex: these dendrites lie in parallel and the post-synaptic currents are long-lasting enough to allow summation (Hari, 1990; Okada et al., 1997).

MEG is most sensitive to currents that are tangential to the surface of the head, which favors detection of neural activity in the cortical sulci, since the pyramidal neurons are perpendicular to the cortical surface. However, only narrow stripes of cortex are perfectly tangential to the local curvature of the skull, and in practice, a more relevant limiting factor may be source depth (Hari, 1998; Hillebrand et al., 2002). The

MEG measurement and analysis techniques, reviewed by e.g. Hämäläinen et al. (2002) and Baillet et al. (2001), will be briefly described in the Methods section.

The main advantages of MEG are millisecond-scale temporal resolution and complete noninvasiveness. Compared with other brain imaging methods, e.g. fMRI, the subject can sit in a relatively open space and in complete silence, which provides more flexibility for experimental design. MEG's ability to identify the site of origin of brain responses depends on the specific conditions. A rough estimate of active cortical areas can be obtained directly from the measured magnetic field pattern, and under optimal conditions, the active brain area can be estimated with millimeter accuracy. Precise location is obtainable when a focal cortical patch is active at a given time, or simultaneously active areas are separate enough (in the order of centimeters). In the case of the visual cortex, multiple nearby areas are often active in parallel (Nowak et al., 1997; Schmolesky et al., 1998; Schroeder et al., 1998; Barbeau et al., 2008). Therefore, localizing activity at the scale of e.g. different retinotopic areas is challenging (Uutela et al., 1999; Stenbacka et al., 2002).

Electroencephalography (EEG) primarily measures the same neuronal activity as MEG. The disadvantage of scalp EEG compared with MEG is its sensitivity to errors caused by differently conducting tissues between the cortical neurons and EEG electrodes, which blur the electric potential distribution but not the magnetic field (Hämäläinen et al., 1993). On the other hand, EEG is more sensitive to deep brain activity and to cortical sources that are oriented radially with respect to head surface. Therefore, simultaneous recording of MEG and EEG can be beneficial. In patients subject to brain surgery, EEG is in some cases recorded intracranially, permitting an excellent signal-to-noise ratio and a more straightforward interpretation of the cortical areas generating the measured signals (Lesser et al., 2005).

The brain's magnetic field was for the first time detected with an induction coil magnetometer in 1968 and with a SQUID magnetometer in 1972 (Cohen, 1968, 1972). Following the introduction of whole-head neuromagnetometers (e.g. Ahonen et al., 1992), MEG has been applied widely to study the human sensory and motor systems, oscillatory brain activity, higher functions such as language processing and action observation, and various brain disorders such as epilepsy and stroke (reviewed in e.g.

Hari, 1990; Del Gratta et al., 1999; Hari et al., 2000; Kakigi et al., 2000; Salenius et al., 2003; Kaneoke, 2006; Salmelin et al., 2006; Shibasaki et al., 2007).

Visual evoked fields (VEFs) were first recorded by Brenner et al. (1975) and Teyler et al. (1975). Subsequent studies have traced the cortical processing of several basic visual attributes, such as spatial frequency (Williamson et al., 1978; Aine et al., 1990), color (Fylan et al., 1997) and motion (Anderson et al., 1996; Uusitalo et al., 1997). Besides studies on the visual system, higher-order VEFs have been extensively characterized in the context of language processing, e.g. picture naming and reading (Salmelin et al., 1994; Salmelin et al., 1996). Visual MEG studies most relevant for the topics of this thesis are discussed in the appropriate sections.

Functional magnetic resonance imaging

Magnetic resonance imaging (MRI) employs nuclear magnetic resonance (NMR) for obtaining non-invasively three-dimensional images of structures of the human body. In MEG studies, anatomical MRIs of the brain are used for constraining and visualizing the generators of the measured MEG signals. Besides anatomy, MRI can be employed to obtain information about brain function as well. The most commonly applied form of functional magnetic resonance imaging (fMRI) is based on local changes in the oxygenation level of blood (blood oxygenation level dependent signal, BOLD; Ogawa et al., 1990). Intensity of the BOLD signal is coupled with the amplitude of local field potentials, which in turn reflect post-synaptic activity in the neuronal dendrites (Logothetis et al., 2001; Logothetis et al., 2004). The coupling between action potentials and the BOLD signal seems variable, depending on correlations in firing rates of neighboring neurons (Nir et al., 2007).

FMRI can distinguish activation of nearby cortical areas at higher spatial resolution than MEG, and it is also sensitive to the deep brain areas that are difficult to reach with MEG. However, since coupling between neural activity and blood oxygenation is slow and not necessarily constant (Henson et al., 2002), fMRI does not provide precise information about the timing of cortical events. A significant proportion of the studies of the visual temporo-occipital cortex, summarized in the preceding section, is based on the BOLD fMRI method.

AIMS OF THE STUDY

This thesis work focused on the processing of visual contours, objects and faces in the human brain, with emphasis on the temporal sequence of cortical activity underlying visual recognition. The studies employed time-accurate MEG, in combination with fMRI and behavioral methods, with the specific aims

1) to characterize how cortical processing of visual scenes advances from local elements to global contours (Study I),

2) to investigate the neural basis of visual recognition by seeking cortical responses that show tight correlation with face recognition performance when the visibility of the faces was manipulated by superimposing noise on the faces (Study II) and limiting the time the faces were available for inspection (Study III),

3) to study whether the single cortical evoked responses to faces, comprising the graded averaged responses observed in Studies II and III, are graded or on-off –type (Study III),

4) to characterize the automaticity vs. task dependence of cortical processing of faces and objects (Study IV), and

5) to study how the speed of visual recognition is affected by such basic visual attributes as luminance, contrast and size of the perceived objects (Study V).

OVERVIEW OF THE METHODS

Magnetoencephalography

Measurement

Whole-scalp neuromagnetic signals were measured in a magnetically shielded room, while the subject was sitting with the head leaning against the measurement helmet of the Vectorview™ 306-channel magnetometer (Neuromag Oy., Helsinki, Finland; currently Elekta Neuromag Oy). The helmet-shaped detector array comprises 102 identical SQUID-based triple sensor units, each housing two planar first-order gradiometers and one magnetometer. The two gradiometers of each unit measure orthogonal tangential derivatives of the magnetic field component approximately normal to the head surface.

MEG signals were filtered to 0.1–172 Hz and sampled at 600 Hz. Signals were averaged over a time window starting 0.2–0.3 s before and ending 1.0 s after the onset of the stimulus. Horizontal and vertical electro-oculograms were recorded for on-line rejection of epochs contaminated by blinks and eye movements.

Before the MEG recordings, four head position marker coils were attached to the subject's scalp. The positions of the coils and of three anatomical landmarks were measured with a 3D digitizer. At the beginning of each recording block, the position of the subject’s head with respect to the sensor array was determined by feeding current to the marker coils and localizing the coils based on the signals measured by the MEG sensors. This information was used afterwards for combining the sources of the measured neuromagnetic signals with the subject’s structural MRIs by identifying the anatomical landmarks in the MR images.

Analysis

The effect of environmental noise on the averaged MEG signals was first attenuated by projecting out noise sub-spaces calculated on the basis of ambient noise measured in the absence of the subject (Parkkonen et al., 1999). Alternatively (for single trial analysis in

Study III and calculation of the minimum norm source estimates in Study I), the signal-to-noise ratio was improved by Signal Space Separation (SSS; Taulu et al., 2004). In all studies, the responses were digitally low-pass filtered at 35 Hz, and a 200–300-ms prestimulus baseline was applied for amplitude measurements.

The averaged evoked responses of each subject were first screened for experimental effects. Since planar gradiometers pick up the strongest sensor signals just above a locally activated brain area, the regions with strongest signals can be readily used as the first guesses of the activated brain areas.

In Study I, the neural generators of the MEG responses were estimated by noise-normalized minimum norm estimation (MNE). The current estimates were calculated using the "MNE Software" package (developed by M. Hämäläinen, http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofMNE.php).
Anatomical MRIs were processed with the FreeSurfer software (http://www.nmr.mgh.harvard.edu/martinos/userInfo/data/sofFreeSurf.php).
A boundary element model (BEM) along the inner skull surface was used as the volume conductor. To obtain the source point set, the gray and white matter border was tessellated (Dale et al., 1999) and decimated to a 5-mm dipole grid. Dipole amplitudes were determined by ℓ^2 MNE that incorporates depth weighting and loose orientation constraints (Lin et al., 2006). The estimated dipole strengths were then normalized by their noise sensitivity, i.e., by the estimated currents due to the noise in the measurement (sometimes referred to as dynamic statistical parametric maps, dSPMs; Dale et al., 2000). The estimates were thereafter transformed to an atlas brain with surface-based morphing (Fischl et al., 1999) and averaged across subjects.

In Study II, the responses that showed clear dependence on the experimental manipulations were modeled with equivalent current dipoles, assuming a spherical volume conductor that was fitted to the posterior part of the intracranial volume (Hari et al., 1986; Sarvas, 1987). These current dipoles served two functions: First, they acted as spatial filters to integrate data from the sensors to yield a better signal-to-noise ratio for the sources they were modeling. Second, they roughly indicated the sites of cortical areas where the observed effects took place. The dipole locations and orientations were searched by a least-squares fit to a subset of sensors around the local signal maxima. The dipoles found in the experimental conditions with the strongest signals were then inserted into a multidipole model that was used to reveal source strengths as a function of time in all conditions. The source coordinates were transformed into standard brain coordinates. This alignment was based on an affine transformation of the individual brains (Woods et al., 1998), followed by a refinement with a non-linear elastic transformation (Schormann et al., 1996) to match a standard atlas brain (Roland et al., 1994).

In Studies III and IV, the quantitative analysis was based on a selection of sensors showing significant responses, and then averaging the rectified signals across these sensors. The noise level was determined separately for each subject and each sensor by estimating the standard deviation of all timepoints within the prestimulus period. The signals were then examined to identify the sensors in which the evoked responses exceeded the baseline variability by at least 8 SDs. This set of sensors was then used to analyze responses for all stimulus conditions.

Functional magnetic resonance imaging

In Study IV, gradient-echo, echo-planar imaging (repetition time 2.5 s, echo time 40 ms) was used to measure the BOLD responses in a GE 3-tesla scanner at the National Institute of Health, Bethesda, USA. Whole brain volumes, comprising 40 contiguous 3.5-mm thick sagittal slices, were obtained.

Time-series data were analyzed on a voxel-by-voxel basis using multiple regression (Friston et al., 1995). The strength the of response for each stimulus condition was taken as the estimated beta-weights associated with each regressor in the general linear model (GLM). Selected contrasts between responses to different task conditions were calculated as effects of interest. Regions of interest were defined as areas showing significant responses relative to the scrambled image control condition ($Z > 5.6$, $p < 10^{-8}$) with a minimum volume of 7 contiguous voxels. Mean timeseries for the selected subregions of cortex were obtained for each subject. The mean strength of response to each stimulus condition, expressed as percent changes in the signal, was calculated for the subregions of each subject. The significance of differences between responses was tested using a random effects repeated measures analysis of variance (ANOVA) with planned comparisons.

Psychophysics

In Study V, a rapid serial visual presentation method (RSVP) was combined with a staircase algorithm to determine how much time per image frame was needed for the identification of a target with a given probability (79%). The task of the observer was to identify the target stimulus shown in the RSVP sequence.

A staircase method was used to determine the threshold frequency: After three consecutive correct responses the temporal frequency of image sequence was increased approximately by a factor of 1.26 (0.1 $\log_{10}$-units), and after each incorrect response the temporal frequency was decreased by the same factor (Wetherill et al., 1965). The algorithm adjusted the presentation frequency close to a level at which the probability of correct responses was 0.79. On each run, the threshold frequency was obtained as the mean of eight reversals. The threshold frequency was measured three times for each experimental condition in a randomized order.

Subjects

In all brain imaging studies, 6–10 healthy members of the laboratory personnel served as subjects; the total number of subjects that contributed to this thesis was 23. Males and females were represented in roughly equal numbers, and the age range was 22–46 years. All subjects had normal or corrected-to-normal visual acuity. The MEG recordings had prior approval by the Ethics Committee of the Helsinki and Uusimaa Hospital District.

Stimulus presentation

In the MEG experiments, stimulus presentation was controlled by Presentation® software (http://www.neurobs.com/) run on a personal computer. The images were displayed on a rear projection screen by a data projector (VistaPro™, Christie Digital Systems Inc., Cypress, CA, USA) based on Digital Light Processing™ and hosting three digital micromirror panels (for details on the projector performance, see Packer et al., 2001). The experiments were run in the standard VGA mode (resolution 640 x 480 pixels, frame rate 60 Hz, 256 gray levels). The subjects viewed the screen binocularly at distances from 88 to 120 cm in a dimly lit room. In the fMRI experiment (Study IV), a mirror was placed in front of the subjects' eyes to allow them to see the rear projection screen outside the scanner.

For behavioral responses in the MEG experiments, the subjects used response pads. Light was fed to the pads via optical fibers, and the subject's finger lifts were detected by changes in the light returned to the receiver unit via another fiber (slightly different apparatuses were used in different experiments, based on either cutting the light beam or reflecting back the light by the skin).

In Study V, stimulus presentation was controlled by custom-made software. The stimuli were presented on a cathode-ray tube (CRT) monitor. Responses were given by placing the mouse cursor on an appropriate icon on the display.

EXPERIMENTS

Higher-order visual areas are involved in contour processing (Study I)

Neurons in the early visual cortices have spatially limited receptive fields and can thus process only local elements of visual information. Processing of more global patterns, starting from contours, requires integration of local information. The neural mechanisms underlying such processes have been under extensive research since the early 1990's (Hess et al., 2003; Roelfsema, 2006; Sasaki, 2007). For example, stimuli surrounding the classical receptive field of monkey V1 cells can modulate the cell's responses and affect the perceived contrast (Knierim et al., 1992; Kapadia et al., 1995). The initial responses in V1 are affected only by the stimulus part landing on the neuron's classical receptive field, whereas a later sustained response is affected by more global properties of the visual stimulus (Lamme, 1995; Zipser et al., 1996). In humans, FMRI studies have pinpointed cortical areas underlying such effects (Altmann et al., 2003; Kourtzi et al., 2003; Dumoulin et al., 2007). However, the techniques used in humans so far have not unraveled the time courses of the cortical processes underlying integration of visual information.

The integration mechanisms can be studied by comparing the processing of arrays of Gabor elements that either form or do not form a global contour (Field et al., 1993). Because the classical receptive field of a single cell in the primary visual cortex, sensitive to the appropriate spatial frequency and orientation, should cover approximately one element, the earliest cortical responses to the contour and no-contour stimuli should be similar. The responses should separate only when the neurons start to integrate local information. The purpose of this study was to characterize the cortical dynamics of contour processing by comparing responses to these two stimulus types. Specifically, the aims were to find out when the cortical responses to contour and no-contour stimuli would separate, to identify the sites of contour-specific activity, and to study the effect of the local orientation of elements on cortical responses.

Methods

Square arrays of Gabor elements (Figure 4) were presented to the center of the visual field. In the *no-contour* stimuli, all elements were oriented randomly and positioned pseudorandomly without overlap. In the *contour* stimuli, a proportion of the elements were oriented and positioned to form an easily detectable double circle.

In the first condition, the orientations of the contour elements were tangential to the circle, and in the second the orientations were radial. In the third condition, the contour comprised only the lower left quadrant of a full circle.

The stimuli were presented once every 2.5 s for 0.5 s with abrupt onsets and offsets. Each recording block consisted of contour trials, no-contour trials, and catch trials presented in random order. A minimum of 120 responses were collected for each stimulus type. In the catch trials, the stimuli were replaced by a question mark, and the subject reported with a finger lift whether a contour or a no-contour stimulus had been presented in the preceding trial. After the MEG experiment, discrimination reaction times were measured to get a coarse measure of task difficulty for the different contour types.

Results

The first cortical responses were detected at the posterior MEG channels close to the midline, exceeding the prestimulus noise level at 69 ± 2 ms (mean ± SEM across subjects) and peaking at 85 ± 3 ms. For tangential contours, the difference between responses to contour and no-contour stimuli exceeded the 2 SD baseline noise threshold at 130 ± 9 ms, i.e. 61 ± 9 ms after the emergence of the first responses ($p < 0.001$; Figure 4). The contour-sensitive effects reached their maximum at 274 ± 35 ms, and the differences vanished at 600–700 ms, i.e. 100–200 ms after the offset of the stimulus. For radial contours, the contour-sensitive activity started at 164 ± 17 ms, and for the quadrant contours at 149 ± 13 ms.

With the minimum norm source modeling, the first cortical responses were detected around the calcarine sulcus (Figure 4). A second peak occurred around 125 ms in the same cortical region. The source estimates of these two first responses did not differ between the contour and no-contour conditions. Instead, clear contour-sensitive activity peaked around 215 ms at several locations in the posterior parieto-occipital (PO) cortex. The most prominent differences in responses to contour vs. no-contour stimuli were observed in the medial surface around the PO sulcus and precuneus. Additional differences were observed in occipital areas spanning from cuneus to the middle occipital gyrus in the left hemisphere, and from middle occipital gyrus to superior occipital gyrus in the right hemisphere. The spatial patterns of the current estimates for the contour vs. no-contour effects were rather similar for all contour types.

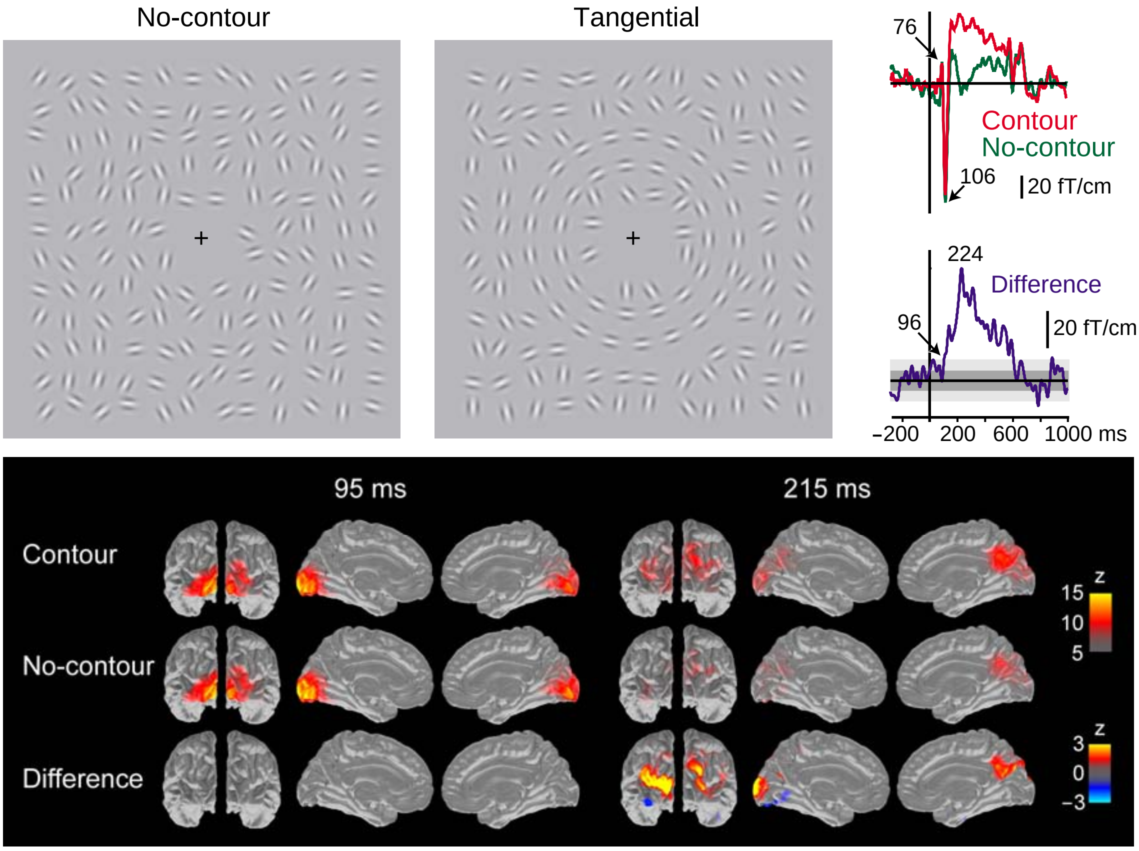

Figure 4. No-contour and tangential contour stimuli (top left), examples of MEG responses of Subject 2 (top right) and cortical activity averaged across all 8 subjects (bottom).

The reaction times for discriminating between the contour *vs.* no-contour stimuli were the shortest, about 550 ms, to the tangential stimuli, and ~50 ms longer to radial and quadrant stimuli ($p < 0.01$ for both categories).

Discussion

The results demonstrated early responses that were identical to contour and no-contour stimuli, and later responses (> 130 ms) that were sensitive to contours and thus to the global form. The early responses were generated around the primary visual cortex, whereas the later responses arose from the more lateral and dorsal occipital and parietal regions.

When the local elements were oriented tangentially to the global contour, the difference emerged on average at 130 ms after the stimulus onset, in good agreement with previous evoked potential studies on visual segmentation and grouping (Bach et al., 1997; Han et al., 2001; Fahle et al., 2003; Ohla et al., 2005; Pei et al., 2005; Mathes et al., 2006). The site of the present contour-sensitive effects around the PO sulcus, suggested to be the functional homologue of macaque area V6/V6a (Colby et al., 1988; Galletti et al., 1991; Portin et al., 1998; Pitzalis et al., 2006) is involved in a wide range of cortical processes (Jousmäki et al., 1996; Vanni et al., 2001; Cavanna et al., 2006). Interestingly, this area is the most prominent generator of the MEG alpha rhythm (Hari & Salmelin, 1997), the level of which is inversely related to the saliency of perceived visual objects (Vanni et al., 1997). Furthermore, the activity in the PO region covaries with the number of attention switches between local and global elements of visual objects (Fink et al., 1997). Patients with lesions in the parieto-occipital cortex, typically bilaterally, fail to perceive more than one object at a time, having difficulty integrating elements of the visual field and switching between them (Rizzo, 1993).

The more postero-lateral contour-sensitive activations spanned from cuneus to middle and superior occipital gyri and thus across several functional areas, with the largest overlap in area V3a. The human V3a is involved in processing of visual objects at a level independent of the type of the visual cues defining an object (Grill-Spector et al., 1998).

The most consistent contour-sensitive cortical effects were produced by the full circular contours in which the local elements were oriented tangentially to the global contour; the effects were weaker when the local orientations were radial or when tangential elements formed only quadrant contours. Accordingly, finding a contour among randomly distributed elements is most efficient when the local elements are aligned along the contour (Field et al., 1993; Kovacs et al., 1993; Saarinen et al., 1997;

Bonneh et al., 1998; Pettet et al., 1998; Saarinen et al., 2001). Besides matching local and global orientations, the observed response might reflect enhanced processing of concentric patterns (Wilkinson et al., 2000; Kurki et al., 2004; Dumoulin et al., 2007).

Face recognition and temporo-occipital responses are tightly correlated (Study II)

In EEG and MEG recordings, a face-sensitive response peaks 140–180 ms after the stimulus onset (often labeled as N170 or M170). The face sensitivity of this response has been demonstrated by showing that it is at least twice as strong for faces than for any control stimuli tested so far, including textures and a large variety of objects (Bentin et al., 1996; George et al., 1996; Sams et al., 1997; Allison et al., 1999; Halgren et al., 2000). Study II aimed at demonstrating the importance of cortical processes underlying M170 for face recognition in a more direct manner: The recognizeability of faces was manipulated parametrically to demonstrate possible correlations between the response strength and successfulness of recognition. The parametric manipulation of recognizeability was achieved by masking the face stimuli by noise with different spatial frequencies (Näsänen, 1999); face recognition is sensitive to noise at certain spatial frequencies, but independent of noise at frequencies higher or lower than the critical frequency (Fiorentini et al., 1983; Hayes et al., 1986; Peli et al., 1994; Costen et al., 1996).

Methods

The stimuli were combinations of synthetic facial images and spatial noise masks with 10 different noise spatial frequencies (NSFs; see Figure 5). A set of eight synthetic face images was adopted from Näsänen (1999). The center NSFs ranged from 2 to 45 c/image, corresponding to 0.28–6.3 c/deg during stimulus presentation. For all noisy faces, the signal-to-noise ratio was 0.5. The signal-to-noise ratio was selected so that face recognition was difficult at NSFs centered on the critical band for face recognition (11–16 c/image), without too much interference at low and high NSFs (Näsänen, 1999).

Besides the noisy faces, ‘low-contrast’ and ‘high-contrast’ noiseless faces were presented.

The stimuli were presented once every 2.5 s for 0.5 s, with abrupt onsets and offsets. All stimulus categories were presented within the same blocks in a random order. Subjects were asked to respond with a finger lift to images representing the target person, indicated before the MEG recording. Before the MEG recordings, the subjects went through behavioral training until they were able to recognize the target person with close to 100% accuracy.

Results

A prominent 100-ms posterior response (M100) was observed in all six subjects, and it was adequately modeled with a current dipole in the occipital region close to the midline (Figure 5). The smallest responses were elicited by the images with the lowest NSF and by the low-contrast noiseless faces. Around the NSF of 5.6 c/image, the responses started to increase and they reached the maximum on average at 20.5 c/image. The responses then decreased again for the highest NSFs. The peak latency of this response was shortest, 85 ± 4 ms (mean ± SEM) for 5.6 c/image, and it then systematically prolonged as a function of NSF, being 113 ± 2 ms for the highest NSF. The responses were statistically significantly ($p < 0.005$) stronger to high-contrast than low-contrast noiseless images.

Another prominent response peaked at 130–180 ms (M170), with sources in the temporo-occipital or posterior temporal cortex. The images with low NSF elicited strong signals. At NSF ≥ 4 c/image the amplitudes started to decrease, reaching the minimum at NSFs of 8–16 c/image, and then increased again for the highest NSFs. The latencies of the temporo-occipital responses were shortest for the high-contrast noiseless faces, on average 144 ± 5 ms. Responses to the low-contrast noiseless faces and faces with noise peaked 10–20 ms later.

Face recognition was close to perfect for stimuli with the lowest and highest NSFs but poor for those with the middle NSFs (11.0–16.0 c/image). The shapes of the face recognition and temporo-occipital source strength curves, plotted as a function of NSF, resembled each other. In line with this, the M170 responses correlated statistically significantly with recognition performance ($r = 0.89$; $p < 0.001$).

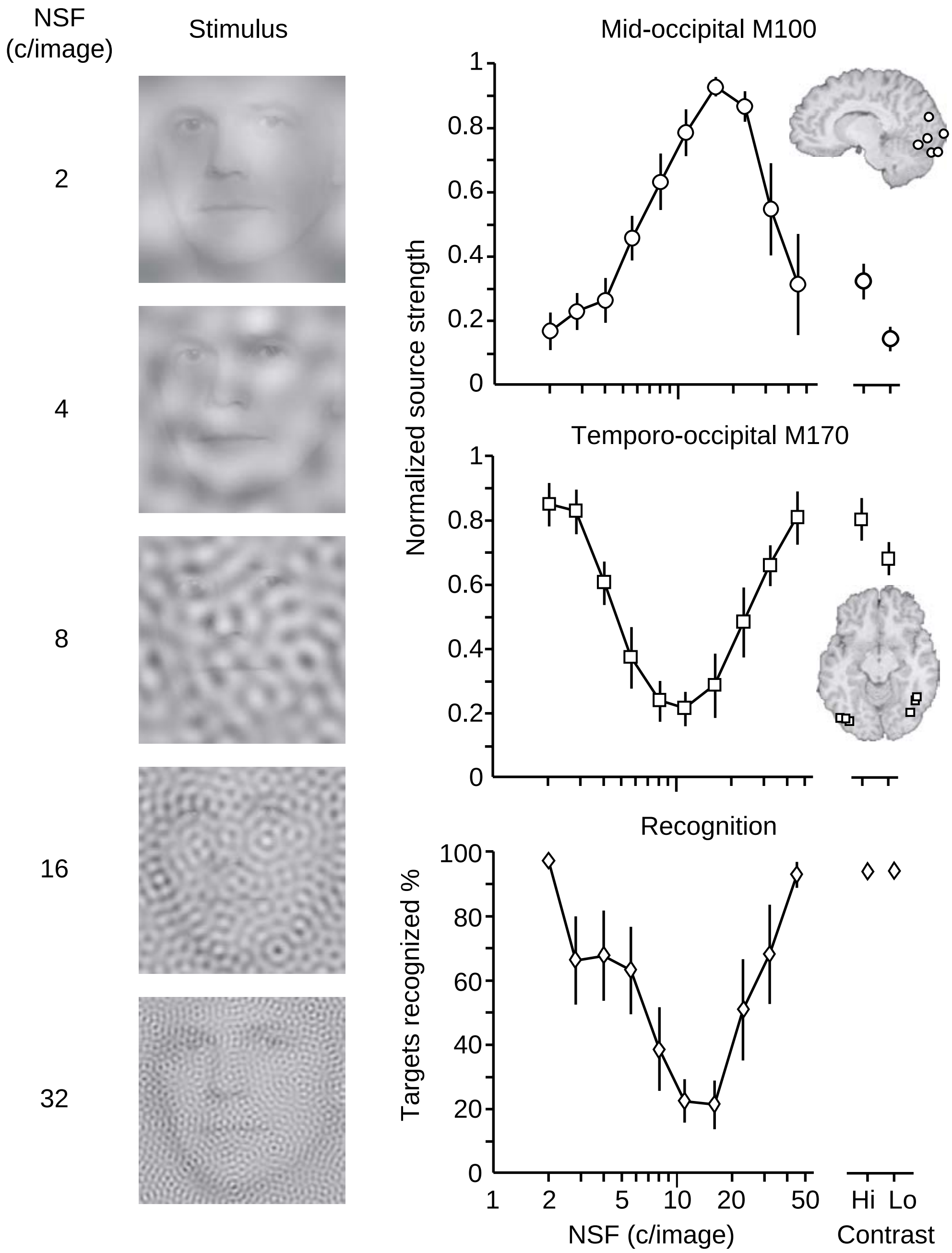

Figure 5. Examples of stimuli (left) and the cortical responses and recognition performance of all subjects (right). Error bars indicate SEM.

To further clarify the functional roles of the mid-occipital M100 and the temporo-occipital M170 responses, we measured responses of two subjects to the noise masks presented alone. The mid-occipital M100 was very similar to plain noise and face + noise stimuli; this result is in line with the very small M100 responses elicited by plain faces. On the contrary, the M170 responses were strongly affected by the presence of a face: For the face + noise stimuli, M170 showed the same U-shaped modulation as was observed in the main experiment, but for plain noise, the response was small and almost independent of NSF.

Discussion

Two cortical responses showed distinct dependence on the NSF. First, the early mid-occipital responses at 70–120 ms (M100) were smallest for low NSFs, increased until central NSFs, and decreased again for the highest NSFs. Second, the temporo-occipital responses at 130–180 ms, likely to correspond to the face-selective 170-ms response (N170/M170) reported previously in both EEG and MEG literature, were strong for images with low and high NSFs that were easy to recognize but tiny for images with NSFs of 8–16 c/image that were difficult to recognize. Thus, behavioral face recognition and the M170 showed similar sensitivity to NSF. A control experiment supported the interpretation that the M100 mainly depended on the spatial frequency of the noise mask, and that the M170 depended on the visibility of a face.

Band-pass characteristics of the mid-occipital 100-ms responses have previously been reported in several studies. Musselwhite and Jeffreys (1985) observed that occipital evoked potentials to black-and-white gratings peak around 4 c/deg. Fylan et al. (1997) demonstrated that MEG responses to chromatic gratings show band-pass characteristics, with strongest responses at 1–2 c/deg. The low-frequency attenuation in the mid-occipital responses most likely reflects the receptive field sizes of V1/V2 neurons which are inversely proportional to the optimal spatial frequency of each neuron; consequently, fewer neurons are needed to cover the stimulus area for low than high spatial frequencies. Indeed, the low-frequency attenuation in the MEG responses can be compensated for by increasing the stimulus area (Fylan et al., 1997). On the other hand, the attenuation of the cortical responses (and behavioral contrast sensitivity) to high spatial frequencies is due to a number of optical and neural factors (e.g., De Valois et al., 1990).

Because face recognition is most sensitive to frequencies at which the mid-occipital regions respond strongly, a possible explanation for the observed tuning would be the signal-to-noise ratio of the representation of the face + noise images in V1/V2 cortex, which in turn could affect the output from these regions to the higher-order areas that are important in face and object recognition. In other words, the results would fit with the idea that the NSF sensitivities of face recognition and face-selective areas reflect properties of the earlier visual areas. However, the center absolute spatial frequency (c/deg) of the critical band for face recognition increases with decreasing stimulus size but more slowly than in inverse proportion. For example, a four-fold increase in viewing distance and thus in absolute spatial frequencies results in a decrease of the relative spatial frequency (c/face) critical for face recognition from 11 to 8 c/face width (Näsänen, 1999). Thus, the critical band is neither scale invariant nor fixed in absolute spatial frequency.

The tight binding between the M170 temporo-occipital responses and face recognition is in good agreement with previous findings on face-sensitive responses in the same region and time window (Allison, Ginter et al., 1994; Bentin et al., 1996; Sams et al., 1997). The cortical generators and the functional interpretation of this response will be discussed in detail in the General Discussion, together with the results from the other face processing studies.

Temporo-occipital responses to faces are graded, not on-off type (Study III)

Study II showed that the amplitude of the face-sensitive M170 response correlates strongly with recognition of facial images that are manipulated by masking them with noise of different spatial frequencies. However, faces also elicit cortical responses beyond 200 ms. The relationship between these long-latency face-sensitive responses and face recognition has remained uncharacterized so far. The first aim of Study III was to obtain further support for the earlier finding on the correlation of M170 strength and face recognition performance (Study II), and to find out whether the more task- and familiarity-dependent responses peaking after 200 ms would also correlate with recognition performance. This time the parametric manipulation of face recognizeability was obtained with temporal instead of spatial masking, i.e. by varying the time the face was available for visual observation before it was replaced by a noise mask (display duration, DUR).

The gradually varying averaged cortical responses observed in Study II, as well as by Tarkiainen et al. (2002), Jemel et al. (2003) and Horovitz et al. (2004), could either be averages of varying proportions of all-or-nothing type single responses (either full response or no response), or averages of single responses that are graded as such. The second aim of this study was to unravel whether the neural responses contributing to M170 are categorical or graded by nature. The improved signal-to-noise ratio of modern neuromagnetometers, together with recent noise-attenuation techniques (Taulu et al., 2004), permits approaching this issue by analyzing the variability of the amplitudes of single responses.

Methods

The face stimuli were black-and-white photographs of 21 men who were unknown to the subjects prior to the study. To identify face-sensitive MEG responses, intact and phase-scrambled faces were presented before the main experiment. The subject’s task was to respond by lifting the right index finger to images of the target person who was indicated before the recording

For the main experiment, the subjects were trained to recognize a set of six faces (out of the original 21) before the recordings. The faces replaced a continuously present noise mask (Gaussian noise with ca. 9 squares per face width) once every 2.5 s for 17–200 ms. A question mark, appearing randomly after 10% of the trials, prompted the subject to report which of the six prelearned persons had been presented just before. Behavioral recognition performance was evaluated before and after the MEG recordings by presenting 24 face images for each DUR; the settings were otherwise the same, but the subjects responded to each stimulus.

Results

The elicited cortical responses clearly differed between the intact and phase-scrambled faces in two time-windows. Therefore, it was possible to define, separately for each subject, the sets of sensors on which the responses to faces exceeded 10 times the standard deviation (SD) of prestimulus baseline noise: “M170 sensors” at 125–200 ms and “M300 sensors” at 200–500 ms. The absolute (rectified) response amplitudes were then averaged across both the selected sensors and all subjects.

At 17- and 33-ms DURs, the M170 responses were similar in amplitude. At 50-ms DUR, a distinct M170 started to emerge, increasing in amplitude at 67–83 ms DURs and reaching its maximum at the DUR of 200 ms. The mean ± SD peak latency of M170 was 180 ± 9 ms at DURs from 67 to 200 ms (range across subjects 167–197 ms); there were no statistically significant differences between the response latencies at these DURs. Responses on the M300 sensors were the smallest at the shortest DUR, and then increased until the DUR of 200 ms. The mean peak latency of M300 across all subjects and conditions was 275 ± 57 ms.

The recognition performance was at chance level (1/6 = 17%) at DURs of 17–33 ms, and then improved at DURs of 50–200 ms so that at 200-ms DUR more than 90% of the faces were recognized correctly. At the group level, correlations between recognition performance and M170 and M300 amplitudes were 0.97 and 0.98 ($p < 0.001$ for both). Individual correlations between recognition and M170 varied from 0.79 to 0.98 ($p < 0.05$ for all subjects), and correlations between recognition and M300 from 0.65 to 0.98 ($P < 0.05$ for six out of nine subjects). To compare the increment rates of recognition performance and cortical responses as a function of DUR, a cumulative normal distribution was fitted to each subject’s data using a least-squares criterion. A statistical comparison indicated that the increment rate of the recognition performance resembled more that of M170 than of M300.

For quantitative analysis of the single-trial response amplitudes, the peak amplitude for each trial was read as the maximum amplitude within 125–200 ms after stimulus onset. The mean M170 amplitude and the corresponding amplitude variability (computed as the mean of individual SDs, therefore showing the variability of response strengths as a function of DUR *within* the subjects) increased very similarly as a function of DUR, and their relationship was linear with a high correlation (Figure 6).

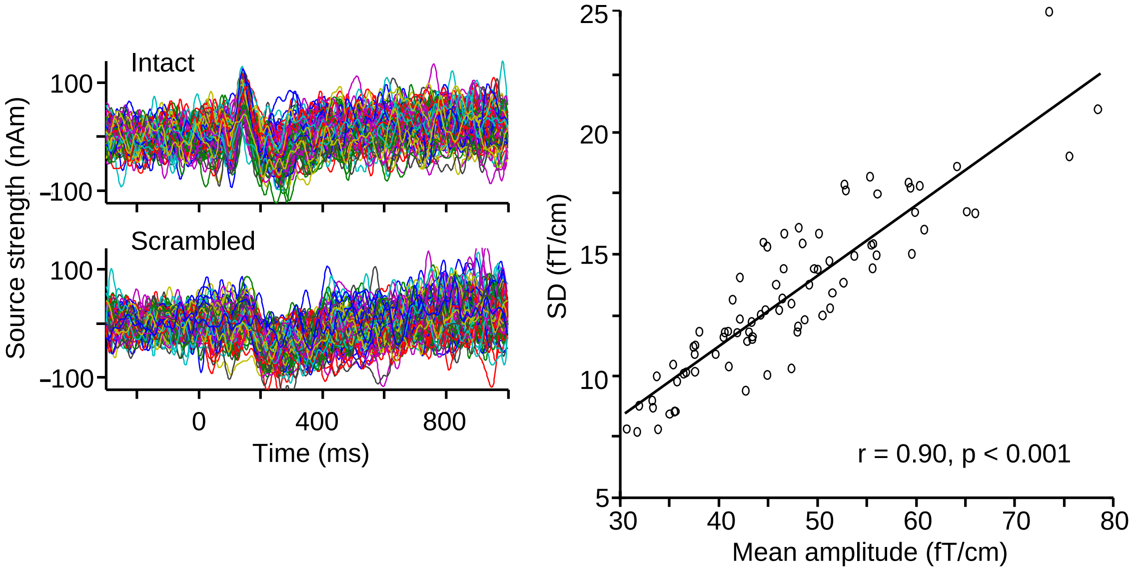

Figure 6. Around 100 overlaid single responses of Subject 9 to intact and scrambled faces (left) and dependence of response variability on response strength (all subjects, right).

Discussion

These results demonstrate a high correlation between face recognition and the cortical activity at 140–200 ms as well as 250–350 ms after the start of stimulus presentation. One possible interpretation for the sub-maximal M170 and M300 responses encountered under backward-masked conditions is that the mask interrupts processing in the temporo-occipital cortex (Rolls & Tovee, 1994; Rolls, Tovee et al., 1994; Kovacs et al., 1995). Alternatively, the mask might affect processing at earlier processing steps and thus lead to decreased input to the processes underlying M170 and M300. The interpretation of these two responses will be further discussed in the General Discussion.

We also observed that the stronger the responses, the more variable (*within* subjects) were their amplitudes from trial to trial. This finding is in line with data from intracortical recordings: For example, when the contrast of a visual stimulus increases, the mean and variance of a neuron's firing rate in the cat and monkey visual cortex increase in parallel (Tolhurst et al., 1981; Tolhurst et al., 1983).

Quantification of the single-response variability allowed addressing whether the sub-maximal averaged responses are built up by a varying proportion of ”all-or-nothing” type single responses or whether they rather reflect an average of graded single responses. If the all-or-nothing alternative is correct, the responses at short DURs should be mainly minimal (and thus with small variance), at medium DURs both full and minimal responses should be elicited (large variance), and at long DURs mainly full responses would occur (again with small variance). Such an inverted-U type dependence of the signal variability on signal strength is in strong contrast to the finding of a strictly linear relationship between the response variability and response strength. Thus, the data indicate that the single responses are graded, increasing as a function of stimulus duration. The graded nature of the single responses goes well together with the finding that for identical stimulus presentations, the likelihood of successful recognition is higher during the trials with higher-amplitude responses (Liu et al., 2002).

Object processing proceeds from stimulus-dependent to task-dependent stage (Study IV)

Human beings receive a continuous flux of information via their multiple senses. As processing capacity is limited, irrelevant information needs to be filtered out. However, the selection process itself requires some preprocessing of information. To what degree all information is processed, and at what stage more selective processing starts, poses the classical question of selective attention. Filtering could occur early, based on sensory characteristics (Broadbent, 1958), or at the other extreme, all information could be processed up to semantic level (Deutsch et al., 1963). The neural substrates underlying the selection mechanisms remain a subject of intensive exploration (Näätänen, 1992; Corbetta et al., 2002; Posner et al., 2007).

This study investigated the effect of selective attention on cortical responses to faces and houses using stimuli that require attention to act directly on object selection. The stimuli were still images of superimposed face and house photographs in which the attended and unattended objects can be segregated based only on the contours that define those objects. Studies before this work had used lower-level stimulus differences such as spatial location (Wojciulik et al., 1998), eye of input (Tong et al., 1998), or differential movement (O'Craven et al., 1999). These differences have potentially affected the the input to object recognition operations, rather than the implementation of object recognition operations themselves.

The experiment tested whether preliminary processing of the unattended face or house occurs during the performance of the task. Such preliminary processing could play a role in segregating the defining contours of the attended and unattended objects. Alternatively, attention could bias object recognition to process only visual information that is consistent with the attended object category.

Methods

Electrophysiological responses were measured with MEG and hemodynamic responses with fMRI. The subjects viewed pictures of faces and houses and determined whether each contained a picture of the same face or house that was shown in the immediately preceding picture (Figure 7). Pictures with different views of the same person or house were used, so that subjects could not base their response on a simple pattern match. Images superimposed on a scrambled background had equivalent luminance, contrast, and spatial frequency spectra as the double-exposure pictures.

MEG responses to single-category pictures of faces or houses, and to pictures of intact faces or houses superimposed on phase-scrambled images of the other category (Fscr and Hscr), were used to identify and quantify category-selective responses at different latencies. fMRI responses to the same, single-category stimuli were used to locate the face-responsive and house-responsive regions of ventral temporal cortex and to measure the magnitude of category-selective, hemodynamic responses in these regions. Responses evoked while subjects viewed double-exposure stimuli and attended selectively to faces or houses (Fattn and Hattn, respectively) were used to quantify the effect of attention.

Results

Faces and houses elicited characteristic fMRI responses in distinct areas of the ventral temporal cortex (Figure 7). These responses showed strong attentional modulation: In the face-responsive fusiform region, the response to the unattended face (i.e, the condition where houses were attended, Hattn) was essentially identical to the response when a face was not present in the stimulus (Hscr). The responses to faces vs. houses on scrambled backgrounds (Fscr vs. Hscr) did not differ quantitatively from the responses when subjects attended to faces vs. houses in stimuli that contained intact images of both object categories (Fattn vs. Hattn; $p > 0.1$). The same was true for activity in the house-responsive regions in medial fusiform and inferior temporal cortex ($p > 0.1$). Other face-responsive regions, in inferior occipital and superior temporal sulcal cortex, showed no trend towards greater responses in the Hattn condition (face present but unattended) as compared with the Hscr condition (face not present; $p > 0.1$). Thus attention strongly suppressed the spatially-defined, category-related hemodynamic responses to the unattended faces and houses.

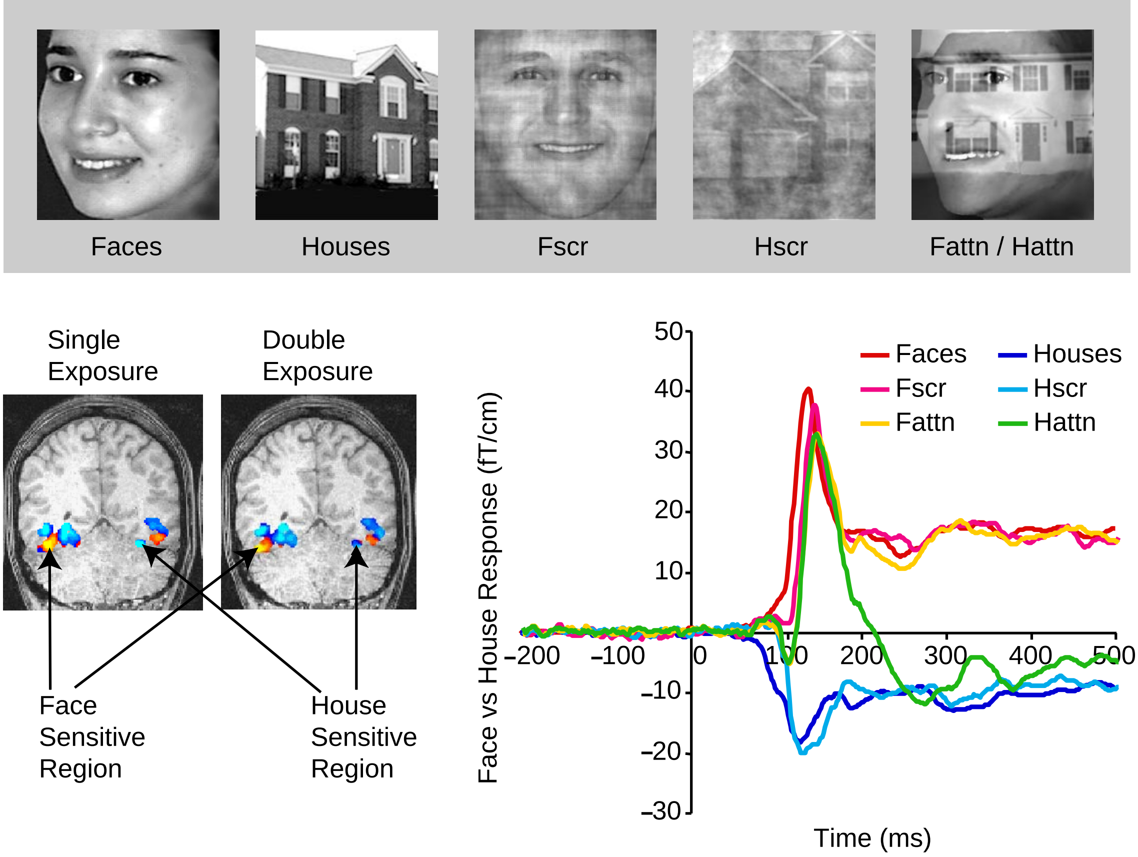

Figure 7. Examples of stimuli (top) and hemodynamic (bottom left) and neuromagnetic (bottom right) responses averaged across all subjects. Scr – scrambled background; attn – attended.

In MEG recordings, the strongest responses to faces peaked at 142 ± 18 ms (mean ± SD) for single-exposure faces (Figure 7). These responses correspond to M170, despite their somewhat shorter latency. The strongest responses to houses peaked 20–33 ms earlier. The face-house differences started in the single-exposure conditions at 103 ± 15 ms and in scrambled-background conditions at 116 ± 9 ms (Faces vs. Houses). Responses during selective attention to faces and houses were indistinguishable from each other until 194 ± 28 ms (Fattn vs Hattn, $p < 0.0001$ compared with the onset of the difference in single-exposure conditions), both displaying a strong face-selective M170. In contrast, late (> 235 ms) responses during selective attention to faces and houses differed markedly from each other and were indistinguishable from responses to faces and houses, respectively, in the single-exposure conditions. Early responses in the attention conditions, however, were not identical to the response to faces on a scrambled background (Fattn and Hattn vs. Fscr): A significant deflection towards the early house response peaked at 120 ms ($p < 0.005$), indicating some early processing of the house that also was not affected by attention but was then obscured by the slightly later and stronger, face-selective M170. The early face–house differences in the non-attention conditions were more restricted to bilateral posterior temporal sites, whereas the late face-house differences were more widely distributed including more anterior temporal sites on the right and a dorsal occipital site. Notably, the face–house differences started significantly earlier for single-exposure pictures than for pictures on scrambled backgrounds ($p < 0.02$). The earlier face-selective response with single-exposure pictures may, therefore, be due to the different spatial frequency spectra of face and house pictures, which were controlled in the pictures of intact faces or houses on scrambled backgrounds.

Discussion

The characteristic face- and house-selective fMRI responses (Kanwisher et al., 1997; McCarthy et al., 1997; Aguirre et al., 1998; Epstein et al., 1998; Haxby et al., 1999; Ishai et al., 1999; Halgren et al., 2000; Haxby et al., 2001) were strongly modulated by attention. Namely, response to the unattended stimulus was suppressed. In contrast, attention had no effect on the face-selective M170 response detected with MEG. Later (> 190 ms) MEG responses elicited by faces and houses, on the other hand, were strongly modulated by attention. The MEG result demonstrates automatic preprocessing of unattended faces, and possibly of houses as well, independently of the subjects' task. Such preliminary processing could play a role in segregating the defining contours of the attended and unattended objects. These results also illustrate that hemodynamic and electrophysiological measures of face-selective cortical processing complement each other. The hemodynamic signals reflect primarily late responses that can be modulated by feedback connections. By contrast, the early, face-specific M170 was not modulated by attention, suggesting that it reflects a rapid, feed-forward phase of face-selective processing.

At sufficient visibility, object identification speed is independent of contrast, size and luminance (Study V)

Temporal resolution of human visual perception depends greatly on the perceptual task and the kind of information that is being processed. For example, simple luminance flicker can be detected up to a rate of 50–100 Hz (Kelly, 1961), whereas the temporal resolution of visual attention has been estimated to be 4-8 Hz in motion tracking tasks (Verstraten et al., 2000). The aim of this study was to psychophysically estimate the temporal resolution of the object recognition system and the dependence of the resolution on various stimulus parameters.

A time estimate obtained from a backward masking or an RSVP experiment may directly reflect a limit set by the processing speed of the high-order cortical mechanisms mediating object identification. However, shortening the presentation time of an image probably reduces the signal-to-noise ratio of low-level neural representation and, hence, the amount of information. Thus, this kind of time estimate might equally well be limited by the amount of information available for the high-order identification mechanisms rather than the speed of processing. To clarify this question we estimated the temporal capacity of object identification by varying image contrast, size, mean luminance, and object type (single characters vs. faces). Contrast, size, and mean luminance, but not object type, have a more or less direct effect on the signal-to-noise ratio or timing of the low-level neural signal representation. We combined the RSVP method with a staircase algorithm to determine how much time per image frame is needed for the identification of a target with a given (79%) probability. The results were expressed as the temporal frequency of image presentation in an RSVP sequence (images per second, Hz).

Methods

The stimuli were either characters or facial photographs shown one at a time as a rapid sequence in the middle of the display. In the sequence of characters, the target could be any of twelve possible uppercase letters (A, B, C, D, H, E, K, N, R, U, V, and Z). The distractors were numerals (0–9). For the face sequences, there were six possible target faces and fifteen possible distractors (black and white photographs of men). The RSVP sequence contained ten images. Target position could be any other than the first or the last in the sequence with equal probability.

In addition to the RSVP threshold frequencies, simple reaction times for detecting a target as a function of luminance were measured. Luminance was varied between 0.05 and 50 cd/m^2 by inserting neutral density filters in front of the observers' eyes. In the simple reaction time task, the observer was required to press the left mouse button as soon as possible after the presentation of a letter stimulus. No recognition of the letters was required. In addition, temporal resolution for character identification as well as simple reaction time for detecting the onset of a letter stimulus as a function of stimulus mean luminance was measured.

Results

For both characters and faces, threshold frequency increased with contrast at very low contrasts but was nearly constant at medium and high contrasts (Figure 8). The critical RMS (root-mean-square) contrast above which temporal resolution was contrast invariant was around 0.05 for characters and around 0.1 for faces. Thus, for a broad range of contrasts the temporal resolution of object identification was contrast invariant. At medium and high contrasts, the temporal resolution for faces (around 10 Hz) was clearly lower than for characters (around 25 Hz).

For small characters (height < 1 deg) and facial images (width < 1.5 deg), the threshold frequency increased with increasing size but reached a plateau at medium sizes (Figure 8). Since the characters were band-pass filtered, the centre spatial frequency was inversely related to character size. For characters, temporal resolution of identification was constant when spatial frequency was smaller than 2.2 c/deg. Above 2.2 c/deg, the temporal resolution decreased with increasing centre spatial frequency.

The temporal resolution for the small characters clearly increased with luminance at low luminance levels, but improved little at higher luminance levels (2–40 cd/m^2). For large characters, the increase of temporal resolution was much smaller. Accordingly, reaction times shortened with increasing luminance. At higher luminance levels (2–40 cd/m^2), the decrease of reaction time became small. For large characters, the decrease of reaction time with increasing luminance was considerably smaller.

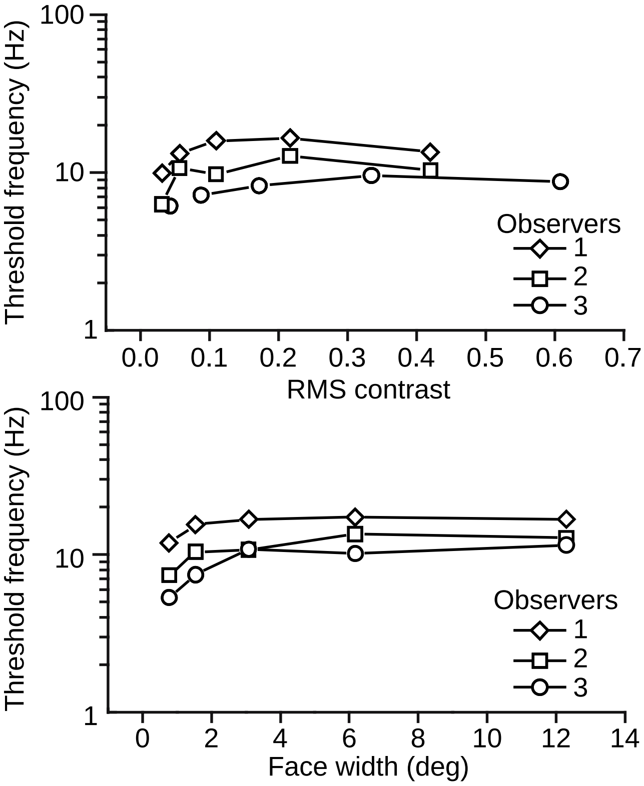

Figure 8. Effects of face contrast (top) and size (bottom) on processing speed**.**

Discussion

At high luminances, the temporal resolution of object identification remained unchanged at a wide range of contrasts and sizes. However, it was reduced at low contrasts, small sizes and low luminances. Furthermore, the processing speed was lower for facial than for character stimuli. Contrast, size, and luminance affect the signal-to-noise ratio and temporal characteristics of the signal processing at low levels of the visual system, in particular at the retina. Thus, at high contrast, luminance and large size, the low-level influence is apparently negligible and the dominant limitation is at a higher level in the visual system. At low contrast, luminance and small size, low-level factors contribute to the limitations of processing speed considerably.

The observed contrast and size invariance coincide with electrophysiological and neuro-imaging findings according to which, in the ventral visual pathway, stimulus representation is independent of size and contrast (McCarthy et al., 1999; Avidan et al., 2002). Furthermore, the lateral occipital complex (LOC) response and the psychophysical recognition performance depend in a similar manner on the duration of object presentation (Grill-Spector et al., 2000; Bar et al., 2001): Performance and responses first increase with object duration and level off after 100 ms.

GENERAL DISCUSSION

The experiments comprising this thesis scrutinized distinct stages in the cortical activation sequence underlying successful visual recognition. Study I demonstrated that the first stage of cortical processing, occurring around 80 ms after stimulus presentation around the calcarine sulcus, is sensitive to local features only, whereas a later stage, starting at 130 ms in the occipital and posterior parietal cortex, reflects processing of more global patterns. In Studies II–III, the subjects were presented with recognition tasks at varying levels of difficulty. The occipito-temporal responses from 150 ms onwards were tightly linked with the subjects' performance, in contrast to the earlier mid-occipital responses that were not. Study III showed that the strength of single cortical responses increases gradually when performance improves, supporting graded rather than on–off-like processing in the temporo-occipital cortex. Study IV addressed the attention dependence of the different processing stages: occipito-temporal responses peaking around 150 ms depended strictly on the physical stimulus, whereas later and more sustained activity was strongly modulated by the observer's attention. Hemodynamic responses followed the pattern of the more sustained electrophysiological responses. Study V assessed behaviorally the temporal processing capacity of the human object-recognition system. Above sufficient luminance, contrast and size, the processing speed was not limited by these low-level factors.

Early stages of processing

In Study I, the first cortical responses exceeded noise level at 70 ms and reached their maximum at 85 ms after stimulus onset. This response most likely corresponds to the first component of a typical VEP, labeled N70/N75 or C1 (Jeffreys et al., 1972). Similar early responses were observed, but not systematically analyzed, in Studies I–IV, since they did not show prominent and consistent effects of the experimental manipulations. In all MEG studies, a strong mid-occipital response (M100) peaked around 100 ms. The M100 response was sensitive to spatial frequency and contrast (Study II), but not global configuration (Study I) or object category (Study IV).

M100 has been shown to be stronger during successful rather than unsuccessful face detection (Liu et al., 2002). On the other hand, Study II clearly shows that faces as such are not a preferred stimulus for the processes underlying M100: This response was considerably weaker to plain faces than to plain noise at certain spatial frequencies. Since face recognition is most sensitive to noise at the frequencies that elicited the strongest M100 responses, M100 might plausibly reflect low-level visual analysis sensitive to spatial frequencies that carry critical information about faces.

Contour and object processing

In Study I, contour-sensitive responses emerged at 130–150 ms and peaked around 270 ms, in accordance with previous studies on visual segmentation and grouping (Bach et al., 1997; Han et al., 2001; Fahle et al., 2003; Ohla et al., 2005; Pei et al., 2005; Mathes et al., 2006). However, Studies II–IV and numerous previous reports demonstrate that cortical responses selective to distinct visual categories can peak as early as at 150 ms (Jeffreys, 1989; Allison, McCarthy et al., 1994). Furthermore, both RT and and VEP recordings imply that 150 ms is enough for the visual processing required for high-level categorization (Thorpe et al., 1996; Fabre-Thorpe et al., 2001). Given these latencies, the contour-sensitive responses of Study I, starting around 130 ms and peaking at 270 ms, are apparently too late to reflect the first steps of integrating local information into global shapes. Rather, these parieto-occipital responses might be related to switching of attention to global aspects of the stimulus (Rizzo, 1993; Fink et al., 1997; Vanni et al., 1997).

Cortical responses to faces

The tight correlation observed between the parametrically manipulated recognition performance and the two cortical responses (M170 and M300) suggests a close link between these responses and the processes that subserve face recognition. However, a correlative approach alone cannot unambiguously determine the chain of events that eventually leads to recognition of a face. Instead, any response from the earliest activity in the primary visual cortex to responses related to the ultimate identification of an individual face could, in principle, correlate with e.g. the visibility of the face and therefore with the recognition performance as well. Fortunately, the interpretation can be constrained by several other findings that are discussed below.

Responses up to 200 ms

Temporo-occipital responses in the time-window of 140–200 ms, commonly labeled as N170 or M170 despite variable latencies across studies, are clearly linked with face perception: The response is at least twice as strong to faces than to any control stimuli (Jeffreys, 1989; Lu et al., 1991; Allison, McCarthy et al., 1994; Bentin et al., 1996; George et al., 1996; Sams et al., 1997; Allison et al., 1999; Halgren et al., 2000), its strength parallels the visibility of faces (Tarkiainen et al., 2002; Jemel et al., 2003; Horovitz et al., 2004, Studies II and III) and it is stronger to recognized than unrecognized faces (Liu et al., 2002). Because N170/M170 is not affected by the familiarity of a face or by the semantic information about a face, it is typically interpreted to reflect visual face analysis preceding identification of individual persons (Puce et al., 1999; Eimer, 2000b; Henson et al., 2003; Paller et al., 2003; Herzmann et al., 2004). This interpretation fits well with two present observations: First, the M170 amplitude increased as a function of increasing contrast and stimulus duration, i.e. increasing amount of visual information, beyond contrast levels and stimulus durations at which recognition performance reached ceiling (Studies II and III). Second, M170 was smaller to faces presented with than without a prestimulus mask, despite perfect recognition in both conditions (Study III). Besides face-specific visual analysis, the N170/M170 responses have been suggested to reflect within-category analysis of objects (Rossion, Curran et al., 2002; Rossion, Gauthier et al., 2002), analogous to the expertise account of face-sensitive responses in the fusiform gyrus observed with fMRI (Gauthier et al., 1999; Gauthier et al., 2000).

The temporo-occipital source of the M170 response, as identified in Study II, had mean Talairach coordinates of ±39, –68, –9, corresponding to the center of the face-selective LO activity reported in fMRI (43, –65, –4; Puce et al., 1996). This location also agrees with the site of LOC reported in Grill-Spector et al. (1999) and is within 1–2 cm from the area in the lateral fusiform cortex that typically shows face-specific activation in fMRI (Puce et al., 1996; Kanwisher et al., 1997; McCarthy et al., 1997; Haxby et al., 1999). Although the face-selective fMRI activation and the M170 response may not reflect identical neuronal processes (Study IV), the good agreement between fMRI activation sites and our M170 source area suggests major contribution to the M170 response from the lateral temporo-occipital and/or fusiform areas. On the other hand, despite the precise coordinates reported for the face-specific fMRI and MEG responses, it is plausible that the M170 reflects rather distributed activity: In the intracranial EEG recordings, the sites of face-sensitive 200-ms responses in the ventral temporo-occipital cortex extend for 7 cm in the anterioposterior and 4 cm in the mediolateral directions, and simultaneous face-sensitive sites can be observed on the lateral surface of the temporal lobe (Allison et al., 1999).

Responses beyond 200 ms

The cortical responses to faces beyond 200 ms are less well characterized than the N170/M170 response. Study III demonstrated a strong correlation between the amplitude of these later responses and face recognition, in line with the intracranially recorded face-sensitive responses that peak around 290 ms and 350 ms (Allison et al., 1999). Such responses were detected in the inferotemporal and lateral temporo-occipital cortices, at sites overlapping with or anterior to the generators of the preceding face-related activation. The face responses at the later latencies (> 200 ms) have been further characterized to be stronger to familiar than unfamiliar faces (Puce et al., 1999; Bentin et al., 2000; Eimer, 2000b; Henson et al., 2003; Paller et al., 2003; Herzmann et al., 2004), suggesting that they reflect processes that are related to identification of individual faces and recall of biographical information. They also show sensitivity to perceived facial expressions of emotion (Carretie et al., 1995; Krolak-Salmon et al., 2001; Ashley et al., 2004). The category-sensitive responses in this time window are strongly modulated by attention (Study IV; Lueschow et al., 2004).

Correlates of recognition

The clear correlation between face recognition and the M170 amplitude (Studies II and III) agrees with an earlier report on stronger M170 during trials associated with successful vs. unsuccessful face recognition (Liu et al., 2002). Similar to the behavior of the M170 and M300 responses in Study III, fMRI activity in the LOC and in the fusiform gyrus anterior to the face-specific region correlates with the success of object recognition in a temporal masking paradigm (Grill-Spector et al., 2000; Bar et al., 2001), and activity in the temporo-occipital regions is, in general, coupled to visual recognition performance (James et al., 2000; Lerner et al., 2002; Grill-Spector, Knouf et al., 2004). Furthermore, in a backward masking study where the subjects had to indicate whether a briefly flashed image contained an animal, both the behavioral accuracy and the evoked potential amplitude increased with an increasing delay between the onset of the test image and mask (Bacon-Mace et al., 2005).

The size and contrast invariance of face and letter recognition, demonstrated in Study V, parallels several electrophysiological and neuroimaging findings. Intracranial N200 responses at ventral face-specific sites are to a large extent independent of image size, suggesting that at these sites the representation of faces is scale invariant (McCarthy et al., 1999). Similarly, the correlation between stimulus contrast and fMRI response gradually declines from V1 to higher-order areas, such as area LO and the posterior fusiform gyrus, where the responses are contrast invariant (Avidan et al., 2002). In line with these findings, the correlation of the fMRI response and recognition performance increases from V1 to the contrast invariant higher-order areas (Avidan et al., 2002).

Effects of attention

In Study IV, the occipito-temporal M170 responses peaking around 150 ms depended mainly on the physical stimulus, in contrast to the later and more sustained activity that was strongly modulated by the observer's attention. The stimulus-dependent responses (< 200 ms) likely reflect a feed-forward phase of perceptual processing that precedes processes modulated by a more distributed neural system (Thorpe et al., 1996; Haxby et al., 2000). The later responses show stronger modulation by experimental manipulations suggesting top-down control, i.e., attention, priming, and familiarity (Puce et al., 1999; Eimer, 2000b; Henson et al., 2003; Paller et al., 2003; Lueschow et al., 2004). An early response that reflects a feed-forward phase, driven by sensory input and the late response that reflects modulation by inter-arel interactions may involve activity in the same neurons. According to single unit recordings from inferior temporal cortex (IT; Sugase et al., 1999), middle temporal cortex (MT; Pack et al., 2001), and earlier visual cortices (Lamme, 1995), the activity in individual cells carries different information during early and late responses. Although the M170 response in Study IV was unmodulated by attention, it can nevertheless be subject to top-down modulations: Recent studies have reported frontal lobe responses that preceded recognition-related responses in the temporo-occipital cortex by 50 ms (Bar et al., 2006; Barbeau et al., 2008), likely reflecting potential for contextual facilitation of visual object recognition.

Under certain circumstances, preliminary processing of an unattended object may play a role in selecting an attended object. In the attention-to-houses condition in Study IV, the undiminished face-sensitive M170 response indicated unaltered early processing of the unattended face. Similarly, in the attention-to-faces condition, features of the early house-sensitive response were seen, suggesting that the unattended house was also processed before the onset of attentional modulation. Other studies indicate that attention can have a small but significant effect on the face-sensitive N170/M170 when task requirements allow the attended non-face stimulus to be selected without processing the unattended face (Eimer, 2000a; Downing, Liu et al., 2001; Holmes et al., 2003). These earlier findings suggest that preliminary, feed-forward processing can be biased to process attended information preferentially if the unattended information is not required for stimulus selection. These attentional effects on the early responses were, however, much weaker (small reductions in amplitude or delays in latency) than the effect of attention on later responses in Study IV, where the response to the unattended face was essentially eliminated.

Methodological considerations

Spatial resolution

Human brain imaging studies often average activity over areas spanning a square centimetre or more. However, because the cortex show functional specialization at the scale of cortical columns (Mountcastle, 1997), the common imaging methods are likely to average activity across units with different feature specificity. The relevance of this concern has been demonstrated by comparing fMRI and intracranial subdural EEG recordings under identical stimulus conditions: At several cortical sites the EEG activity was confined to a single visual category, showing much higher selectivity than the fMRI responses (Privman et al., 2007).

The discrepancy between the MEG results in Study I and single-unit recordings further illustrates the limits of imaging methods that measure neural processes at the population level: Study I did not demonstrate a transition from local to global processing in the primary visual cortex, although single unit recordings in monkeys provide strong evidence for processing of such information at the level of V1 (Li et al., 2006). Since stimulus elements with similar orientations typically facilitate neural responses and suppress responses to different orientations (Kapadia et al., 1995), it is possible that at the population level these effects cancel each other.

Promisingly, the resolution of functional magnetic resonance imaging (fMRI) has recently reached the scale of cortical columns in the primary visual cortex (Cheng et al., 2001; Sun et al., 2007). Since specialization at the columnar scale is a feature of the higher-order visual areas as well (Fujita et al., 1992; Tanaka, 1993), successful application of columnar-resolution imaging to the human occipito-temporal cortex would permit the studying the organizing principles of the visual object processing areas in humans at a more relevant scale than what has been possible until now. Column-resolution imaging would also facilitate the integration of information across single-unit recordings in monkeys and brain imaging studies in humans. Besides increasing spatial resolution, innovative experimental designs such as adaptation paradigms (Grill-Spector & Malach, 2001) can provide alternative means for bypassing some limitations set by the imaging methods.

Psychophysical paradigms and brain imaging

In Study III, temporal masking was achieved with pixel noise, whereas the behavioral Study V employed an RSVP paradigm where faces were masked with other faces. These two masking paradigms might affect the visual system at different levels of processing, hindering the integration of findings from these two studies. However, different approaches were chosen since each stimulus in an RSPV paradigm elicits a complex sequence of cortical responses, and when these responses overlap at high stimulation rates, disentangling the contributions of different areas with MEG is not straightforward. Promisingly, new analysis techniques might facilitate the implementation of more complex stimulus paradigms in MEG experiments. For example, neural processes underlying reading have been studied with an RSVP stimulus paradigm by analyzing the coupling of oscillatory activity between different brain areas (Kujala et al., 2007).

MEG vs. fMRI

In principle, MEG/EEG and fMRI complement each other: MEG and EEG signals are generated directly by neural activity and can thus resolve temporal events on a millisecond time scale. On the other hand, the spatial resolution for multiple, distributed sources is limited (Hämäläinen et al., 2002). FMRI, in turn, has relatively high spatial resolution but coarse temporal resolution because it measures slower hemodynamic changes elicited by neural activity (Buxton et al., 1998). FMRI BOLD signal reflects metabolic demand, whereas MEG/EEG measures reflect electrophysiological activity and are influenced strongly by the degree of synchrony of neural activity (Hari, Salmelin et al., 1997). Consequently, these measures may be dissociated when the activity is synchronous but brief, resulting in a small metabolic demand, or when it reflects a resetting of the phase of the spontaneous activity with no change in power.

The similar cortical locations of the face-sensitive fMRI response in the fusiform gyrus and the face-sensitive electrophysiological M170/N170 response suggest that they reflect, at least to some degree, the same neural activity. The results from Study IV indicate, however, that these responses can be dissociated by the effect of attention:

MEG detected a strong, but brief, early face-selective response that was not evident in the hemodynamic signal that fMRI measured.

The fraction of the cortical response to faces generating the M170 appears too small to be detected by ordinary fMRI. Alternatively, the hemodynamic signal associated with this transient response might be masked by the signal generated by longer-lasting activity. The possibility thus remains that a brief hemodynamic response, reflecting the same neural processes as the M170, occurs in the fusiform gyrus. Although such evidence was not detected in Study IV, an event-related fMRI design that allows a more detailed analysis of the earliest phases of the hemodynamic response might provide more information on this issue. Combined MEG–fMRI studies that systematically vary the strength of the early and late face-selective responses may also detect a hemodynamic correlate of the early response.

Natural vs. artificial stimulation

The studies constituting this thesis, as well as almost all current knowledge on the human visual system, are based on experimental settings that differ from natural viewing conditions in a number of ways. In a typical laboratory setup, the subjects hold their gaze at a constant location at which still images of single objects are presented one by one. Under normal viewing conditions, on the contrary, we freely and actively move our gaze across complex scenes that are under continuous change.

The traditional experimental approach, manipulating one variable at a time to isolate its effects on other variables, has been and will remain crucial for unraveling how the visual system functions. But in parallel, it is necessary to explore whether the findings obtained under simplistic laboratory conditions hold in more natural settings as well. Unexpected findings under natural conditions can then, in turn, elicit new questions that have to be tackled under strictly controlled conditions. Recently, several lines of research have evolved to study visual processing under (semi)-natural conditions (Kayser et al., 2004). In a pioneering study, Gallant et al. (1998) showed that neural activity in the visual cortex of monkeys was higher during viewing of flashed gratings and flashed patches of natural scenes than during free viewing of natural scenes. Subsequent refinements of the paradigm demonstrated that natural stimulation increases the selectivity and sparseness of neural responses in primary visual cortex, as well as information transmission efficiency (Vinje et al., 2000, 2002). Furthermore, natural stimuli seem to strengthen the nonlinearities of neural responses (Smyth et al., 2003; David et al., 2004) and decrease the size of the receptive fields of category-selective neurons in the inferior temporal cortex of experimental animals (Rolls et al., 2003).

In humans, natural stimulation (a 2D movie) and traditional paradigms elicit partly similar patterns of hemodynamic activity, but natural stimuli activate more areas with a higher specificity (Bartels et al., 2004). The sequence of such activation patterns shows surprisingly high similarity across subjects, demonstrating the feasibility of the approach (Hasson et al., 2004). Proceeding towards natural stimulation in a different way, Ponkanen et al., (2008) showed differences between the response elicited by a real human being vs. an image of the same person. Similarly, Järveläinen et al. (2001) showed that the actions of a real person elicit stronger modulation of the cortical action observation system than a video clip presenting the same action. Taken together, the studies in both monkeys and humans point out similarities as well as clear differences in the cortical processing of artificial vs. natural stimuli. This conclusion agrees with a large body of behavioral studies showing that natural context facilitates processing of visual information (Bar, 2004).

CONCLUSIONS

Taken together, the studies comprising this thesis demonstrate several distinct stages in the cortical activation sequence underlying visual object recognition, reflecting the level of feature integration, difficulty of recognition, and direction of attention. In the context of other literature, the observed functional characteristics of processing under different time windows can be interpreted as described below.

Mid-occipital responses up to 100 ms (M100) are consistently elicited by a wide range of visual inputs. They are highly sensitive to such low-level attributes as contrast and spatial frequency (Study II), but remain unaffected by the global configuration of the stimulus (Study I). These early responses have previously been suggested to be sensitive to more abstract features as well, for example to object category. The experiments included in this thesis were not designed to provide direct evidence either for or against such claims, but no positive correlation between face recognition and the M100 responses was observed. In line with this, Study II indicated that the presence of a face in the stimulus has little effect on the M100 responses. Furthermore, Study II demonstrated the sensitivity of the early responses to spatial frequency and contrast, thereby emphasizing the importance of careful control for low-level attributes when claims in support of higher-order selectivity in early visual responses are made.

Temporo-occipital responses peaking at 140–180 ms (M170) have consistently been shown to be stronger for visually presented faces than for other categories. Extending this finding, Studies II and III demonstrate these responses to be strongly linked with face recognition performance. However, the processes underlying M170 still bear some relation to low-level visual properties: They are affected by for example the contrast of a face (Study II). Study IV showed that the M170 responses are automatic, triggered independently of the observer’s attention.

From 200 ms onwards, the cortical activity elicited by different objects seems less focal. Clearly different patterns of responses are evoked by faces and houses (Study IV), and like the M170 response, the subsequent processing shows high correlation with recognition performance (Study III). The major difference is the strong task dependence of the later responses (Study IV). Other studies indicate that the processing beyond

200 ms is affected by the familiarity of the faces, indicating links to memory related processes.

The studies presented here demonstrate correlation with recognition, but do not distinguish processes related to the detection, categorization and identification of a face. Neither did the experiments probe whether the responses were modulated by mere visual features of the face, or the identity of the person. Therefore, the characterization of the exact nature of the processing within the different time windows will require extensive further research. In parallel with such careful experimental dissection of the processes underlying visual recognition, another important direction is to probe and extend these findings under viewing conditions as natural as possible. This will be important for validation of results obtained under rather artificial conditions, as well as for exploring the effects of temporal and spatial context and active viewing on visual cortical processing.

ACKNOWLEDGMENTS

This research work was carried out in the Brain Research Unit of the Low Temperature Laboratory at the Helsinki University of Technology. I deeply respect late Academician Olli V. Lounasmaa for his talent and efforts that made the establishment and growth of this true centre of excellence possible. I am indebted to him and the current leader of the laboratory, Professor Mikko Paalanen, for the superb research environment.

Financially, this thesis work was supported by the Finnish Graduate School of Neuroscience, the Academy of Finland (National Centers of Excellence Programs 2000-2005 and 2006-2011), the Finnish Cultural Foundation, the Emil Aaltonen Foundation, the Ella and Georg Ehrnrooth Foundation, and the Sigrid Jusélius Foundation.

Professor Riitta Hari's impact on this thesis and the development of my research skills has been fundamental. As the head of the Brain Research Unit, she has provided unique facilities for the work, and as my supervisor, an infinite resource for learning to understand and conduct neuroscience. In particular, her around-the-clock devotion to improving the presentation of the studies, from tick marks to clarity of thought, has been remarkable. Moreover, I do acknowledge her efforts in stretching her patience and tolerance through these years…

I thank Professors Jari Hietanen and Kristian Donner for their contribution as the members of my FGSN follow-up committee, and Professor Kimmo Alho for the smooth handling of the final stages of the project. I am indebted to Professors Jari Hietanen and Moshe Bar for their suggestions that greatly improved the summary part of this thesis, despite the very tight pre-review schedule to which they kindly adhered. The quick and thoughtful comments by Docent Simo Vanni, Dr. Tiina Parviainen, Dr. Hanna Renvall and MSc Lauri Parkkonen, always accompanied with optimistic encouragement, were crucial for the completion of the summary. Special thanks are due to Dr. Cathy Nangini for revising my English, to MA Marikki Heinomo for saving me from the hatred of Terho Itkonen, and to series editor, Professor Veijo Virsu for guidance in formatting the thesis.

Conducting the studies that comprise this thesis has given me the privilege to collaborate with a number of skilled scientists. I owe my gratitude to Docent Risto

Näsänen for his expertise in visual perception and psychophysics, and to Professor James Haxby, Dr. Michael Beauchamp and Dr. Maura Furey for the educative and enjoyable visits that attracted me to the world of face and object perception. Lauri Parkkonen's in-depth understanding of everything from dynamite to Josephson junctions, together with the extraordinary skills in communicating his knowledge, has been a real blessing. I am thankful to Dr. Kimmo Uutela for his expertise in computation and data analysis, and to Docent Jussi Saarinen for the guidance and support since the times of my master's thesis. I thank Sari Avikainen, Teresa Montez, Helena Ojanpää and Juha Päällysaho for their contributions to conducting, analyzing and reporting the experiments.

I owe my sincere gratitude to the former and current BRU seniors, especially Acad. Prof. Riitta Salmelin, Professor Matti Hämäläinen, Docents Nina Forss and Päivi Helenius, and Drs Ole Jensen, Mika Seppä and Matti Kajola for sharing their time and knowledge whenever I have had something to ask about. In particular, I want to express my gratitude to Docent Simo Vanni for the devoted introduction to the world of MEG during my first year in the laboratory, and for the subsequent discussions and advice. The vision and working memory seminars at the Department of Psychology, led by Docent Pentti Laurinen and Dr. Virpi Kalakoski, and the "vision and hearing evening meetings" have been of great value in broadening my understanding of the field.

I am truly grateful to the brains and hands behind the development and maintenance of the top-class environment for running and analyzing MEG experiments, as well as to the members of the computer support team that have faithfully kept the systems running and provided their immediate, proficient and kind help whenever problems have occurred. I am thankful to the LTL secretaries for handling the financial, travel and purchase bureaucracies, and a variety of other practicalities.

During my years in the BRU, quite a number of researchers have been coming, staying, and leaving. Therefore, preparing a comprehensive and error-free list of the names unfortunately exceeds my capacities. Nevertheless, although the flux of people has given different eras somewhat different flavors, the warm and fun atmosphere among the students has prevailed. It is this very atmosphere that has made coming to the lab so pleasant from day to day, including the occasional times when the research itself might not have been progressing at the speed of light. My warmest thanks to everyone responsible!

Some 95% of this thesis work was conducted in room F347, but what made the years at this southeast end of the Axis so irreplaceable was learning to know Hanna and next door’s Tiina. Infinite and multidimensional thanks to the two of you, just as something to start with. Major thanks are due to Jan and Lauri as well for pleasantly sharing the entire time course of scientific and social activity, and to Miiu for the happy and supporting company in the old and new office.

Whether or not growing up in the Athens of Finland has affected my subsequent academic activities, it has profoundly influenced my social life in the form of Taru, Matti, Jussi, Jenni, Kati, Simo, Marikki, Sami, Toni, Riikka, Anssi, Mari, Timo, Johanna, Solja, Piia, Teppo, Janne, Sanna, Jani, Joni and Leila. Thank you for all the happy, interesting and relaxing times!

To my parents Kaisu and Antti, I am ever grateful for their love and support through the years. With respect to the years spent with this thesis, I am particularly thankful for the project Lojo: The pumpkins, gadgets, eatings, floatings and a lot more have brought me endless excitement and relaxation. To my sister Katri, valtavat thanks are due for the shared time and discussions, and of course for introducing Mr. Vedran and wonderful Croatia to the family. Warm thanks also to Riitta and Martti, and to the unique Matsumoto family.

Finally, how can I ever thank dear Tomoko enough for leaving the tasty foods of Japan to live with a person who is angry because he should finish his thesis? Endless hugs for your too kind caring and all the happiness of our daily life.

Töölönkadulla, aurinkoisena vappuviikonloppuna 2008

Topi Tanskanen

REFERENCES

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron, 21*, 373–383.

Ahonen, A. I., Hämäläinen, M. S., Kajola, M. J., Knuutila, J. E. T., Laine, P. P., Lounasmaa, O. V., Parkkonen, L. T., Simola, J. T., & Tesche, C. D. (1992). A 122-channel SQUID instrument for investigating the magnetic signals from the human brain. *Physica Scripta, T49*, 198–205.

Aine, C. J., Bodis-Wollner, I., & George, J. S. (1990). Generators of visually evoked neuromagnetic responses. Spatial-frequency segregation and evidence for multiple sources. *Advances in Neurology, 54*, 141–155.

Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., & Spencer, D. D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology, 71*, 821–825.

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 4*, 544–554.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*, 415–430.

Altmann, C. F., Bulthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology, 13*, 342–349.

Anderson, S. J., Holliday, I. E., Singh, K. D., & Harding, G. F. (1996). Localization and functional analysis of human cortical area V5 using magneto-encephalography. *Proceedings. Biological Sciences / The Royal Society, 263*, 423–431.

Andrews, T. J., Schluppeck, D., Homfray, D., Matthews, P., & Blakemore, C. (2002). Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *Neuroimage, 17*, 890–901.

Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J. M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *Journal of Neuroscience, 22*, 8633–8646.

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport, 15*, 211–216.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., & Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *Journal of Neurophysiology, 87*, 3102–3116.

Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. *Journal of Cognitive Neuroscience, 17*, 1150–1167.

Bach, M., & Meigen, T. (1997). Similar electrophysiological correlates of texture segregation induced by luminance, orientation, motion and stereo. *Vision Research, 37*, 1409–1414.

Bacon-Mace, N., Mace, M. J., Fabre-Thorpe, M., & Thorpe, S. J. (2005). The time course of visual processing: backward masking and natural scene categorisation. *Vision Research, 45*, 1459–1469.

Baillet, S., Mosher, J., & Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 14–30.

Baker, C. I., Hutchison, T. L., & Kanwisher, N. (2007). Does the fusiform face area contain subregions highly selective for nonfaces? *Nature Neuroscience, 10*, 3–4.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*, 600–609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617–629.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hamalainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 449–454.

Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., Rosen, B. R., & Dale, A. M. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron, 29*, 529–535.

Barbeau, E. J., Taylor, M. J., Regis, J., Marquis, P., Chauvel, P., & Liegeois-Chauvel, C. (2008). Spatio temporal dynamics of face recognition. *Cerebral Cortex, 18*, 997–1009.

Bartels, A., & Zeki, S. (2004). The chronoarchitecture of the human brain—natural viewing conditions reveal a time-based anatomy of the brain. *Neuroimage, 22*, 419–433.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron, 34*, 149–159.

Bentin, S., Allison, T., Puce, A., Perez, A., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*, 551–565.

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology, 17*, 35–54.

Bonneh, Y., & Sagi, D. (1998). Effects of spatial configuration on contrast detection. *Vision Research, 38*, 3541–3553.

Braeutigam, S., Bailey, A. J., & Swithenby, S. J. (2001). Task-dependent early latency (30–60 ms) visual processing of human faces and other objects. *Neuroreport, 12*, 1531–1536.

Brenner, D., Williamson, S. J., & Kaufman, L. (1975). Visually evoked magnetic fields of the human brain. *Science, 190*, 480–482.

Broadbent, D. E. (1958). *Perception and Communication*. Oxford: Pergamon.

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*, 369–384.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77 (Pt 3)*, 305–327.

Bullier, J., & Nowak, L. G. (1995). Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Current Opinion in Neurobiology, 5*, 497–503.

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine, 39*, 855–864.

Carretie, L., & Iglesias, J. (1995). An ERP study on the specificity of facial expression processing. *International Journal of Psychophysiology, 19*, 183–192.

Catani, M., Howard, R. J., Pajevic, S., & Jones, D. K. (2002). Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage, 17*, 77–94.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology, 88*, 2530–2546.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology, 88*, 2547–2556.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*, 564–583.

Cheng, K., Waggoner, R. A., & Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron, 32*, 359–374.

Clark, V. P., Keil, K., Maisog, J. M., Courtney, S., Ungerleider, L. G., & Haxby, J. V. (1996). Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. *Neuroimage, 4*, 1–15.

Cohen, D. (1968). Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science, 161*, 784–786.

Cohen, D. (1972). Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer. *Science, 175*, 664–666.

Colby, C. L., Gattass, R., Olson, C. R., & Gross, C. G. (1988). Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative Neurology, 269*, 392–413.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201–215.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics, 58*, 602–612.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*, 179–194.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26*, 55–67.

Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology, 32*, 331–341.

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *Journal of Neuroscience, 24*, 6991–7006.

De Monasterio, F. M., & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *The Journal of Physiology, 251*, 167–195.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. New York: Oxford University Press.

Del Gratta, C., Pizzella, V., Torquati, K., & Romani, G. L. (1999). New trends in magnetoencephalography. *Electroencephalography and Clinical Neurophysiology. Supplement, 50*, 59–73.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Downing, P., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia, 39,* 1329–1342.

Downing, P. E., Chan, A. W., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex, 16,* 1453–1461.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293,* 2470–2473.

Dumoulin, S. O., & Hess, R. F. (2007). Cortical specialization for concentric shape processing. *Vision Research, 47,* 1608–1613.

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Research. Cognitive Brain Research, 10,* 145–158.

Eimer, M. (2000b). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology, 111,* 694–705.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392,* 598–601.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13,* 171–180.

Fahle, M., Quenzer, T., Braun, C., & Spang, K. (2003). Feature-specific electrophysiological correlates of texture segregation. *Vision Research, 43,* 7–19.

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex, 17,* 2400–2406.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1–47.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research, 33,* 173–193.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1997). Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120 (Pt 10),* 1779–1791.

Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception, 12,* 195–201.

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping, 8,* 272–284.

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *Neuroimage, 2,* 45–53.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature, 360,* 343–346.

Fylan, F., Holliday, I. E., Singh, K. D., Anderson, S. J., & Harding, G. F. (1997). Magnetoencephalographic investigation of human cortical area V1 using color stimuli. *Neuroimage, 6,* 47–57.

Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science, 259,* 100–103.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology, 76,* 2718–2739.

Gallant, J. L., Connor, C. E., & Van Essen, D. C. (1998). Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. *Neuroreport, 9,* 85–90.

Galletti, C., Battaglini, P. P., & Fattori, P. (1991). Functional properties of neurons in the anterior bank of the parieto-occipital sulcus of the macaque monkey. *European Journal of Neuroscience, 3,* 452–461.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103,* 322–330.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3,* 191–197.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2,* 568–573.

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Brain Research. Cognitive Brain Research, 4,* 65–76.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20–25.

Grill-Spector, K. (2003). The functional organization of the ventral visual pathway and its relationship to object recognition. In N. Kanwisher & J. Duncan (Eds.), *Functional Brain Imaging of Visual Cognition* (Vol. XX, pp. 169–193). London: Oxford University Press.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7,* 555–562.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41,* 1409–1422.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24,* 187–203.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron, 21,* 191–202.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3,* 837–843.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica, 107,* 293–321.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*, 649–677.

Grill-Spector, K., Sayres, R., & Ress, D. (2006). High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. *Nature Neuroscience, 9*, 1177–1185.

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex, 10*, 69–81.

Han, S., Song, Y., Ding, Y., Yund, E. W., & Woods, D. L. (2001). Neural substrates for visual perceptual grouping in humans. *Psychophysiology, 38*, 926–935.

Hari, R. (1990). The neuromagnetic method in the study of the human auditory cortex. In F. Grandori, M. Hoke & G. L. Romani (Eds.), *Auditory Evoked Magnetic Fields and Electric Potentials* (Vol. 6, pp. 222–282). Basel: Karger.

Hari, R. (1998). Magnetoencephalography as a tool of clinical neurophysiology. In E. Niedermeyer & F. Lopes da Silva (Eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Baltimore: Williams & Wilkins.

Hari, R., & Ilmoniemi, R. J. (1986). Cerebral magnetic fields. *Critical Reviews in Biomedical Engineering, 14*, 93–126.

Hari, R., Levanen, S., & Raij, T. (2000). Timing of human cortical functions during cognition: role of MEG. *Trends in Cognitive Sciences, 4*, 455–462.

Hari, R., & Salmelin, R. (1997). Human cortical oscillations: a neuromagnetic view through the skull. *Trends in Neurosciences, 20*, 44–49.

Hari, R., Salmelin, R., Makela, J. P., Salenius, S., & Helle, M. (1997). Magnetoencephalographic cortical rhythms. *International Journal of Psychophysiology, 26*, 51–62.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32*, 203–218.

Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience, 15*, 419–431.

Hasson, U., Hendler, T., Ben Bashat, D., & Malach, R. (2001). Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *Journal of Cognitive Neuroscience, 13*, 744–753.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron, 34*, 479–490.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science, 303*, 1634–1640.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*, 2425–2430.

Haxby, J. V., Gobbini, M. I., & Montgomery, K. (2004). Spatial and Temporal Distribution of Face and Object Representations in the Human Brain. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences III* (pp. 889–904): MIT Press.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B., & Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America, 88*, 1621–1625.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223–233.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron, 22*, 189–199.

Hayes, T., Morrone, M. C., & Burr, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception, 15*, 595–602.

Hecaen, H., & Angelergues, R. (1962). Agnosia for faces (prosopagnosia). *Archives of Neurology, 7*, 92–100.

Hegde, J., & Van Essen, D. C. (2006). Temporal dynamics of 2D and 3D shape representation in macaque visual area V4. *Visual Neuroscience, 23*, 749–763.

Hegde, J., & Van Essen, D. C. (2007). A comparative study of shape representation in macaque visual areas v2 and v4. *Cerebral Cortex, 17*, 1100–1116.

Hendry, S. H., & Reid, R. C. (2000). The koniocellular pathway in primate vision. *Annual Review of Neuroscience, 23*, 127–153.

Hendry, S. H., & Yoshioka, T. (1994). A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science, 264*, 575–577.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*, 793–805.

Henson, R. N., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. *Neuroimage, 15*, 83–97.

Herzmann, G., Schweinberger, S. R., Sommer, W., & Jentzsch, I. (2004). What's special about personally familiar faces? A multimodal approach. *Psychophysiology, 41*, 688–701.

Hess, R. F., Hayes, A., & Field, D. J. (2003). Contour integration and cortical processing. *Journal of Physiology, Paris, 97*, 105–119.

Hillebrand, A., & Barnes, G. R. (2002). A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage, 16*, 638–650.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3*, 80–84.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research, 16*, 174–184.

Horovitz, S. G., Rossion, B., Skudlarski, P., & Gore, J. C. (2004). Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage, 22*, 1587–1595.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*, 574–591.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, 195*, 215–243.

Hulme, O. J., & Zeki, S. (2007). The Sightless View: Neural Correlates of Occluded Objects. *Cerebral Cortex, 17*, 1197–1205.

Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394*, 784–787.

Hämäläinen, M., & Hari, R. (2002). Magnetoencephalographic Characterization of Dynamic Brain Activation: Basic Principles and Methods of Data Collection and Source Analysis. In A. W. Toga & J. C. Mazziotta (Eds.), *Brain Mapping. The Methods* (2nd ed., pp. 227–253). Amsterdam: Academic Press.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics, 65*, 413–497.

Ishai, A. (2008). Let's face it: It's a cortical network. *Neuroimage, 40*, 415–419.

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage, 17*, 1729–1741.

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron, 28*, 979–990.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 9379–9384.

James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology, 10*, 1017–1024.

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research, 78*, 193–202.

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Experimental Brain Research, 16*, 1–21.

Jemel, B., Schuller, A. M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., & Bruyer, R. (2003). Stepwise emergence of the face-sensitive N170 event-related potential component. *Neuroreport, 14*, 2035–2039.

Jousmäki, V., Hämäläinen, M., & Hari, R. (1996). Magnetic source imaging during a visually guided task. *Neuroreport, 7*, 2961–2964.

Järveläinen, J., Schürmann, M., Avikainen, S., & Hari, R. (2001). Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts. *Neuroreport, 12*, 3493–3495.

Kaas, J. H., & Lyon, D. C. (2007). Pulvinar contributions to the dorsal and ventral streams of visual processing in primates. *Brain Research Reviews, 55*, 285–296.

Kakigi, R., Hoshiyama, M., Shimojo, M., Naka, D., Yamasaki, H., Watanabe, S., Xiang, J., Maeda, K., Lam, K., Itomi, K., & Nakamura, A. (2000). The somatosensory evoked magnetic fields. *Progress in Neurobiology, 61*, 495–523.

Kaneoke, Y. (2006). Magnetoencephalography: in search of neural processes for visual motion information. *Progress in Neurobiology, 80*, 219–240.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 361*, 2109–2128.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*, 843–856.

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 12073–12078.

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology, 84*, 2048–2062.

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America, 83*, 2755–2757.

Kayser, C., Kording, K. P., & Konig, P. (2004). Processing of complex stimuli and natural scenes in the visual cortex. *Current Opinion in Neurobiology, 14*, 468–473.

Kelly, D. H. (1961). Visual response to time-dependent stimuli. I. Amplitude sensitivity measurements. *Journal of the Optical Society of America, 51*, 422–429.

Kleinschmidt, A., Buchel, C., Hutton, C., Friston, K. J., & Frackowiak, R. S. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron, 34*, 659–666.

Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67*, 961–980.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20*, 3310–3318.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293*, 1506–1509.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: monkey and human FMRI studies. *Neuron, 37*, 333–346.

Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 5587–5591.

Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 7495–7497.

Krolak-Salmon, P., Fischer, C., Vighetto, A., & Mauguiere, F. (2001). Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. *European Journal of Neuroscience, 13*, 987–994.

Kujala, J., Pammer, K., Cornelissen, P., Roebroeck, A., Formisano, E., & Salmelin, R. (2007). Phase coupling in a cerebro-cerebellar network at 8-13 Hz during reading. *Cerebral Cortex, 17*, 1476–1485.

Kurki, I., & Saarinen, J. (2004). Shape perception in human vision: specialized detectors for concentric spatial structures? *Neuroscience Letters, 360*, 100–102.

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15*, 1605–1615.

Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cerebral Cortex, 12*, 163–177.

Lesser, R. P., & Arroyo, S. (2005). Subdural electrodes. In E. Niedermeyer & F. Lopes da Silva (Eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Baltimore: Williams & Wilkins.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*, 533–539.

Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron, 50*, 951–962.

Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping, 27*, 1–13.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters*, *253*, 147–150.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience, 5*, 910–916.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*, 150–157.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology, 66*, 735–769.

Lu, S. T., Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Lounasmaa, O. V., Sams, M., & Vilkman, V. (1991). Seeing faces activates three separate areas outside the occipital visual cortex in man. *Neuroscience, 43*, 287–290.

Lueck, C. J., Zeki, S., Friston, K. J., Deiber, M. P., Cope, P., Cunningham, V. J., Lammertsma, A. A., Kennard, C., & Frackowiak, R. S. (1989). The colour centre in the cerebral cortex of man. *Nature, 340*, 386–389.

Lueschow, A., Sander, T., Boehm, S. G., Nolte, G., Trahms, L., & Curio, G. (2004). Looking for faces: Attention modulates early occipitotemporal object processing. *Psychophysiology, 41*, 350–360.

Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences, 6*, 176–184.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 8135–8139.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature, 379*, 649–652.

Mathes, B., Trenner, D., & Fahle, M. (2006). The electrophysiological correlate of contour integration is modulated by task demands. *Brain Research, 1114,* 98–112.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9,* 431–444.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9,* 605–610.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11,* 8–15.

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *Journal of Neuroscience, 19,* 8560–8572.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16,* 369–402.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6,* 414–417.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain, 120 (Pt 4),* 701–722.

Musselwhite, M. J., & Jeffreys, D. A. (1985). The influence of spatial frequency on the reaction times and evoked potentials recorded to grating pattern stimuli. *Vision Research, 25,* 1545–1555.

Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., & Malach, R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Current Biology, 17,* 1275–1285.

Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In J. Kaas, K. Rockland & A. Peters (Eds.), *Extrastriate Cortex. Cerebral Cortex, vol 12.* (pp. 205–241). New York: Plenum Press.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39,* 3824–3833.

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401,* 584–587.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013–1023.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America, 87,* 9868–9872.

Ohla, K., Busch, N. A., Dahlem, M. A., & Herrmann, C. S. (2005). Circles are different: the perception of Glass patterns modulates early event-related potentials. *Vision Research, 45,* 2668–2676.

Okada, Y. C., Wu, J., & Kyuhou, S. (1997). Genesis of MEG signals in a mammalian CNS structure. *Electroencephalography and Clinical Neurophysiology, 103,* 474–485.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607–609.

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature, 409*, 1040–1042.

Packer, O., Diller, L. C., Verweij, J., Lee, B. B., Pokorny, J., Williams, D. R., Dacey, D. M., & Brainard, D. H. (2001). Characterization and use of a digital light projector for vision research. *Vision Research, 41*, 427–439.

Paller, K. A., Ranganath, C., Gonsalves, B., LaBar, K. S., Parrish, T. B., Gitelman, D. R., Mesulam, M. M., & Reber, P. J. (2003). Neural correlates of person recognition. *Learning and Memory, 10*, 253–260.

Parkkonen, L. T., Simola, J. T., Tuoriniemi, J., & Ahonen, A. I. (1999). An interference suppression system for multichannel magnetic field detector arrays. In T. Yoshimoto, M. Kotani, S. Kuriki, H. Karibe & N. Nakasato (Eds.), *Recent advances in biomagnetism: Proceedings of the 11th International Conference on Biomagnetism* (pp. 13). Sendai: Tohoku University Press.

Pei, F., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2005). Event-related potentials show configural specificity of global form processing. *Neuroreport, 16*, 1427–1430.

Peli, E., Lee, E., Trempe, C. L., & Buzney, S. (1994). Image enhancement for the visually impaired: the effects of enhancement on face recognition. *Journal of the Optical Society of America A. Optics and Image Science, 11*, 1929–1939.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47*, 329–342.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 223*, 293–317.

Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research, 38*, 865–879.

Pitzalis, S., Galletti, C., Huang, R. S., Patria, F., Committeri, G., Galati, G., Fattori, P., & Sereno, M. I. (2006). Wide-field retinotopy defines human cortical visual area v6. *Journal of Neuroscience, 26*, 7962–7973.

Ponkanen, L. M., Hietanen, J. K., Peltola, M. J., Kauppinen, P. K., Haapalainen, A., & Leppanen, J. M. (2008). Facing a real person: an event-related potential study. *Neuroreport, 19*, 497–501.

Portin, K., Salenius, S., Salmelin, R., & Hari, R. (1998). Activation of the human occipital and parietal cortex by pattern and luminance stimuli: neuromagnetic measurements. *Cerebral Cortex, 8*, 253–260.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58,* 1–23.

Privman, E., Nir, Y., Kramer, U., Kipervasser, S., Andelman, F., Neufeld, M. Y., Mukamel, R., Yeshurun, Y., Fried, I., & Malach, R. (2007). Enhanced category tuning revealed by intracranial electroencephalograms in high-order human visual areas. *Journal of Neuroscience, 27*, 6234–6242.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205–5215.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9*, 445–458.

Rizzo, M. (1993). 'Bálint's syndrome' and associated visuospatial disorders. *Baillieres Clinical Neurology, 2*, 415–437.

Robinson, D. L., & Petersen, S. E. (1992). The pulvinar and visual salience. *Trends in Neurosciences, 15*, 127–132.

Roelfsema, P. R. (2006). Cortical Algorithms for Perceptual Grouping. *Annual Review of Neuroscience, 29*, 203–227.

Roland, P. E., & Zilles, K. (1994). Brain atlases--a new research tool. *Trends in Neurosciences, 17*, 458–467.

Rolls, E. T., Aggelopoulos, N. C., & Zheng, F. (2003). The receptive fields of inferior temporal cortex neurons in natural scenes. *Journal of Neuroscience, 23,* 339–348.

Rolls, E. T., & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings. Biological Sciences / The Royal Society, 257*, 9–15.

Rolls, E. T., Tovee, M. J., Purcell, D. G., Stewart, A. L., & Azzopardi, P. (1994). The responses of neurons in the temporal cortex of primates, and face identification and detection. *Experimental Brain Research, 101,* 473–484.

Rossi, A. F., Desimone, R., & Ungerleider, L. G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience, 21,* 1698–1709.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain, 126,* 2381–2395.

Rossion, B., Curran, T., & Gauthier, I. (2002). A defense of the subordinate-level expertise account for the N170 component. *Cognition, 85,* 189–196.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science: A Journal of the American Psychological Society, 13*, 250–257.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience, 8,* 107–113.

Saarinen, J., & Levi, D. M. (2001). Integration of local features into a global shape. *Vision Research, 41*, 1785–1790.

Saarinen, J., Levi, D. M., & Shen, B. (1997). Integration of local pattern elements into a global shape in human vision. *Proceedings of the National Academy of Sciences of the United States of America, 94,* 8267–8271.

Saint-Amour, D., Saron, C. D., Schroeder, C. E., & Foxe, J. J. (2005). Can whole brain nerve conduction velocity be derived from surface-recorded visual evoked potentials? A re-examination of Reed, Vernon, and Johnson (2004). *Neuropsychologia, 43*, 1838–1844.

Salenius, S., & Hari, R. (2003). Synchronous cortical oscillatory activity during motor action. *Current Opinion in Neurobiology, 13,* 678–684.

Salmelin, R., Hari, R., Lounasmaa, O. V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature, 368,* 463–465.

Salmelin, R., & Kujala, J. (2006). Neural representation of language: activation versus long-range connectivity. *Trends in Cognitive Sciences, 10,* 519–525.

Salmelin, R., Service, E., Kiesila, P., Uutela, K., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology, 40,* 157–162.

Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J., & Lounasmaa, O. V. (1997). Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience, 77,* 49–55.

Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology, 32,* 11–22.

Sasaki, Y. (2007). Processing local signals into global patterns. *Current Opinion in Neurobiology, 17,* 132–139.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976a). Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology, 39,* 1288–1319.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976b). Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. *Journal of Neurophysiology, 39,* 1320–1333.

Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976c). Quantitative studies of single-cell properties in monkey striate cortex. III. Spatial frequency. *Journal of Neurophysiology, 39,* 1334–1351.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology, 79,* 3272–3278.

Schormann, T., Henn, S., & Zilles, K. (1996). *A new approach to fast elastic alignment with applications to human brains.* Paper presented at the Visualization in biomedical computing : 4th international conference, Hamburg.

Schormann, T., Henn, S., & Zilles, K. (1996). A new approach to fast elastic alignment with applications to human brains. In K. H. Höhne & R. Kikins (Eds.), *Visualization in biomedical computing: 4th international conference* (Vol. 1131, pp. 337-342). Hamburg: Springer-Verlag.

Schroeder, C. E., Mehta, A. D., & Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex, 8,* 575–592.

Schwabe, L., Obermayer, K., Angelucci, A., & Bressloff, P. C. (2006). The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. *Journal of Neuroscience, 26,* 9117–9129.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain, 115 Pt 1,* 15–36.

Shibasaki, H., Ikeda, A., & Nagamine, T. (2007). Use of magnetoencephalography in the presurgical evaluation of epilepsy patients. *Clinical Neurophysiology, 118,* 1438–1448.

Simmons, W. K., Bellgowan, P. S., & Martin, A. (2007). Measuring selectivity in fMRI data. *Nature Neuroscience, 10,* 4–5.

Smyth, D., Willmore, B., Baker, G. E., Thompson, I. D., & Tolhurst, D. J. (2003). The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *Journal of Neuroscience, 23*, 4746–4759.

Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *Neuroimage, 35*, 836–852.

Stenbacka, L., Vanni, S., Uutela, K., & Hari, R. (2002). Comparison of minimum current estimate and dipole modeling in the analysis of simulated activity in the human visual cortices. *Neuroimage, 16*, 936–943.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*, 869–873.

Sun, P., Ueno, K., Waggoner, R. A., Gardner, J. L., Tanaka, K., & Cheng, K. (2007). A temporal frequency-dependent functional architecture in human V1 revealed by high-resolution fMRI. *Nature Neuroscience, 10*, 1404–1406.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262*, 685–688.

Tanaka, K. (1997). Columnar Organization in the Inferotemporal Cortex. In K. S. Rockland, J. H. Kaas & A. Peters (Eds.), *Extrastriate Cortex in Primates* (Vol. 12, pp. 469–498). New York: Plenum.

Tarkiainen, A., Cornelissen, P. L., & Salmelin, R. (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain, 125*, 1125–1136.

Taulu, S., Kajola, M., & Simola, J. (2004). Suppression of interference and artifacts by the Signal Space Separation Method. *Brain Topography, 16*, 269–275.

Teyler, T. J., Cuffin, B. N., & Cohen, D. (1975). The visual evoked magnetoencephalogram. *Life Sciences, 17*, 683–691.

Thorell, L. G., De Valois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research, 24*, 751–769.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23*, 775–785.

Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research, 41*, 414–419.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*, 753–759.

Tootell, R. B., Nelissen, K., Vanduffel, W., & Orban, G. A. (2004). Search for color 'center(s)' in macaque visual cortex. *Cerebral Cortex, 14*, 353–363.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience, 15*, 3215–3230.

Tootell, R. B., Silverman, M. S., Switkes, E., & De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science, 218*, 902–904.

Tootell, R. B., Tsao, D., & Vanduffel, W. (2003). Neuroimaging weighs in: humans meet macaques in "primate" visual cortex. *Journal of Neuroscience, 23*, 3981–3989.

Tootell, R. B. H., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition: fMRI in human visual cortex. *Trends in Cognitive Sciences, 2,* 174–183.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*, 670–674.

Ungerleider, L. G., & Pasternak, T. (2003). Ventral and Dorsal Cortical Processing Streams. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (Vol. 1, pp. 541–562). Cambridge: MIT Press.

Uusitalo, M. A., Virsu, V., Salenius, S., Näsänen, R., & Hari, R. (1997). Activation of human V5 complex and rolandic regions in association with moving visual stimuli. *Neuroimage*, *5*, 241–250.

Uutela, K., Hämäläinen, M., & Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *Neuroimage, 10*, 173–180.

Van Essen, D. C. (2003). Organization of Visual Areas in Macaque and Human Cerebral Cortex. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (Vol. 1, pp. 507–521). Cambridge: MIT Press.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron, 56,* 366–383.

Vanni, S., Revonsuo, A., & Hari, R. (1997). Modulation of the parieto-occipital alpha rhythm during object detection. *Journal of Neuroscience, 17*, 7141–7147.

Vanni, S., Tanskanen, T., Seppa, M., Uutela, K., & Hari, R. (2001). Coinciding early activation of the human primary visual cortex and anteromedial cuneus. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 2776–2780.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex, 3,* 79–94.

Verstraten, F. A., Cavanagh, P., & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research, 40,* 3651–3664.

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology, 18,* 1–10.

Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). An fMRI study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Current Biology, 10,* 1455–1458.

Williamson, S. J., Kaufman, L., & Brenner, D. (1978). Latency of the neuromagnetic response of the human visual cortex. *Vision Research, 18,* 107–110.

Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, *287*, 1273–1276.

Vinje, W. E., & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *Journal of Neuroscience*, *22*, 2904–2915.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79*, 1574–1578.

Woods, R. P., Grafton, S. T., Watson, J. D. G., Sicotte, N. L., & Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography, 22*, 153–165.

Zeki, S. M. (1973). Colour coding in rhesus monkey prestriate cortex. *Brain Research, 53*, 422–427.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology, 236*, 549–573.

Zipser, K., Lamme, V. A., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience, 16*, 7376–7389.