The behavioural and electrophysiological aspects of the temporal context of face processing

“I shouldn’t know you again if we did meet,” Humpty Dumpty replied in a discontented tone. ... “You are so exactly like other people.” “The face is what one goes by, generally,” Alice remarked in a thoughtful tone. That’s just what I complain of,” said Humpty Dumpty. “Your face is the same as everybody has – the two eyes, so – “ (marking their places in the air with his thumb) “nose in the middle, mouth under. It’s always the same. Now if you had the two eyes on the same side of the nose, for instance – or the mouth at the top – that would be some help.”

(Lewis Carroll: Alice in Wonderland)

Zimmer, Márta
Dissertation

Budapest
2009
Introduction and the concept of context in visual perception 4
Temporal context 7
1. Priming and/or adaptation 9
   1.1 Priming 9
   1.2 Adaptation 14
      1.2.1 Adaptation effects in the human visual system 15
      1.2.2 Facial adaptation 21
         1.2.2.1 Distortion 22
         1.2.2.2 Attractiveness 23
         1.2.2.3 Ethnicity 24
         1.2.2.4 Expression 24
         1.2.2.5 Identity 25
         1.2.2.6 Gaze 26
         1.2.2.7 Viewpoint 26
         1.2.2.8 Gender 27
      1.2.3 Faces and the brain – adaptation effects mirrored in brain activity 28
         1.3.1 Non-human primate single cell recordings 28
         1.3.2 Electrophysiology 29
            1.3.2.1 Intracranial recordings 29
            1.3.2.2 Scalp recordings – the N170 and later components 30
            1.3.2.3 Magnetoencephalography (MEG) studies 32
         1.3.3 Imaging studies 32
            1.3.3.1 fMRI adaptation 33
      1.4 Adaptation and/or priming? – fMRI data 34
      1.5 Models of adaptation and functional benefits 36

Experimental results
2. How can our perception be biased when a stimulus appears BEFORE the target image? – adaptational results 41
   2.1 Electrophysiological correlates of visual adaptation to faces and body-parts in humans 41
   2.2 Position-specificity of facial adaptation 55
   2.3 Adaptation duration affects the spatial selectivity of facial after-effects 61
   2.4 Electrophysiological correlates of face distortion after-effects 73
   2.5 Modulation of the facial after-effects by transcranial direct current stimulation 81
   2.6 Discussion 88
3. How can our perception be biased when a stimulus appears AFTER the target image? – backward masking results 104
   3.1 Early correlates of face visibility 104
3.2 Discussion
4. General discussion
5. Acknowledgements
6. Publications related to thesis points
7. References
Introduction and the concept of context in visual perception

From an economical and mathematical point of view we can regard the world around us as a huge set of information (objects and/or events). Much of this information is irrelevant to our daily needs, while some is absolutely essential. To exploit this information effectively, human beings are equipped with specialized sensory machinery that captures this information and translates it into the language of the nervous system. The brain refines this translated information into neuronal “description” of behaviour-relevant objects and events in the environment. Some of these descriptions reach awareness, allowing us to formulate conscious plans in order to be able to interact with those objects; other descriptions are used to guide immediate or reflexive reactions to objects and events (Sekuler and Blake, 2002). Our perception provides us with a useful representation of the world, one that allows us to interact safely and effectively with our environment. Thus, perception accentuates important and diminishes irrelevant information. In any given situation our attentional system aids us in distinguishing between relevant and irrelevant information, so that we can focus on solely the relevant information (object or event) and in consequence we restrict the set of information to a hopefully minimal subset. However, not a single sensory stimulus appears isolated and alone in a hermetically closed world. Therefore both the perception and neurophysiological responses to a targeted input signal depend strongly on both its spatial context (in the case of object recognition what surrounds a given object) and its temporal context (recent experiences in the past or even those that are going to be observed later in future, in a very small time frame later) (Schwartz et al., 2007). Though contextual influences (both in time and space) extend to different sensory modalities (for example to audition (Oxenham, 2001), somatosensory processing (Wallace et al., 2004), olfaction Kadohisa et al., 2006), taste (Gent et al., 1978; Bujas et al., 1991), most studies focus on visual perception, including brightness (Eagleman et al., 2004), orientation (Bednar, 1997), colour (Engel, 2005), motion (Kohn and Movshon, 2004, Mather et al., 1998), shape (Suzuki and Cavanagh, 1998) and more complex objects (such as faces (for a review see Webster et al., 2004)). Spatial and temporal contextual modulations have typically been treated separately, distanced to the point that a rupture in their nomenclature has appeared: for spatial context the terms perceptual illusions/after-effects are in use whereas for temporal contextual effects the phrases non-classical receptive fields and adaptation are in common usage. Nevertheless, they are closely related both functionally and in their impact on vision despite their rather different demands on their respective neuronal substrates (horizontal intra-areal interactions for spatial contextual effect...
and memory for temporal contextual effect). Figure I. shows a summary by Schwartz et al. (2007) depicting the behavioural (psychophysical) and neuronal (electrophysiological) parallels of spatial and temporal context of orientation processing.

**Spatial** context (Tilt illusion)  **Temporal** context (Tilt after-effect)

---

**Figure I.** Example for experimental data on spatial and temporal context effects. Columns show the similarity of spatial and temporal context effects on perception and neuronal response. First row: tilt illusion (spatial, left) and tilt after-effect (temporal, right) biases as a function of the angle difference between context and target orientation. Second row: Orientation discrimination thresholds for spatial and temporal data. Third row: Response suppression in cortical neurons as a function of the angle difference between the target and its context. Fourth row: repulsing tuning shifts in V1 neurons for space and time. Fifth row: contrast-response functions for space and time in cortical neurons, either without (circle) or with (triangle) a context orientation equal to the test. (Schwartz et al., 2007)
A central question of the review was whether spatial and temporal context share a functionally important commonality, in other words if there is a mechanism that underlies similarities between them or whether these similarities are mere coincidences. Most functional characterization of cortical visual areas suggests that their processing either implicitly or explicitly reflects the statistical structure of visual inputs. One obvious source of commonality would be if these statistics themselves exhibited similar properties. Despite this common underlying factor, the functional goals of contextual effects are still hotly debated. Summarizing the above it is safe to state that spatial and temporal contexts induce strong and similar biases, even for low-level features of an image, such as tilt. The similarity renders explanations that work in one, but not in the other domain, less appealing. Current psychophysical and physiological data are far from complete, particularly in terms of comparison between and integration of spatial and temporal contextual effects. Furthermore, some results remain controversial. For example, psychophysical studies of spatial, but not temporal, context have systematically investigated the separation of the target and contextual stimuli (for example Durant and Clifford, 2006), whereas in adaptation, mostly simpler manipulations of adaptation timescale have been tested (Magnussen and Johnsen, 1986). The neurophysiological data are even less comprehensive. Although some theoretical overlap between spatial and temporal context can be demonstrated, it is very hard to bind these data together or align them, because of the many differences in experimental design, such as the species, the alertness of the subjects (anesthetized or awake), the stimulus configuration or the time-scales of adaptation. Because of these experimental and methodological difficulties most researchers investigate spatial and temporal context separately.

Thus here we only focus on temporal context narrowed down to the area of face perception. First, I will summarize our current knowledge about temporal context then I review our own experimental data.

**Temporal context**

By varying the temporal context of a given visual stimulus we can achieve that sometimes the same retinal image may lead to substantially different perceptual experiences. Perception of a given object stimulus (target, S₂) can be biased or impaired by another stimulus, presented before (S₁) the target. We can refer to this phenomenon as forward masking (Eriksen and Collins, 1965; Smith and Schiller, 1966)/priming (Henson et al., 2000, 2002, 2003)/adaptation (Thorndike, 1901; Webster and MacLin, 1999). On the other hand, presenting a stimulus (S₃)
after the given target changes the perception of the former stimulus as well, a phenomenon called *backward masking* (BWM, Raab, 1963; Eriksen and Collins, 1965; Schiller, 1965; Smith and Schiller, 1966; Breitmeyer, 1984) (see also Figure II.).

**Figure II.:** Demonstration of the concept of temporal context in an experimental design and various names of these phenomena. Perception of a target image (S₂) can be biased either by a stimulus presented before S₂ (S₁) or by a stimulus presented after S₂ (S₃). While the former is often referred to as forward masking/priming/adaptation the latter is usually mentioned as backward masking.
1. Priming and/or adaptation

1.1 Priming

One of the clearest and simplest demonstrations of temporal context effect is the repetition effect. The processing of a repeated stimulus is often associated with a reduction in neuronal activity, as measured by single-cell recording (Li et al., 1993) or by haemodynamic imaging techniques like PET (Squire et al., 1992) and fMRI (Grill-Spector et al., 1999). The phenomenon of reduced haemodynamic response for repeated relative to un-repeated stimulus, is called repetition suppression (RS) and its neuronal analogue is “response suppression” what refers to a reduced firing rate of neurons to repeated stimulus. This effect has also been demonstrated using electroencephalography (EEG) (Gruber and Muller, 2005).

For repeated stimuli (for the second, third and consequent presentations of a given stimulus ‘A’, A-A-A-...) we observe reduced firing rate in macaque IT neurons using single cell recording technique (Li et al., 1993, see Figure 1.1) while we do not find such a suppression effect for a new stimulus which does not excite the observed neuron.

Figure 1.1: Object repetition effect in a neuron in macaque IT cortex. Averaged response to stimuli appearing as samples (first presentation), non-matches (first presentation of a new stimulus after a sample) or matches (repeated presentation of the sample stimulus). (Li et al., 1993; Grill-Spector et al., 2006)

Figure 1.2: Object repetition effects measured with fMRI: single subject data. Time course of fMRI activation in object-selective cortex. The black line shows the first presentation of a target stimulus, while the red line illustrates the repeated versions of the same stimulus. (Sayres and Grill-Spector, 2006; Grill-Spector et al.,
At the level of haemodynamic changes this phenomenon is often referred to as fMRI adaptation (fMRIa or fMR-A, Grill-Spector et al., 1999; Grill-Spector and Malach, 2001; Sayres and Grill-Spector, 2006) (Figure 1.2, see later in Chapter 1, Box 1.).

At the level of electric/magnetic field-change exploiting techniques we can concentrate our analyses on event related potentials (ERPs and the magnetic analogue of it is MEP), especially to changes in amplitude (see also in Figure 1.3) and/or in latency. This effect occurs both in conceptual (Matsumoto et al., 2005) and perceptual priming tasks as we have demonstrated as described later in our own findings, (Kovács et al., 2005, 2006, 2007 or Zimmer and Kovács, 2009) (for theoretical differences between conceptual and perceptual priming see later in this chapter).

We can also concentrate on changes in brain oscillations. Typically authors find decreased high frequency (>40Hz, gamma-) power and repeated stimulus presentation also results in a decrease in phase synchrony between distant electrode sites (as was found by Gruber et al. (Gruber et al., 2004) who used line drawings as stimuli, see Figure 1.4).
With the help of EEG/MEG registration we are able to localize the source of these electric/magnetic signals – these results point to the role of prefrontal cortex.

Desimone and his colleagues suggest that the theoretical background of this repetition effect may be a “sharpening” effect (Desimone, 1996 or for a review see: Henson and Rugg, 2003). According to this explanation repeated presentation of a stimulus results in sparser representation. At the first presentation of a stimulus we can get a wider neuronal representation, however, for subsequent presentation neurons, that code features irrelevant to the identification of the given repeated stimulus show RS. So it is a learning-process where the distributed representation will be sparser or sharpened and correlates with a decreased neuronal/haemodynamic activity. Furthermore, Wiggs and Martin (1998) proposed that this “sparser” representation allows more efficient stimulus processing, which can result in behavioural priming effects. However, it is difficult to link neuronal and haemodynamic changes because the small temporal effects of RS are undetectable for fMRI, the single-cell recordings are biased towards larger excitatory neurons and when we would like to concentrate on LFP we should not forget, that changes in presynaptic activity do not perfectly correlate with (and cause) changes in postsynaptic activity (Grill-Spector et al., 2006). Other explanations about the theoretical background of RS will be discussed later in Chapter 1.5.

Behavioural measurements show a difference in accuracy caused by the repeated presentation, a phenomenon called priming. Priming refers to a change in the speed (faster reaction time), bias (examples include a bias to produce that stimulus when generating responses) or accuracy of processing the following stimulus (more accurate identification of a degraded version of the stimulus), following prior experience with the same (repetition priming) or a related (priming) stimulus (Henson, 2003). Although the content of priming is also used in the literature of implicit memory (where it is called conceptual priming) in my dissertation I will focus on this construct as a tool for mapping the stages of stimulus processing (perceptual priming). In the latter case the hypothesis is the following: if a given brain region (R) shows a reduced repetition effect to a target stimulus that only differs from the prime on one given dimension (D) then the processes subserved by R are sensitive to D. However, there is a dissociation between conceptual and perceptual priming. For example Alzheimer’s patients have significantly lower accuracy in conceptual, but not in perceptual priming experiments (Gabrieli et al., 1994). The relevance of such discrimination between conceptual and perceptual priming is also confirmed by Schacter and his colleagues (2004, 2007), who emphasize that the reduction in cortical activity during priming involves at least two different
mechanisms that are associated with different cortical areas. Such effects could occur either early in the processing stream at the level of object recognition in sensory cortices (perceptual priming – middle and inferior occipital gyri, fusiform gyrus in a modality-dependent manner) or at later stages during semantic classification in frontal and temporal cortices (conceptual priming – left inferior frontal gyrus/BA44 in a modality-independent manner) (Wig et al., 2005, Schacter et al., 2007). This hierarchical distinction is also mirrored in task-dependency - the prefrontal areas are significantly modulated by task changes, while the fusiform gyrus appears insensitive to changes in tasks.

There is another question to be answered: is there any connection between behavioural and neuronal/haemodynamical repetition related effects? If so, how can it be localized in the brain? To answer this question Wig et al. (2005) combined transcranial magnetic stimulation (TMS) and fMRI to provide compelling evidence of a causal link between neuronal and behavioural priming. TMS of the left prefrontal cortex during an object-classification task disrupted subsequent markers of both behavioural and neuronal priming in frontal and lateral temporal regions, whereas neuronal priming in early perceptual regions was unaffected. In an object-classification task Sayres and Grill-Spector (2006) obtained evidence that repetition of objects was accompanied by neuronal priming in the lateral occipital complex (LOC) and the posterior fusiform gyrus (FG), however, neuronal priming in earlier visual areas was unrelated to the facilitation in response time (behavioural data) that accompanied repeated classification. In another study Xu and colleagues (2007) have found a full dissociation between repetition suppression and behavioural performance in the scene-selective region of the ventral visual pathway, the parahippocampal place area (PPA, Epstein et al., 1998, 1999), whereas responses from two frontal regions, especially the left insula and the anterior cingulate cortex, mirrored this performance, consistent with their role in decision-making. These data are in line with non-human primate single-cell recording data, as McMahon and his colleagues (2007) have found that inferior temporal (IT) cells of macaque monkeys show no significant correlation between behavioural and neuronal priming effect.

To interpret how priming works Schacter and his colleagues (2007) suggest a two-way model (see also in Figure 1.5):

- “Sharpening” period: exposure to a stimulus in a sharper, more precise neuronal representation of that stimulus (for more details see later in Chapter “Models and functional benefits”). This function is located in the posterior regions (the most important role of these areas is coding of the perceptual representation of the perceived object) and the anterior regions (conceptual properties)
A system to account for response-specific effects and showing higher correlation values between behavioural and neuronal effects (in the prefrontal cortex in a top-down manner).

Figure 1.5: Proposed components of priming. (a) A lateral view of the left hemisphere, tilted in the dorsal-ventral plane to expose the ventral surface. (b) A ventral view of the left and right hemispheres. Abbreviations: A – anterior, D – dorsal, L – left, R – right. The colour coding of anatomical regions is intended to serve as a heuristic for the proposed components. The colour gradient within the ventral visual stream (blue) is intended to represent approximately the gradient of stimulus specificity that has been observed within these regions (Schacter et al., 2007).

Other researchers suggest that no complete homology can be established among this plethora of similar phenomena such as repetition priming, behavioural priming, repetition suppression in neuroimaging and response suppression in single unit responses in monkeys. Several important findings describe results in which repeated exposure to an object caused increase and not the typical decrease in activation (for example in an experiment of Henson and colleagues (2000) using familiar and non-familiar faces as stimuli, a repetition suppression effect could only be demonstrated for familiar faces, however, novel faces led to an enhancement following repetition). Single cell studies similarly suggest that response suppression is not ubiquitous and that repetition effects in single units may be sensitive to the contexts and demands of different experimental procedures (Bichot and Schall, 1999, 2002).
1.2 Adaptation

It’s a fundamental property of all living creatures that they adapt to their environment, thus improving their chances of survival. This phenomenon, as I have already mentioned, could appear in each sensory modality – and the effect can be experienced not only in artificial experimental situations but also in everyday life. Just think about the following situation:

- In a perfumery we can not “process” more than 3 or 4 fragrances in a row - we have to smell a very intensive non-perfume scent (like coffee) inbetween scents in order to smell well again.
- After a very loud rock concert we have to talk to each other louder in order to hear the others’ voice.
- In a snowy evening playing snowballs without putting gloves on at home when washing hands cold water could appear warmer (or even hot).
- After eating a very salty soup for dinner an extremely delicious meal could appear saltless.
- Or let us think about men who work in a sewage farm! Without adaptation they must have given up their work in a moment due to the terrible whiff.

So we can talk about this phenomenon or about its behavioural manifestation, the after-effects in each sensory modality. Adaptation is often referred as “the psychologists’ microelectrode” (Frisby, 1979) using the logic that if adaptation specific to a stimulus attribute can be demonstrated, then it can be concluded that the given attribute is encoded explicitly by the visual system.

Although on the one hand some form of visual adaptation occurs over seconds or minutes and can produce dramatic changes in perception (after-effects), on the other hand, let us think about the development of the human visual system or the evolution that we also refer to as adaptation and that takes years or millions of years. Although the timescale of the phenomenon is widely variable the function is common: “to fit the mind to the world” – as Rhodes and her colleagues would say (Rhodes et al., 2003).
Studying the human sensory system is getting more and more elaborated; and in certain modalities (such as vision and hearing) our knowledge is deeper and more precise. In the field of perception research most studies engage in vision. In the next section of my dissertation I will summarize recent results regarding adaptation and after-effects with special attention to what they can add to our knowledge about face perception and representation in the human brain. However, one of the biggest problems of the visual perception is that the visual environment is not fixed - it changes from time to time, and thus in principle it could not be optimally represented by a visual system with fixed properties. In the lack of such a fixed coding/representational network we can also define adaptation as a tool that recalibrates our (visual) coding so that our vision is normalized according to the stimuli that we are currently exposed to. In other words what our visual system is able to do is difference detection between the target stimulus and the recently defined norm.

1.2.1 Adaptation effects in the human visual system

If we base our judgement on raw data of the human brain we will come to the conclusion that vision is “overrepresented” in our brain. More than half of the brain is involved in visual processing, distributed in more than twenty different brain areas. Nevertheless, the processing of the visual information does not begin at the level of the brain. The first stage of the processing of photoinformation is the retina. The light adaptation of the retinal photoreceptors is a well-known phenomenon and underlies our ability to function equally in a dark room and in wonderful sunshine. Adapting an earlier interpretation in such a situation the norm (average or the origin) of light intensity range will be biased and the whole system will be recalibrated around this newly defined norm. Since our perceptual system always detects the difference, it is possible that viewing a picture for a few seconds with high contrast difference, and later fixating to a homogenous surface (without any pattern) we can see the contrast-polarity changed version of the original image. The presence of such afterimages is also evidence for the existence of the phenomenon of adaptation. However, this phenomenon also appears at the cortical level where we process the pattern of an image. Such adaptable properties are the orientation (Gibson and Radner, 1937), spatial frequency (Pantle, 1970), motion (Mather et al., 1998) or proportions of a stimulus/image (Suzuki and Cavanagh, 1998).
In 1937 Gibson and Radner have found that after staring at a pattern of tilted lines or gratings, subsequent lines appear to have a slight tilt in the opposite direction, a simple but intriguing visual phenomenon, called tilt after-effect (TAE, Gibson and Radner, 1937; Bednar, 1997, Figure 1.6).

Whereas the earliest TAE studies focused only on horizontal and/or vertical directions later it was found that although subjects do not report all difference in the perceived orientation spontaneously (neither in vertical nor in horizontal lines after adaptation), when they had to choose a line parallel to the test one they made systematic errors similar to the original results (Mitchell and Muir, 1976; Bednar, 1997). Therefore researchers conclude that the angular separation between the adaptor and test line (or grating) is the important parameter.

The prevailing theory for the TAEs attributes them to lateral interactions between orientation-specific feature-detectors in V1 (Tolhurst and Thompson, 1975). While a subject fixates on a tilted line-pattern the inhibitory connection strength between neurons showing bigger activation increase temporarily, causing changes in the perception of subsequent orientations. This is the combined result of the tuning curve and the columnar/hypercolumnar representation of the orientation-selective V1 single cells (Hubel and Wiesel, 1968). Therefore when the orientation of the test stimulus is only slightly different from the orientation of the adaptor the most strongly responding units are the ones with orientation preferences further away from the adapting line, resulting in a change in the angle perceived (Bednar, 1997). This can mean that the neuronal mechanisms underlying this effect are located in the primary visual cortex. However, it has been found that the orientation-sensitivity of neurons involved in TAE is independent from the defining cue. Such neurons have been found in monkey V1 (Grosot et al., 1993), V2 (Peterhans et al., 1986) and IT (Sáry...
et al., 1995) as well. Consequently our current hypothesis about TAE is that it’s caused by more, hierarchically located areas simultaneously.

Another example in the topic of visual motion perception is described first by Aristotle in *Parva Naturalia* (1955) and later by a Scottish painter, Scott Addams (Addams, 1834) who wanted to paint a waterfall. He noticed that after viewing a waterfall moving monotonously downwards, when he blinked on the static canvas it appeared to be moving in the opposite direction, a phenomenon that later came to be called waterfall illusion or motion after-effect (MAE, see also in nowadays Mather et al., 1998; Vidnyánszky et al., 2002). As a summary of the two cases mentioned above we can say that these changes in the perceptual experience are specific to the induced stimuli and the underlying factor is that the adaptor and test image is similar to each other – thus they activate partially or totally overlapping neuron populations with the result that the adaptor decreases the sensitivity (or desensitizes) these neurons selectively and for a longer time. Because of this sensitivity-decrease, the representation of the test image will be biased and we can see it more distinct from the adaptor than it is in reality.

From these examples we might think that adaptation to any visual object is a well-defined phenomenon with clear underlying neuronal mechanisms. Unfortunately, this is just a dream. Whereas the orientation of a tilted line or the direction of motion can be described well by a number between 0 and 360 (the degree of direction or tilt) most visual images are multidimensional from a lot of view and these dimensions are not as easy defined as direction.

Notwithstanding with the development of methods and techniques in the last decades the number of adaptation studies that used images, which are more complex visual objects than two dimensional line drawings or human faces, is steadily increasing. From a mathematical point of view the task of vision researchers was to find stimuli which are more complex than a tilted line but because of their properties (e.g. symmetry) they are parametrically well defined. In 1944 Koehler and Wallach have shown that after a horizontally elongated ellipse adaptor subjects judged the image of a test circle as a vertically elongated ellipse (Koehler and Wallach, 1944). The name ‘Opponent shape after-effect’ refers to the phenomenon in which adaptation to a certain shape attribute distorts a subsequently presented neutral test shape in a repulsive manner (Suzuki, 2003). The first systematic evidence of adaptation to a two-dimensional image attribute (not explained by local contour adaptation) was reported by Regan and Hamstra (1992). By examining both the discriminability and after-effects of vertical and horizontal elongation while the scale and the
exact shape (e.g., ellipses versus rectangles) of the stimuli were varied, the authors provided evidence that the visual system has a mechanism to directly code global aspect ratio. Similar opponent shape after-effects have been subsequently demonstrated for other shape attributes. In a given situation we can get adaptation effect even if the geometry of the adaptor and test image is very different. For example the shape of an hourglass (concave) also can bias the perception of the 4 test squares to the opposite (convex) direction.

Figure 1.7: The adaptors and test stimulus used in the experiments of Suzuki (Suzuki, 2001). (a) shows the contour-line adaptors, (b) shows the line-texture adaptors, (c) illustrates the test stimulus used to measure the orientation after-effect. (d and e) show the convex and concave distortion is induced by the hourglass (and X-patterned lines) and diamond pattern,(and its line-analogue), respectively (Suzuki, 2001).

In 2001 Suzuki also demonstrated that brief adaptation dissociates global convexity after-effects from local tilt after-effects. He used tilted gratings as adaptors and found that with prolonged adaptation, the four tilted grating squares locally produce TAEs, resulting in the overall convex appearance of the four test squares, similar to after-effects induced by a concave hourglass (see also in Figure 1.7). With brief adaptation (less than 134ms), however, the gratings produce no after-effect (Suzuki, 2001, 2003).

Maybe the most important question is whether these opponent shape after-effects can also be described with the help (perhaps the summation) of low-level mechanisms (like local contour after-effects) or whether they are global shape AEs. Based on our knowledge about the size and organization of the visual receptive fields at the level of higher vision we can say that if the adaptor and test stimuli are different in a given property (for example relative position or
but the neurons are invariant to this altered stimulus (namely the neurons respond similarly to the adaptor and test stimuli that appeared in a new relative position or changed in size) and that’s why AE appears, then we can conclude that the neurons involved in the effect can be located in higher level visual areas (like inferior temporal cortex, IT). As I’ve already written Suzuki and Cavanagh (1998) and Suzuki (2001) have found that the convexity-concavity AEs were not sensitive to the changes in relative position, size and orientation. We know that the IT cortex of primates is strongly sensitive to the global shape of the stimulus (Logothetis and Sheinberg, 1996, Tanaka, 2000), the neurons located in IT are size-invariant (Schwartz, 1983) and are sensitive to a wider orientation range like V1, V2 or V4 neurons (Vogels and Orban, 1994). These results suggest that the shape AEs can be located at a higher level visual area such as IT.

The next step of the hierarchy of object-complexity that was observed in an adaptation experiment was cars and 3D wire-like images. The main problem with such complex images is their multidimensionality. We can neither parameterize them well nor sort them to simple opponent subcategories and create fine gradation among them by special techniques like morphing\(^1\) (Glasbey and Mardia, 1998). Perhaps the only feature that can be well parameterized is the viewpoint. From a methodological point of view we have a relatively easy manner to manipulate the viewpoint of a complex image.

---

\(^1\) The animated transformation of one image into another by gradually distorting the first image so as to move certain chosen points to the position of corresponding points in the second image.
In 2005 Fang and He have found that after adaptation to a car or a three dimensional wire-like image viewed either 15 or 30 degrees from one direction, when the same object was subsequently presented near the frontal view, the perceived viewing directions were biased in a direction opposite non-overlapping adaptor and test images (see also in Figure 1.8). They also found that the magnitude of this viewpoint AE depends on the angular difference between the adapting and test viewing angles and grows with increasing duration of adaptation.

In one of our earlier adaptation experiments (Kovács et al., 2005b, non-published data) we used computer generalized 3D objects consisted of 5 geons (see Figure 1.9, more details about geons from Biederman, 1987). The goal of this experiment was to test directly whether prolonged viewing of a visual non-face object can result in adaptation of the neuronal mechanisms responsible for both viewer-centred and object-centred object coding and thus lead to corresponding view invariant as well as viewpoint-specific after-effects. We aimed at revealing different types of the after-effects by instructing the observers to make different judgements about the test stimuli. In particular, we argued that viewpoint after-effects can be uncovered by asking the observers to report the angle from which they perceived the test object, whereas shape-specific view-invariant after-effects will be found when observers are required to judge the identity of the test object. In the experiments aiming to reveal shape-specific after-effects, we used a side view of a fish or submarine like object as adaptor (Figure 1.9) and the test images (as the central image of Figure 1.9) were chosen from morph series that were created by morphing fish and submarine images into each other: a full side view image pair and 4 rotated views (15, 30, 45 and 60 deg) of the fish and submarine images. To test for viewpoint after-effects our experimental design was similar to that used by Fang and He (Fang and He, 2005) and the stimuli were different views of the same fish and submarine objects that were used to reveal shape-specific after-effects.

Figure 1.9: Sample 100% fish (left) and submarine (right) stimuli and the 50% morph level (centre) between them from full side view (0 deg) rotated in counter-clockwise direction.
We can summarize our results as follows:

1. Adaptation to a visual object led to shape-specific after-effects. As a result of 5 second exposure to specific 3D rendered objects, in our case images of a fish and submarine, observers’ judgements regarding the identity of the test stimuli were biased and distanced away from that of adaptor image.

2. This shape-specific AE was to some degree viewpoint-invariant. It was present when adaptor and test objects were presented from different viewing angles for up to 30 deg, i.e. it generalized completely across different object views. The transfer was partial, but still significant for test images which viewing angle differed by 45 or 60 deg from that of the adaptor image.

3. Essentially the same view-invariant shape-specific AE was observed for the rotating adaptor condition as well.

4. Adaptation to the same objects that were used to show shape-specific AEs resulted also in viewpoint AEs, which could be revealed by asking the observers to report from which angle they perceived the test objects, instead of judging their identity.

Altogether, our results suggested that both view-specific and view-invariant object representations are susceptible to selective adaptation, and depending on the nature of the observers’ task can lead to view-specific and view-invariant AEs, respectively.

Examining the neuronal background of the opponent shape-after-effects researchers have also located these effects to higher-level visual processing stages, such as IT and/or superior temporal sulcus (STS).

1.2.2 Facial adaptation

Faces are one of the most important - and from another point of view - most complex visual stimuli in our life. Not only do they have complicated three-dimensional structures, but they convey a multiple of perceptual data, including information about identity, gender, race, expression, and direction of gaze, among others (Fox et al., 2008). The perception of these above-mentioned features (or properties) is usually effortless even though faces have very similar structure as visual patterns. So we can say that we are face-experts (Carey, 1992; Gauthier et al., 2000) from a point of view of categorization since we have to decide about a face at the level of exemplars. Therefore we must have a very precisely built representation of faces in our coding system (especially in our brain). One of the most prominent and widely
known explanations about representation of faces is a norm-based or prototype-referenced coding of faces, a multidimensional ‘face-space’, suggested first by Valentine and colleagues (Valentine, 1991; Valentine and Endo, 1992). Since a face is a very complex image AEs can be demonstrated in a lot of different properties (Figure 1.10).

1.2.2.1 Distortion

The AE first described in the face adaptation literature was the so-called face distortion after-effect (FDAE) (Webster and MacLin, 1999). By expanding or contracting the inner features of a face we can create artificially distorted faces. In such a situation the relative position of the eyes, nose and mouth will be increased or decreased, respectively. Adaptation to a expanded/contracted face strongly biases face perception by causing the original (normal) face to appear distorted in a direction opposite to the adapting distortion, namely contracted/expanded, respectively (see also in Figure 1.10., first row).

This effect transferred across different faces and was similar to upright and inverted faces however, there was only weaker adaptation transfer between adaptor and test faces that are in different orientations. The authors have suggested that AEs depend on the whole structure of a face and not simply on the local features of it. We also know that adaptation to an asymmetrically distorted face also causes FDAE, but doesn’t effect the perception of its mirror image (Morikawa, 2005). Robbins and collaborators have shown that the size of FDAE varies as a function of the distortion level of the adaptor (Robbins et al., 2007).
If we try to find the neuronal mechanisms underlying FDAE we will have two possibilities. Either we use a well-planned experimental design (namely “we use the psychologists’ microelectrode” by Frisby) in a psychophysical experiment or we measure brain activity using electrophysiological and/or haemodynamical methods. Researchers manipulated different variables in the experimental design. They hypothesized that varying the size (Zhao and Chubb, 2001) orientation (Watson and Clifford, 2003, Rhodes et al., 2004), spatial frequency content, contrast, colour (Yamashita et al., 2005) or the viewpoint (Jeffery et al., 2006, 2007) of the adaptor and test images, by getting any statistically significant adaptation effect we can say that these FDAEs are not mediated by low-level retinotopic visual mechanisms but higher level of object-based processing. In all of these above-mentioned cases authors were able to show stronger or weaker, but significant adaptation effects.

1.2.2.2 Attractiveness

Some theorists think that the most attractive face is also the most average face (namely the origin /norm or prototype/ of our perhaps evolutionally coded multidimensional face-space) (Langlois and Roggman, 1990 or for reviews see Rhodes and Zebrowitz, 2002). With the help of adaptation technique we can test this hypothesis. Based on earlier findings we can hypothesize that this multidimensional coordinate-system is not fixed – exposure to a given face that is special in a given aspect (since it is distorted, or it has an own identity, gender, race or anything else) is able to recalibrate the whole system, especially this transformation defines a new origin. Therefore, after adaptation to a distorted face biases not only the magnitude of endorsed ratio of “normality” of a face but also the rating of its attractiveness (Rhodes et al., 2003). So we can conclude that the most attractive faces are always the ones that can be found very near to the actual norm.

In these earlier two cases authors used distorted (and original) faces as stimuli which can be relatively easily generated by different programs or mathematical algorithms. In spite of the multidimensionality of faces by a special technique, called morphing or warping (Glasbey and Mardia, 1998) we are also able to form a finely graded series of intermediate images where neighbouring faces differ equally from each other in a lot of different aspects (as relative position of face parts, the identity or the gender of a face). With the help of morphing we are able to show face adaptation related after-effects not only on distorted faces but also on natural facial categories.
1.2.2.3 Ethnicity

After-effects also occur in the perception of races. Adaptation to a Japanese or Caucasian faces significantly biases the perceived neutral image in the sequence between them so that the adaptor face itself appears more neutral since long exposure shifts the perceived neutral point towards the adapting image (Webster et al., 2004). In 2007 Jaquet and her collaborators investigated whether this after-effect is generated in high-level visual coding that is sensitive to the social category information in faces, or in earlier visual coding sensitive to simple physical differences between faces (Jaquet et al., 2007). They caricatured the race information in the stimuli and created wide race-continua between SuperCaucasian and SuperChinese. In a FDAE experiment the authors have found larger AE following adaptation to faces from different race categories, than for faces that were placed at the same physical distance on the morphed continua, but fell within the same race category. These results suggest that these AEs for Chinese and Caucasian faces reflect the recalibration of face neurons tuned to high-level social category information. In a category contingent AE study Little and his collaborators (Little et al., 2008) have asked their subjects to make judgements about the eye-spacing of face stimuli which differ in ethnicity (European or African), age (adult versus infant) or species (human or monkey faces). Based on their own results they concluded that functionally distinct neuronal population code faces of different ages, ethnicities and species.

1.2.2.4 Expression

Face adaptation related after-effect can be observed between not only faces of different people but also between different emotional states (expressions) of the same face. Since emotions are well distinguishable, involve an innate and universal pattern of behaviour, and are perceived in terms of a small number of relatively discrete categories by morphing we are able to generate intermediate levels of emotional continua. Webster and his colleagues (Webster et al., 2004, see for an example Figure 1.10 fourth row) have shown that in a happy – angry morph, previous exposure to an angry/happy face induced clear differences in the expression boundary shifting it systematically towards the adapting emotion. Fox and Barton have found (Fox and Barton, 2007) that this adaptation was not based on low-level image properties, as a different image of the same person displaying that expression produced equally robust AEs. Smaller but still significant AEs were generated by images of different individuals, however, no after-effect have found for non-face visual, auditory, or verbal emotional information. These results are clear evidence underpinning the idea that adaptation affects at least two neuronal representations of emotions – one specific to the individual and other that represents
expression across different facial identities. In a study published in 2008 the same group have reported that facial expression AEs are not due to local adaptation to image elements but due to higher-level adaptation of neuronal representations that involved both facial features and facial configuration (Butler et al., 2008).

### 1.2.2.5 Identity

One of the most cited studies from the adaptation literature has shown that subjects can be also adapted to the identity of a given face (Leopold et al., 2001). From their face image database they created a multidimensional face-space by morphing (see also in Figure 1.11).

Figure 1.11: A multidimensional face-space generated by Leopold et al. (Leopold et al, 2001). Green ellipses show the original faces, whereas the blue ellipse marks the average face. Each grey line is an ‘identity-trajectory’, where numbers indicate the ‘identity strength’ of a given face. Anti-faces on the opposite side (orange ellipses) of the mean created from the original faces have negative identity strengths.

They also created the morphed average of the faces (norm) and the anti-faces of each experimental stimulus by measuring the differences between the original and the average face and distorting the norm with these differences to the opposite direction for each feature. Thus an anti-face is on the same identity trajectory but on the opposite side where the original face is. Leopold and his colleagues have found that longer adaptation to an individual face generated a significant bias in the subsequent perception of face identity (face identity after-effect, FIAE). They have also shown that when the adapting face was on the same identity trajectory but on the opposite side of the average face (anti-face), perception was biased in the direction of the test face, even at values quite close to the average face subjects identified stimuli as the test face. In contrast, when the adapting and test faces were on different trajectories, performance was impaired since perception was biased in a different direction. FIAE appears not only for real faces but also for imagined ones (Ryu et al., 2008) so identity information contained in imagined and real faces produces similar behavioural results. These
findings are evidence for the shared neuronal networks that underlie perceived and imagined complex visual objects. Another important result from Moradi and colleagues is that FIAE requires that the face be visible, which is supported by the fact that the effect can be cancelled by binocular suppression or by inattentiuonal blindness, whereas cross-modal interference that does not affect the visibility of the face does not interfere with the AE (Moradi et al., 2005).

1.2.2.6 Gaze

Gaze direction is also a very important signal in social interaction. Jenkins et al. have found that adapting to averted gazes reduces dramatically the sensitivity to gaze in the adapted direction, without reducing sensitivity to gaze opposite to the adapted direction (Jenkins et al., 2006). Since these after-effects generalized across different identities, size and head orientation results lead the authors to conclude that adaptation operates at the level of gaze perception. Altogether these findings suggest that humans have distinct populations of neurons that are selectively responsive to particular directions of seen gaze. Schweinberger and his colleagues observed the neuronal correlates of gaze adaptation by recording event-related brain potentials (ERPs) (Schweinberger et al., 2007). They have shown that the difference in scalp activity between adapted and non-adapted (control) conditions does not start before ~250-350ms. They located this adaptation effect at more anterior right-temporal areas. Inspired by earlier monkey data Calder and his collaborators (Calder et al., 2007) in an fMRI adaptation experiment investigated whether the human anterior STS contains representations of different gaze directions. Adaptation to leftward/rightward gazes produced decreased BOLD signal to left/right relative to right/left and direct gaze probes in the anterior STS and inferior parietal sulcus presenting clear evidence for dissociable neuronal systems for left and right gaze, which has already been hypothesized in the psychophysical experiment of Jenkins et al.

1.2.2.7 Viewpoint

Fang and He tested whether there are neurons representing specific views of objects in the visual system in humans (Fang and He, 2005). One of the tested object categories was human faces where they also found viewpoint after-effect. The magnitude of this AE depended on the difference in the viewing angle between the adapting and test stimulus and grew with the exposition time of the adaptor. In another experiment they measured the face viewpoint AEs for upright faces using adaptors and test faces differing in identity, gender or orientation (by using inverted adaptors) (Fang et al., 2007). The AE showed a strong transfer following
adaptation to other faces however, the transfer in the inverted face adapted conditions was much smaller, indicating that the neuronal coding of upright and inverted faces is different. However, later adaptation experimental results suggest that human neuronal coding of faces is view-specific. For example Jeffery and her colleagues (Jeffery et al., 2006) have found that a FDAE induced in one viewpoint showed only weak transfer to other views. Jiang et al. have shown that FIAEs show only limited transfer across views (Jiang et al., 2006). In another FDAE experiment Jeffery and collaborators obtained the result that cancellation of an AE in one view by an opposing AE in another view decreased as the angular distance between the adapt views increased, suggesting broadly tuned, view-specific coding of face shape (Jeffery et al., 2007).

1.2.2.8 Gender

By morphing between pairs of female and male faces we can create a continuum of more or less androgynous faces. Using these faces as test images we can get adaptation effect at the level of gender processing. After adaptation to a male face Webster et al. have shown that a previously ambiguous face appeared more female and vice versa (Webster et al., 2004). Anthony C. Little and his colleagues have emphasized that gender is a very important factor in face related after-effects (Little et al., 2005). They have shown that visual AEs occurred only when the gender of the faces at exposure and post-adaptation testing were congruent with each other. Sex-contingent effects can be observed for adaptation to eye-spacing, identity and sexual dimorphism and these effects depend on perceptual category rather than structural coding (Bestelmeyer et al., 2007). In view of these findings we can hypothesize that distinct neuronal population code female and male gender information in face perception. As we used the same (gender-based) after-effect in most of our experiments I will elaborate further on the behavioural effect and the underlying neuronal mechanism of this AE in later chapters.

Later I will summarize the behavioural correlates of face adaptation related after-effects and their implications to the underlying neuronal mechanisms. In the next chapter I review the neuronal background of face perception and face adaptation. These data come from monkey single-cell recordings and human experiments using electrophysiological (electro- or magneto-encephalography, EEG, MEG) and imaging techniques (functional magnetic resonance imaging or positron emission tomography, fMRI or PET).
1.3 Faces and the brain – adaptation effects mirrored in brain activity

1.3.1 Non-human primate single cell recording

We know that the inferior-temporal cortex (IT) of the rhesus monkey is involved in the recognition of objects (Logothetis and Sheinberg, 1996; Tanaka, 1996). Furthermore, single-cell recordings by Gross and colleagues (1972) revealed cells related to object and especially face recognition in the superior temporal sulcus (STS) and inferior temporal cortex. Examination of the response properties of these face cells has revealed a great variability showing for example that neurons in IT respond selectively to complex visual images differing in shape, colour, texture and three dimensional properties. The size of the receptive fields (RFs) of IT neurons is large, not retinotopically organized and usually cover both visual fields. The responsibility and stimulus selectivity of the neurons is independent from the location, size (Schwartz et al., 1983), colour (Komatzu and Ideura, 1993), contrast (Nakamura et al., 1994) and other properties of the seen object. The IT cortex also plays an important role in perceptual categorization (Vogels, 1999). It has also been found that the face-selective cells in IT and superior temporal sulcus (STS) have large RFs (Gross et al., 1972) therefore they are less sensitive to the relative position on the RF and to the size of the stimulus (Perrett et al., 1982, Rolls and Baylis, 1986). The response of some IT cells was view-selective whereas others respond to all views (Perrett et al., 1985; Desimone et al., 1984). Recent fMRI studies with monkeys yielded results that are consistent with these early single-cell recording works. Furthermore, studies with humans have also reported neuronal responses that are highly selective to faces as detailed below.
1.3.2 Electrophysiology

1.3.2.1 Intracranial recordings

Studies with temporal epileptic patients enable intracranial recordings of electrophysiological responses in humans to complex visual patterns such as faces. In these studies, we are able to record both evoked potentials over the sulci of the brain via macroelectrodes and single-cell activity via microelectrodes (see also Figure 1.12).

One of the most cited studies from the literature of human electrophysiology of face perception is a pioneering work by Allison and collaborators (Allison et al., 1999a, b) who recorded event-related potentials (ERPs) intracranially in 98 epileptic patients. In a series of experiments, they found distinct types of face-specific ERPs.

Figure 1.12: Illustration of electrode localization in intracranial recordings. Coronal slice of an MRI. An 8×8 grid of electrodes was placed subdurally on the right parietal and temporal lobes (from Allison et al., 1999a).

The earliest activity evoked in visual cortex is reflected by the $P100$ which is sensitive to luminance and other primitive stimulus features. Most $P100$ sites were located in the primary visual cortex, V1 and in the vicinity to V2. The centre of this activity was located in the medial occipital lobe. However, they also found later components specific to face processing. They focused on three cortical regions, namely the ventral face area, the lateral face area, and the anterior face area. The VENTRAL face area consists of four visual regions: the major part is formed by the lateral fusiform gyrus and adjacent IT but medial fusiform gyrus and V4 can also be involved in generating components located here. The LATERAL face area is located in the middle temporal gyri, whereas the ANTERIOR regions are the anterior fusiform gyrus and the entorhinal cortex. Table 1 shows which components are located in the above-mentioned areas.

<table>
<thead>
<tr>
<th>$P150$</th>
<th>$N200$</th>
<th>$P290$</th>
<th>$VP350$</th>
<th>$AP350$</th>
<th>$LP350$</th>
<th>$N700$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventral</td>
<td>Ventral</td>
<td>Ventral</td>
<td>Ventral</td>
<td>Anterior</td>
<td>Lateral</td>
<td>Ventral</td>
</tr>
<tr>
<td>Lateral</td>
<td>Lateral</td>
<td>Lateral</td>
<td>Lateral</td>
<td>Lateral</td>
<td>Lateral</td>
<td>Lateral</td>
</tr>
</tbody>
</table>

Table 1: The face-specific early and late ERPs and their loci in the brain found by Allison and collaborators (Allison et al., 1999a).
Allison et al. (Allison et al., 1999a, b) emphasized that although sometimes already $P150$ shows face-specificity, the earliest clear evidence of face-specific processing is $N200$ which may reflect the structural encoding of a face and its strength (amplitude) was larger in the right hemisphere. The authors also observed top-down influences on these ERP components and found the followings: the ventral $N200$ was unaffected by emotional stimuli, familiarity and identity and it showed no semantic priming, therefore they concluded that $N200$ reflects the operation of a module specialized for the perception of faces. $N200$ was also insensitive to the colour, size and spatial frequency of faces but by presenting inverted faces the amplitude and latency of the component were enhanced (McCarthy et al., 1999). Top-down effects, however, were mirrored only on later components. Such later components were recorded by Dietl et al. in 2005, namely the rhinal $AMTL-N400$ and the hippocampal $P600$ for familiar faces which were recorded by an intrahippocampal depth electrode in temporal epileptic patients (Dietl et al., 2005). For repeated presentation of faces they got decreased $AMTL-N400$ and $P600$.

### 1.3.2.2 Scalp recordings – the N170 and later components

Parallel to intracranial recordings ERP studies measuring scalp electrical responses also appeared in neuroscience research. These studies have found the scalp analogue of $N200$. Because of its shorter latency (~140-200ms) it was called $N170$ (for an example see Figure 1.13).

![Figure 1.13: Demonstration of the N170 ERP component. The figure shows the right laterality of N170, the inversion effect and also demonstrates that the amplitude of this component is much larger to faces than to any other non-face stimuli (from Itier and Taylor, 2004).](image)

Bentin et al. have recorded this negative bilateral component that peaked about 170ms after the presentation of a face stimulus with largest amplitude at the occipito-temporal sites that was much larger to faces than any other non-face images and they also emphasized its right
laterality (especially on electrodes P8 and PO10 according to the international 10/20 system) (Bentin et al., 1996). However, Proverbio et al. in 2006 have shown that this right hemispheric dominance occurs only in male subjects, female individuals show bilateral activity at occipito-temporal sites (Proverbio et al., 2006). Faces of animals were also associated with N170 and its amplitude was similar compared to N170 evoked by human faces (Rousselet et al., 2004). Scrambling, contrast reversal and orientation inversion affect the amplitude and latency of N170 which results in larger amplitude and in slightly but significant delay in latency (George et al., 1996). This latter effect is called face inversion effect which appears only for faces but not for any other inverted objects and this effect can be located in the Fusiform Face Area (FFA) (Rossion et al., 2000; Rossion and Gauthier, 2002; Yovel and Kanwisher, 2005). Furthermore, eyes could be central to what makes faces so special (Itier et al., 2007, Bentin et al., 1996). The question of whether this component encodes only the structural or also the identity information of a face is under hot debate. Most studies suggest no familiarity effect on this early component (Eimer, 2000; Bentin and Deouell, 2000). The effect of emotions on N170 is debated, while some studies suggest that N170 is unaffected by emotional expression (Eimer and Holmes, 2002) others have shown that N170 is modulated by facial expressions (Batty and Taylor, 2003; Blau et al., 2007). Specifically, the amplitude of N170 is greater for fearful than neutral faces (Jiang et al., 2009) and the effect of context (fearful vs neutral) in which a face appears is also reflected on N170 (Righart and de Gelder, 2006). However, not only the spatial but also the temporal context can modulate the N170. For example some ERP studies have investigated the changes of this and later components by repeated stimulus presentation or in experiments based on facial after-effects. It was shown that repeated presentation of the same facial expression (independently from the identity of the face) reduced the amplitude of N170 (Campanella et al., 2002). However, Werheid and her colleagues have shown that ERPs showed priming effect in an emotional 2-AFC experiment both in early and late part of the evoked signal (Werheid et al., 2005). Nevertheless, the early priming effect seemed to start earlier than the N170 component, whereas the later effect occurred between 500 and 600ms after stimulus onset. In another experiment authors have investigated whether the N170 would gradually decreased to repeated faces when face identity is irrelevant to the experimental task (Heisz et al., 2006) and they have shown significant decrease in the amplitude of N170 for repeated presentation of an unattended face. In a viewpoint after-effect study Caharel and her colleagues have also shown amplitude reduction in N170 at right occipito-temporal sites when the identity of the adapting and test faces were the same (Caharel et al., 2008). After 250ms
following the test stimulus onset they also found adaptation effect but its scalp distribution differed from the first one, because this effect was bilateral. However, from the results of intracranial recordings by Allison et al. (Allison et al., 1999a, b) we know that there are also later evoked potentials related to face perception which can also show priming or adaptation effects. Eimer has shown that familiar faces evoked an enhanced negativity at ~300-500 ms after the stimulus onset (N400f) which was followed by an enhanced positivity at 500ms post-stimulus (P600f) (Eimer, 2000). Stefan R. Schweinberger and his collaborators, however, emphasize that the earliest face-related ERP component that shows strong modulation to repeated stimulus presentation is the N250 (or N250r) component over the inferior-temporal regions (Schweinberger et al., 2002, 2004). A parietal N400 also shows modulation by priming (Schweinberger et al., 2002; Cooper et al., 2007).

1.3.2.3 Magnetoencephalography (MEG) studies
In MEG studies we measure the magnetic fields produced by electrical activity in the brain. Liu et al. have found the magnetic analogues of early ERP components, called M100 and M170. These components behave so similar to their “electrical relatives”, however, that M170 amplitude evoked by face inversion does not differ from the one that evoked by upright face stimulus furthermore the M170 amplitude is similar over the left and the right hemispheres (Liu et al., 2000). Liu and colleagues have also shown that M170 was sensitive to success in face categorization and identification (Liu et al., 2002). Adaptation can also be observed on earlier or later MEG components however, studies suggest controversial effects. Whereas Harris and Nakayama found face-selective adaptation on M170 (Harris and Nakayama, 2008) not only for full faces but also for scrambled and even isolated face parts, but the results Schweinberger and his colleagues suggest that both M100 and M170 are insensitive to repetition. They also found the MEG analogue of N250, called M250r which was sensitive to repeated stimulus presentation (Schweinberger et al., 2007). According to the results of Alumit Ishai and her collaborators repetition suppression is also reflected in earlier (~160-185ms) and later (~220-250ms) components of MEG (Ishai et al., 2006).

1.3.3 Imaging studies
Cognitive neuroimaging is a rapidly developing field. The number of found items for a search in Medline for “fMRI and face perception” is more than 800 and the earliest study was published no more than 20 years ago. Here I will not summarize all of these results I just review some facts about face processing and its neuronal background. Four different brain
areas are usually mentioned in connection with face perception – the Fusiform Face Area (FFA, Kanwisher et al., 1997), the Lateral Occipital Complex (LOC, Malach et al., 1995), a region in the posterior STS called Occipital Face Area (OFA) and the superior temporal sulcus (STS) whose importance was emphasized by Haxby et al. (Haxby et al., 2000). It is worth noting that neuroimaging studies also confirm the multidimensional face-space explanation. Loffler and his colleagues systematically varied the facial geometry of their stimuli and found that the fMRI signal increased with increasing distance from the mean face (Loffler et al., 2005).

1.3.3.1 fMRI adaptation
As I have already mentioned before repetition effects are also reflected in haemodynamic signal changes. The blood-oxygen-level-dependent (BOLD) signal in different brain areas will be reduced when the same stimulus is repeatedly presented – a phenomenon called repetition suppression/repetition attenuation/fMRI adaptation (fMRIa or fMR-A) (Grill-Spector et al., 1999, Grill-Spector and Malach, 2001, Cant et al., 2009). fMRI adaptation is a sensitive tool that allows us to investigate the selectivity/invariance of a given neuron population. The underlying idea is the following: recovery of adaptation suggests neuronal representations selective for specific stimulus properties, while adaptation across changes between two stimuli provides evidence for a common neuronal representation insensitive to that change.

Furthermore, the cortically defined fMRIa effect and the behaviourally defined priming effect have both been associated with improved performance and a decrease in activation (in almost all cases) in similar visual areas as a result of stimulus repetition in past functional neuroimaging research.

fMRI-a is also a perfect tool to examine how and where the visual information about the ethnicity, gender and identity of a given face processed. Ng et al. (Ng et al., 2006) have found a distributed system to these processes located in the inferior occipital cortex, fusiform gyrus and the anterior cingulated cortex (ACC). In another adaptation experiment where subjects had to make expression or identity judgements Furl et al. have shown that in both cases a decrease in anterior medial temporal lobe activity appeared in adapted compared to non-adapted conditions (Furl et al., 2007).

If we are to continue this assignment however, it is worth focusing on one of the biggest questions of brain research today which is to what extent BOLD signals are direct reflections of neuronal spiking responses or synaptic activity, or BOLD signals simply provide a general
measure of neuronal “activity” (Krekelberg et al., 2006). To put it more directly: does the change in BOLD signal reflect the same neuronal activity (or changes in activity) as that measured by single-cell recordings. Finally, focusing on the topic of my thesis, the question is the following:

1.4 Adaptation and/or priming? – fMRI data

Some authors like Murray et al. (2006) or Kourtzi and Grill-Spector (2005) refer to repetition suppression and adaptation as the same phenomenon. They suggest that adaptation may be a manifestation of the basic phenomenon of habituation, whereby the system suppresses temporal repetitive stimuli. The behavioural correlate of prolonged habituation is typically reduced sensitivity of the observer to test stimuli which have similar properties as the adaptor. However, Schacter and Buckner (1998) claimed that in contrast to habituation, visual priming can be manifested after a single exposure to a stimulus (prime) and it is preserved in time-scales ranging from seconds to even a year. The underlying mechanisms could be the priming of selective neurons and/or the inhibition of non-selective neurons (irrelevant in the given situation). So if we would like to answer the question of this chapter first of all we have to separate two levels of effect. At the level of behaviour we can clearly distinguish these two phenomena. Priming always (or in most cases) shows a relatively better performance, like shorter reaction time or higher scores (more hits) in accuracy, whereas adaptation generally results in “worse” performance compared to the control (baseline) situation, as if our perception were “deceived” by our sensory system about the actual situation of the environment. However, at the neuronal level, it is not clear whether when we say repetition/response suppression or fMRI adaptation we refer to the same neuronal mechanisms or not. Or to put it more dramatically, are we able to distinguish these two phenomena at the neuronal level? The scientific literature that I summarized above till now was unable to clarify the neuronal borders between priming and adaptation. Although there are some results in the literature further investigations are necessary to clarify this question. For example, Ganel and colleagues (2006) examined whether or not the two aforementioned phenomena reflect the same underlying neuronal mechanism. In the study the authors manipulated adaptation and priming orthogonally and they found some evidence to the dissociation between two phenomena:
• Left LO showed fMRIa but no repetition priming was observed.
• Reduced activation effect in the regions that showed both priming and fMRIa effect were additive rather than interactive.

They concluded that it may be the case that fMRIa reflects a rapid and short-lasting effect related to changes in neuronal firing whereas repetition priming engages slower and more stable changes in synaptic efficacy of the neuronal population (Henson et al., 2003).
1.5 Models of adaptation and functional benefits

There are multiple potential neuronal causes of the repetition suppression (RS) measured with single-cell recording, fMRI and/or EEG/MEG. Here we discuss three main models that have been proposed to account for repetition related reductions in neuronal activity: a; the Fatigue model, whereby the amplitude of firing of stimulus-responsive neurons decreases (Miller and Desimone, 1994; Grill-Spector and Malach, 2001); b; the Sharpening model, whereby fewer neurons respond (Li et al., 1993; Desimone, 1996; Wiggs and Martin, 1998); and c; the Facilitation model, whereby the latency (James and Gauthier, 2006) and/or duration of neuronal activity are shortened (Sobotka and Ringo, 1996; Henson and Rugg, 2003).

1. “Fatigue model” (called “Suppression Model” by James and Gauthier, 2006)

Figure 1.13: The Fatigue model on a schematic three layered neural network. The left panel shows the activity in the input layer caused by a visual stimulus for the first presentation. Blue graphs illustrate the spiking of the neurons with highest response at each stage. According to this model the cause of the reduced BOLD signal for repeated presentation is lower firing rates (central panel). The right panel shows the implication of the model (black line – initial response, grey line – repeated response). The Fatigue model suggests that repetition will reduce overall response in proportion to initial response, but the preferred stimulus and tuning width will not change (Grill-Spector et al., 2006).

According to this model all neurons involved in the first presentation of a stimulus (in the sense that they show increased activity to this stimulus) respond to the repeated presentation but in a reduced manner (see also Figure 1.13). So there is no difference in the pattern of relative responses across neurons, but the mean firing rate declines. One of the biggest problems with this explanation is that it is not reconcilable with the
phenomenon of behavioural priming, especially with shortened reaction time and enhanced accuracy (one of the possible solutions: a reduction in the mean and variance of firing rate allows for greater synchrony of neuronal activity).

2. “Sharpening model”

Figure 1.14: The Sharpening model on a schematic three layered neural network. The left panel shows the activity in the input layer caused by the first presentation of a visual stimulus. Blue graphs illustrate spiking of the neurons with highest response at each stage. According to this model the cause of the reduced BOLD signal for repeated presentation is that fewer neurons respond (central panel). The right panel shows the implication of the model (black line – initial response, grey line – repeated response). The Sharpening model predicts that repetition will narrow the tuning bandwidth, that is, sharpen the response around the preferred stimuli (Grill-Spector et al., 2006).

According to this model repetition results in sparser representation of stimuli (Figure 1.14). Specifically, neurons, that code features irrelevant to identification of a stimulus that exhibit RS, at the first occasion generate neuronal response to a stimulus, but show RS to subsequent presentation of that stimulus. So we can refer to this effect as a learning-process, where the distributed representation will be sparser or sharpened. The advantage of this interpretation is that it is more adaptive – in a sharper representation neurons may become more sensitive to changes, it makes perception more efficient and faster.

According to this model (Figure 1.15) the subsequent presentations of a given stimulus cause faster processing, that is, shorter latencies or shorter durations of neuronal firing.

Some authors directly tested which one of these aforementioned proposals is the best and reflects best the real behavioural and neuronal aspects of repetition effects. For example, James and Gauthier (2006) isolated the Fatigue (Suppression) and the Facilitation (Accumulation) Model from each other using a backward masking (BWM) paradigm. They hypothesized that the Fatigue Model would predict that effectively masked primed objects will produce less activation than non-primed objects, whereas the Facilitation Model would predict that effectively primed objects will produce more activation than non-primed objects (see also Figure 1.16). The authors have found that the Fatigue Model alone can not account for RS in fMRI, while the Facilitation Model could account for the findings.
After we summarized the behavioural and neuronal (haemodynamic) effect of repeated presentations we can observe another question from an economical point of view – what are the functional benefits of these effects and phenomena. The clearest example of a beneficial effect of adaptation is light adaptation in the retina. With the help of this phenomenon we are able to discriminate relatively small changes in luminance. We can refer to this effect as a horizontal shift in the contrast response function of the retinal ganglion cells and the cortical neurons, as well. Adaptation should thus lead observers to discriminate better between contrasts that are similar to the recently encountered mean. Adaptation re-centres tuning around prevailing stimulus conditions so as to improve discriminability. Second, adaptation improves the detectability or discriminability of novel or rare stimuli. It is an attractive idea, because it can be viewed as an extension of the general predictive coding strategy of the visual system, which improves efficiency by encoding the environment as differences in stimulus strength in space or time. Third, changes in tuning after adaptation serve to improve representational efficiency. The biggest benefit of this decreased redundancy is metabolic savings. When we define action potentials as a metabolically expensive neuronal signal then fewer active neurons and weaker responsiveness after adaptation may be of significant benefit.

Above we reviewed three proposals on the benefit of adaptation but it is worth noting that these explanations do not explain most adaptation effects.
The main questions of the thesis are the followings:

1. Which are the levels of visual processing that object-selective after-effects can be linked to?
2. Is there category specificity in adaptation?
3. Are there hemispheric asymmetries in the strength of the adaptation effect (both in the behavioural and the electrophysiological domains)?
4. What role does the duration of adaptation play in shape-selective after-effects?
5. Which level of face processing can be regarded as being reflected in consciousness?
2. How can our perception be biased when a stimulus appears *BEFORE* the target image? – adaptational results

2.1 Electrophysiological correlates of visual adaptation to faces and body-parts in humans

As it could have been seen in the Introduction the existence of adaptation to basic low-level visual dimensions – i.e. motion, orientation, spatial frequency or texture (for review see Anstis et al., 1988; Frisby, 1979; Durgin and Proffitt, 1996; Clifford, 2002) - has been known for a long time and the investigation of visual after-effects provided crucial information about the mechanisms involved in the processing of specific visual attributes. However, it has also been shown that adaptation is not an exclusive characteristic of the early stages of feature processing. Similarly to the after-effects caused by adaptation to low-level visual features prolonged exposure to a visual object - even if it is such a complex object as a face - will also lead to visual after-effects that are shape-selective and cannot be explained based on combination of adaptation to low-level features (Kohler and Wallach, 1944; Webster and MacLin, 1999; Leopold et al. 2001, Webster et al. 2004; Fang and He, 2005). In the topic of face adaptation related after-effects it has been found that these effects were to a large extent size (Zhao and Chubb, 2001, Rhodes et al., 2004; Anderson and Wilson, 2005) and position (Leopold et al., 2001) invariant as well as insensitive to the difference in the orientations of the adapting and test faces (e.g. if adaptor is rotated by +45 and the test is by −45 degrees, Rhodes et al., 2003). As I have already mentioned these findings altogether strongly suggest that neuronal processes at the higher stages of visual processing – containing neurons with large receptive fields and selective for complex shape information of the faces - might also adapt. Whether shape-selective adaptation is a general property of the mechanisms of visual object coding and whether its mechanisms are similar in case of different object categories are important questions waiting to be answered.

Neurophysiological research resulted in substantial progress towards uncovering the mechanisms of neuronal adaptation in recent years. It was found that in the early stages of
cortical processing, adaptation leads to desensitization of the neurons selective for the basic low-level dimensions of the adapting stimulus (Movshon and Lennie, 1979; Petersen et al., 1985; Carandini and Ferster, 1997; Dragoi et al., 2000). It has also been shown that adaptation can happen simultaneously at different stages/areas of cortical processing (for a review see Niedeggen and Wist, 1998) and that the mechanisms of neuronal adaptation in the case of short (a few hundred ms) and long (several seconds or more) adaptation might be different. For example, it appears that adaptation in motion direction selective neurons of macaque area MT takes place after short (< 500 ms; Priebe et al., 2002) but not after long (several seconds) exposure to their preferred motion direction (Kohn and Movshon, 2003). Thus, taken together, the physiological results point to the conclusion that neuronal adaptation to basic visual dimensions is a complex process, which might include different adaptation mechanisms acting simultaneously at different stages of visual processing.

We know surprisingly little, however, about the mechanisms of neuronal adaptation underlying shape-specific after-effects. To our knowledge there are no published results on the neuronal correlates of long-term adaptation (several seconds or more) to complex visual objects. This is because the related previous studies – including electrophysiological studies on animals (for review see Ringo, 1996), human neuroimaging (for review see Henson, 2003a) as well as ERP studies (Schweinberger et al., 1995, 2002; Campanella et al., 2002; Henson et al., 2003b; Werheid et al., 2005) – were testing how repeated short presentation (< 1 sec) of a specific image – rather than prolonged adaptation to it - affects the magnitude of neuronal responses to this stimulus in different visual cortical areas. The results of these studies provide converging evidence that – just as in the case of adaptation to low level visual attributes at the early visual areas – repeated exposure to the same object results in decreased neuronal responses in the cortical areas with neurons selective to complex shape properties of the visual objects.

However, although it has been shown that shape contrast - which is a negative after-effect for simple shape properties (i.e. convex or concave contours) - can be induced by short (< 1 sec) adaptation (Suzuki and Cavanagh, 1998), in all previous studies reporting shape-specific after-effects, a prolonged adaptation period was used, lasting for several second or even for minutes. In fact, Leopold et al. (Leopold et al., 2001) reported that short adaptation failed to evoke identity specific facial adaptation. These, taken together with the physiological results showing differential neuronal adaptation mechanisms after short and long adaptation in motion visual processing (Priebe et al., 2002; Kohn and Movshon, 2003) pose serious concerns regarding the possibility of making inferences about the neuronal mechanisms of
shape-specific adaptation based on the physiological results obtained with short adaptation using repetition suppression paradigm.

The goals of the present study were the following

- First, we aimed at investigating the ERP correlates of facial adaptation at the higher shape-selective stages of visual processing. ERP responses to faces – in addition to the early $P100$ component - are known to consist of a face-specific negative component (Bentin et al., 1996; Rossion et al., 1999; Itier and Taylor, 2004), the so-called $N170$ - peaking around 140-200 ms over the lateral occipito-temporo-parietal regions and its counterpart, the vertex positive peak ($VPP$, Jeffreys 1989; Rossion et al., 1999) over fronto-central sites. They are believed to reflect the structural encoding of facial features. Based on the known fact that facial after-effects are to a large extent invariant to changes of the low-level properties of the face stimuli (Leopold et al., 2001; Zhao and Chubb, 2001; Rhodes et al., 2003; Rhodes et al., 2004), we predicted that the neuronal effects of prolonged facial adaptation should be reflected primarily in the face-specific $N170$ and $VPP$ components of the ERP responses.

Previous studies testing the effect of repetition of shortly presented faces on $N170$ led to somewhat conflicting results: while in most of the studies $N170$ was found to be insensitive to repeated presentation of the same faces (Schweinberger et al., 1995, 2002; Henson et al., 2003; Werheid et al., 2005) there are studies showing that stimulus repetition decreased the amplitude (Campanella et al., 2002; Itier and Taylor, 2002) as well as the latency of $N170$ (Itier and Taylor, 2002).

- Second, to investigate whether high-level shape-selective adaptation is a general mechanism of visual object coding we tested whether prolonged adaptation to visual images of human hands can induce similar after-effects as adaptation to faces. The relevant physiological results suggest that human faces and body parts might be processed by different neural networks. fMRI studies have shown that viewing body parts activate a specific region of the lateral occipito-temporal cortex, the so called extrastriate body area (EBA, Downing et al., 2001) and importantly, it was also found that brain regions activated by body parts do not overlap with those that specifically respond to faces (Downing et al., 2001; Peelen and Downing, 2005). Moreover, it is known that the ERP evoked by human hand stimuli also consists of an $N170$ component (Mouchetant-Rostaing et al., 2000), providing us an opportunity to test whether neuronal adaptation effects in the case of faces and hands exhibit similar properties.
• Third, we also aimed at investigating the stimulus (category) specificity of the adaptation mechanisms at the higher stages of visual object processing. For this, we tested whether adaptation has any cross-category effects, by using different object categories during adaptation and testing; i.e. adapt with a face and the test stimulus is a hand or vice versa.

I will interpret our findings in separated blocks focusing on the responses to the above-mentioned experimental questions.

**Materials and Methods**

*Subjects:* 12 naïve, healthy volunteers (4 female and 8 male) participated in the study (age range: 16-27 years, mean 23 years). They all had normal or corrected-to normal vision, no previous history of any neurological or ophthalmologic diseases and were not under medication. The procedures were approved by the Ethical Committee of the Budapest University of Technology and Economics.

*Stimuli:* Face stimuli were grey-scale full-front digital images of six young males and six young females (chosen from a larger face database of our laboratory). Faces had no obvious gender-specific features, such as facial hair, jewellers, glasses or make-up. They were fit behind an oval mask (fit into a square of 400 x 400 pixels, 7.3 deg) eliminating the outer contours of the faces (see a sample image on Figure 2.1).

Hand stimuli were grey-scale digital images of six male and six female upright hands (see a sample image on Figure 2.1). They were matched in size and orientation. The luminance of the stimuli was subjectively equated (mean for faces is 1.17 and 2.1 cd/m$^2$ for hands). These pre-processed images were then set into pairs of male and female faces or hands and were entered into a morphing algorithm (WinMorph 3.01.), using landmark based morphing. 100 faces were created gradually along the female-male axis between each pairs of the specific face and hand stimuli. From each morphed series 6 images were chosen - corresponding to 80% female/20% male, 70%/30%, 60%/40%, 40%/60%, 30%/70% and 20% female/80% male images – and used as test images in the experiments.

We used three different stimuli as adaptors. A typical female face (FACE) was chosen as face adaptor (luminance 1.1 cd/m$^2$); a typical female hand was used as hand adaptor (HAND, luminance 1.8 cd/m$^2$). Finally, the Fourier phase randomized versions of the adaptors were used as controls (Control). These were created by an algorithm (Nasanen, 1999) that replaces the phase spectrum with random values (ranging from 0 to 360 deg), leaving the amplitude spectrum of the image intact, while removing any shape information.

Figure 2.1: Procedure and example stimuli.
Timeline depicts the Control, FACE and HAND adaptor images and the face and hand target
**Task, procedure:** Stimuli were presented centrally (on a 17” monitor, 1024 x 768 pixel resolution, 75 Hz vertical refresh rate; with a viewing distance of 70 cm) on a uniform grey background (luminance 1.3 cd/m²). Control, FACE and HAND conditions with face or hand test stimuli were given in separate blocks. Block order was randomized across subjects. All software was written in MATLAB 6.5. (Mathworks Inc.) using Psychtoolbox 2.45 for Windows (Pelli, 1997). Subjects were tested in a dimly lit room (average background luminance <1 cd/m²). They were instructed to fixate a spot in the center of the monitor and to perform a two-alternative forced choice gender discrimination task by pressing left mouse button when the face was perceived as a female and right button for male faces. Various stimulus values were presented according to a method of constant stimuli. In each trial (Figure 2.1), after a random interval (randomized between 500 and 700 ms), the adaptor was presented for 5000 ms, followed by a 200 ms gap, and finally the test face or hand stimulus, presented for 200 ms. Within a block each test stimulus was presented 6 times. Altogether ERPs were recorded for 70-80 minutes. During the experiments subjects were asked to refrain from movements and blinking.

**Electrophysiological recordings:** ERPs were recorded via 23 Ag/AgCl electrodes placed according to the 10/20 system with the left earlobe as ground and nasion as the reference lead. Impedances were kept below 5 kOhm. The sampling rate was 1024 Hz. EEG was segmented offline (using BrainVision Analyser (Brain Products GmbH, Munich, Germany)), into 1100 ms long trials, using a 100 ms prestimulus interval. Trials containing blinks, movements, A/D saturation or EEG baseline drift were rejected on the basis of visual inspection of each recording by semi-automatic artefact detection. ERPs were averaged separately for each subject, condition and channel. Averages were then digitally filtered (0.1-70 Hz) with a zero phase shift digital filter and average re-referencing was calculated.

**ERP analyses:** Analyses of ERP waves included the amplitude and latency of three major components: (1) P100, defined as a main positive deflection around 110 ms, (2) the N170 negative component and (3) the vertex positive peak (VPP), measured around 170 ms. For the P100 and N170 components, the parietal and occipital sites, corresponding to P7, P8, O1 and O2 electrode positions, were used, while VPP was measured at Cz. These electrode positions were selected on the basis of face-selective effects in previous studies (Eimer, 2000; Rossion et al., 2000). After averaging, the individual peak amplitudes were measured for each subject and each condition using a ± 30 ms window, centred on the maximum of the grand-average. The latencies of the ERP components were measured at the site of the peak amplitude. The magnitude of adaptation effects was determined by comparing the ERP responses found in the main adaptation conditions to those found in the control condition. Both amplitude and latency values were entered into a three-way repeated measures analysis of variance (ANOVA) with adaptor type (2, Control and FACE or HAND), hemisphere (2) and electrode (2) as within subject factors (except for VPP). To test if FACE and HAND had different adaptation effect on face and hand test stimuli we carried a four-way ANOVA with electrode (2), hemisphere (2), adaptation (2) and stimulus (2) out. All analyses involved Greenhouse-Geisser adjusted degrees of freedom for correction for non-sphericity. Post-hoc t-statistics were performed by Fisher LSD tests.
Results

**R1 / Facial adaptation and its electrophysiological correlates**

**Behavioural results**

Adaptation to a female face resulted in a strong perceptual aftereffect (Figure 2.2). In the FACE adapted conditions test faces were perceived more masculine as compared to the Control condition (F(1, 156)=16.67, P=0.00007), where a Fourier randomized female face image was used as an adaptor.

![Figure 2.2: Mean ratio of stimuli endorsed as female as a function of gender morphed level (% female). Face stimuli during Control, FACE and HAND adaptation. Data are modelled by Weibull psychometric function, using the psignifit toolbox (ver. 2.5.6.) for Matlab (http://bootstrap-software.org/psignifit/).](image1)

**ERP correlates of facial adaptation**

Face test stimuli evoked an ERP with clearly identifiable **P100** (measured in the time segment 90 - 120 ms) and **N170** components (measured in the time segment 140-200 ms) from O1, O2 and P7, P8 electrodes (Figure 2.3) as well as **VPP** from Cz electrode in all conditions.

![Figure 2.3: Grand average ERPs displayed between -250 and 550 ms at occipito-parietal (P7, P8) and medio-frontal (Cz) sites for the face (upper blocks) and for the hand stimuli (lower blocks) in the Control (thick line), FACE (thin line) and HAND (dashed line) adapted conditions.](image2)
**Amplitude**

*N170* showed a significant main effect of adaptation (F(1, 12)=24.1, *P*=0.0004) due to the *N170* amplitudes in the Control conditions being larger than those in the FACE adapted condition for each electrode and hemisphere on the post hoc contrasts (*P*<0.05 for all comparisons, Figure 2.4a). No overall effect of hemisphere was found for *N170* amplitude, while it was smaller for occipital than parietal electrodes (F(1, 12)=4.55, *P*<0.05). The adaptation effects on *N170* amplitude were significantly larger on the right side (adaptation by hemisphere interaction: F(1, 12)=5.7, *P*=0.03). Although there was no significant interaction of adaptation by electrode and hemisphere, post hoc tests revealed that the adaptation effects are larger on the parietal (*P*<0.000007 for both hemispheres) than on the occipital electrodes (*P*<0.003).

![Figure 2.4: ERP correlates of adaptation for face stimuli. Mean N170 (recorded from O1, O2, P7 and P8) and VPP (recorded from Cz; a) and P100 (b) amplitudes and N170, VPP (c) and P100 latencies (d). Control adapted condition: black, FACE adapted condition: hatched.](image)

Facial adaptation also decreased significantly the amplitude of *VPP* (F(1, 10)=7.5, *P*=0.02), measured at Cz. *P100* showed a significant main effect of adaptation (F(1, 11)=11.115, *P*=0.007) as well, due to the *P100* amplitudes in the FACE adapted conditions being larger than in the Control conditions (Figure 2.4b). This effect was larger over the right hemisphere (significant interaction of hemisphere by adaptation: F(1, 11)=5.57, *P*=0.038).
Latency

Facial adaptation caused a significant delay of the $N170$ component, reflected in the main effect of adaptation (Figure 2.4c; $F (1, 11) = 8.3$, $P = 0.015$. However, although significant for both parietal and occipital electrodes, this effect was larger on the parietal sites, reflected in the significant interaction of electrode by adaptation ($F (1, 11) = 6.3$, $P = 0.028$). The latency of $VPP$ was also significantly longer in the FACE adapted relative to the Control condition ($F (1, 11) = 6.0$, $P = 0.03$). $P100$ latencies were longer in the FACE adapted than in the Control condition (Figure 2.4d; $F (1, 11) = 11.0$, $P = 0.007$) as well.

R2 / The generality of high-level shape-selective adaptation

Behavioural results

It was also tested whether adaptation to a human female hand could bias the perceived gender of the subsequently presented test hand stimulus. As it can be seen on Figure 2.5, adaptation to a female hand resulted in a strong perceptual aftereffect: the test hand stimuli were perceived as more masculine, as compared to the Control condition ($F (1, 156) = 16.97$, $P = 0.00006$).

![Figure 2.5: Mean ratio of stimuli endorsed as female as a function of gender morphed level (% female). Hand stimuli during Control, FACE and HAND adaptation. Data are modelled by Weibull psychometric function, using the psignifit toolbox (ver. 2.5.6.) for Matlab (http://bootstrap-software.org/psignifit/).](image)

We also investigated whether the hand after-effects we found in our main experiment are invariant to changes in the size or orientation of the test hand stimuli compared to that of the adaptor hand image. For this we slightly modified our hand adaptation paradigm. In the conditions used to test for size invariance, within the same block 50% of the test hand images matched the size of the adaptor image whereas the area of the other half of the test images was 2.5 times larger than the adaptor hand. In the condition designed to test the orientation specificity of the adaptation effects, observers were tested with blocks, containing test hands
that matched the orientation of the adaptor (upright) and test hands that were rotated by 90 deg in the frontal plane. The adaptor image in both conditions – testing for size as well as orientation invariance – was a male hand (as opposed to the main experiment, where female hand was used as an adaptor). Our results clearly show that the adaptation effect was significant in both conditions (main effects of adaptation for size and rotation: F (1, 368) =55.9, p=.00001, F (1, 304) =13.3, p=.0003, respectively, Figure 2.6), meaning that test hands were judged more feminine as a result of adaptation to a male hand.

Moreover, the adaptation effect generalized across both size and orientation changes (as suggested by the non-significant interaction of adaptation with size and rotation: F (1, 368) =0.4, p=.54 and F (1, 304) =.97, p=.33 for adaptor-size and adaptor-rotation, respectively, Figure 2.6). Interestingly, size had a main effect on the observer’s gender judgments (F (1, 368) =353.1, p=0.0001) but this effect was independent of the adaptation: i.e. it was present both in the hand-adapted as well as in the Control conditions. Namely, the same hand images were judged more masculine when their size was increased, independently of presence or absence of adaptation.

Figure 2.6: Mean ratio of hand stimuli endorsed as female as a function of gender morph level (% female). Upper row: size invariance experiment. Original size hand stimuli (a) and 2.5 times enlarged size hand stimuli (b). Lower row: orientation invariance experiment. Original upright hand stimuli (c) and 90 deg rotated hand stimuli (d) during Control and male HAND adaptation. Data are modeled by a Weibull psychometric function.
**ERP correlates of adaptation to human hands**

**Control condition**

Similarly to faces, human hand test stimuli evoked an ERP with clearly identifiable *P100* and *N170* components from O1, O2 and P7, P8 electrodes (Figure 2.3) as well as *VPP* from Cz electrode in all conditions. Comparison of the *N170* components to hands and faces revealed that in the Control condition face test stimuli evoked a *N170* with significantly larger amplitude than hand test stimuli on P8 (post-hoc, P=0.00002) and significantly smaller *N170* on occipital sites (Figure 2.3; P=0.000003 and P<0.05 for left and right, respectively). The amplitude of *VPP* was not significantly different for faces and hands (F (1, 12) =2.9, P=0.11). When the *P100* component is considered, hand test stimuli evoked significantly larger amplitudes than face stimuli in the Control condition (F (1, 11) =7.38, P=0.02). The latency of the *P100*, *N170* and *VPP* components were not different for faces and hands in the Control condition.

**Amplitude**

*N170* showed a significant main effect of adaptation for hand test stimuli (Figure 2.8a; F(1, 10)=18.4, P=0.0016) due to the *N170* amplitudes being larger in the Control conditions than those in the HAND adapted condition for each electrode and hemisphere on the post hoc contrasts (P<0.01 for all comparisons).

![Figure 2.8: ERP correlates of adaptation for hand stimuli. Mean N170 (recorded from O1, O2, P7 and P8) and VPP (recorded from Cz; a) and P100 (b) amplitudes and N170, VPP (c) and P100 latencies (d). Control adapted condition: black, HAND adapted condition: hatched.](image-url)
Just like in the case of FACE adapted condition, the effect of adaptation on $N170$ was significantly larger on the right side for HAND as well (adaptation by hemisphere interaction: $F(1, 10)=5.15, P=0.047$). Adaptation and electrode interaction also reached significance, suggesting that adaptation effects in the HAND condition are larger on the occipital electrodes than on parietal ones ($F(1, 10)=5.7, P=0.04$). Adaptation to human hands also decreased significantly the amplitude of $VPP$ obtained for the hand test stimuli ($F(1, 10)=34.1, P=0.0002$), measured at Cz. The amplitude of the $P100$ component evoked by the hand test stimuli was not different in the HAND adapted condition from that measured in the Control condition ($F(1, 11)=2.97, P=0.112$).

**Latency**

Adaptation to a human hand caused a significant delay of the $N170$ component evoked by the subsequently presented hand test stimulus (Figure 2.8c), reflected in the main effect of adaptation: $F(1, 11)=7.1, P=0.022$). Adaptation effect for HAND was larger on the parietal sites, reflected in the significant interaction of electrode by adaptation ($F(1, 11)=8.67, P=0.01$). The latency of $VPP$ was also significantly longer in the HAND adapted than in the Control condition ($F(1, 11)=5.89, P=0.03$). $P100$ latencies were longer in the HAND adapted condition than in the Control as well (Figure 2.8d, $F(1, 11)=12.11, P=0.0055$) and this increase in latency was larger for parietal than for occipital electrodes (electrode by adaptation interaction effect ($F(1, 11)=7.5, P=0.02$)).

**R3 / The category-specificity of the adaptation mechanisms**

**Behavioural results**

The effects of adaptation to faces and/or to human hands appeared highly category-specific, since adaptation to faces and/or human hands did not affect the gender judgments of human hand and/or human face test stimuli, respectively (main effect of adaptation: $F(1, 156)=0.06, P=0.8$ and $F(1, 156)=2.1, P=0.15$, respectively). Thus, the main findings of our behavioural experiments can be summarized as follows. Adaptation to both faces and human hands results in strong category-specific perceptual after-effects. The judgments about the gender of the face and hand test stimuli were biased towards more masculine following adaptation to female faces or hands, respectively. We also have shown that the newly described hand after-effects are robust to changes in hand size or orientation, suggesting that these after-effects arise from the adaptation of the higher-level shape-specific mechanisms of object.
ERP correlates of cross-category adaptation effects

Amplitude

The ERP results are in accordance with our behavioural findings, showing no cross-category adaptation effects. Adaptation with a human hand did not affect the ERP responses to the test face stimulus. \(N170\), \(VPP\) (Figure 2.9a) and \(P100\) amplitudes (Figure 2.9b), obtained in the conditions of HAND adapted face test stimuli were not different from those obtained in the Control condition with face test stimuli (\(F(1, 11)=2.27, P=0.159\) for \(N170\), \(F(1, 11)=0.86, P=0.37\) for \(VPP\) and \(F(1, 11)=0.48, P=0.504\) for \(P100\)).

Similarly, adaptation to FACE did not affect the \(N170\) and \(VPP\) amplitudes evoked by the test hand stimuli (Figure 2.9c; \(F(1, 11)=2.24, P=0.16\) for \(N170\) and \(F(1, 11)=2.99, P=0.011\) for \(VPP\)). There was only one significant cross-category effect in the FACE adaptor and hand test condition: the amplitude of \(P100\) was significantly larger than in the Control condition (Figure 2.9d; \(F(1, 11)=13.3, P=0.004\)).
Latency

N170, VPP and P100 latencies were not different for HAND adapted face and FACE adapted hand conditions when compared to CONTROL adapted conditions (Figure 2.10a-d; F (1, 11)>0.3, P>0.3 for all conditions).

![Figure 2.10: Cross-category adaptation for face and hand stimuli: ERP latencies. Mean N170 / VPP (a) and P100 latencies (b) obtained for face test stimuli adapted by Control and HAND adaptors and N170/VPP (c) and P100 (d) latencies for hand test stimuli adapted by Control and FACE adaptors. Control adapted condition: black, FACE or HAND adapted condition: hatched.](image-url)
Summary of the results

In the mirror of our experimental questions we can summarize our findings as follows:

- We could replicate the behavioural results of Webster et al. (Webster et al., 2004) as adaptation to a female face has a strong effect on our perception – a subsequent androgynous face appears more male. Moreover, we also found significant effects on the early components of ERP responses. After adaptation the \( P100 \) component was larger and delayed, whereas the \( N170 \) was dramatically reduced and delayed.

- Similar effects can be found when we use human hands as stimuli. These effects mirrored not only at the behavioural but also at the neuronal level.

- Hand adaptation effects were to a large extent size and orientation invariant, implying that these effects are rather higher-level effects and can not be described as a combination of local, low-level effects.

- Our main finding is the category-specificity of adaptation both at the behavioural and at the neuronal level since cross-category adaptation has no effect on the psychometric functions or on the amplitude and/or latency of \( P100, N170 \) or \( VPP \).
2.2 Position-specificity of facial adaptation

In our previous experiment we suggested that hand adaptation is a higher-level effect and it is not described as a combination of local adaptation to low-level features. Therefore we can hypothesize that effects reflect adaptation at the higher stages of visual object processing, most likely in the regions or subregions of IT cortex (Leopold et al., 2001; Webster et al., 2004), where neuronal processes are not retinotopically organized (Schwartz et al., 1983). In accordance with this, it has been shown that facial after-effects are to a large extent size invariant (Zhao and Chubb, 2001; Rhodes et al., 2004, Kovács et al., 2006) and also generalize across conditions where the retinal position of the adaptor and test stimuli differs, suggesting that they can be results of higher-level effects (Rhodes et al., 2003, Kovács et al., 2006).

However, recent single-unit recording studies have provided evidence that some positional information is preserved in the inferior temporal cortex of the macaque brain (for a review see Rousselet et al., 2004). It was found that the receptive fields of inferior temporal neurons can differ in size, have “hot spots”, where they are most sensitive to stimulation, and are typically biased towards the contralateral hemifield (DiCarlo and Maunsell, 2003; Rolls et al., 2003; OpDeBeeck and Vogels, 2000). Therefore, if adaptation of these neurons is indeed involved in shape-selective after-effects, they should show at least partial position-selectivity.

We investigated the position invariance of the behavioural as well as the event related potential (ERP) effects of adaptation to upright and upside-down faces. We compared both the magnitude of facial after-effects and the adaptation effects on the N170 component of the ERP responses – the primary correlate of prolonged face adaptation (Kovács et al., 2006) - when the adaptor and test face stimuli were presented on the same retinal location (overlapping conditions; OL) to that when they were presented in different hemifields (non-overlapping conditions; non-OL).
Materials and Methods

**Subjects:** Fifteen naïve, healthy volunteers (8 females) with normal or corrected-to-normal vision participated in the study (mean age: 23 ± 3 years). The procedures were approved by the Ethical Committee of the Budapest University of Technology and Economics.

**Stimuli:** Face stimuli were identical to those used in the previous experiment. A simultaneously presented face and a Fourier image (Nasanen, 1999; Kovács et al., 2006), displayed peripherally 5 deg (closest side) to the left and to the right from the fixation point along the horizontal meridian (Figure 2.11) was used as test stimulus. There were three different adapting conditions. In the Control condition, the adaptor stimulus was composed of two Fourier images presented on each side of the fixation point. In the two other conditions one female face and one Fourier image was presented during adaptation and a test face and a Fourier image during testing. In the OL condition the position of the adaptor and test face was the same, whereas in the Non-OL condition, their location was different.

In all three conditions, within a block (consisting 120 trials), the different types of adaptor and test stimuli were displayed randomly either in the left or in the right visual field. Upright and upside-down stimuli as well as different adapting conditions were presented in separate blocks.

**Task and EEG recording procedures:** In each trial, (after an interval randomized between 500 and 700 ms), the adaptor was presented for 5 s, followed by a 500 ms gap, and finally the test stimulus, presented for 200 ms. Subjects were instructed to fixate a dot in the centre of the monitor and to perform a two-alternative forced choice gender discrimination task. The eye-movements of the subjects were controlled by an infrared eye-tracking system (IView X RED, SMI, Germany). ERPs were recorded via 25 Ag/AgCl electrodes placed according to the 10/20 system (ground: left earlobe, reference: nasion; impedance < 5 kOhm). EEG was segmented offline (BrainVision Analyser (Brain Products GmbH)), into 1100 ms long trials, using a 100 ms prestimulus interval.

Trials containing blinks, movements, A/D saturation or EEG baseline drift were rejected. ERPs were averaged separately for each subject, condition and channel. Averages were then digitally filtered (0.1-70 Hz) with a zero phase shift digital filter and average re-referencing was calculated.

**Behavioural data analysis:** A three-way analysis of variance (ANOVA) was performed on the subjects’ performance with the visual field of the adaptor (left vs. right), the spatial relation of the target and adaptor faces (OL vs. non-OL) and ratio of sexual dimorphism, i.e. morph level (10) as within subject factors.

**ERP data analysis:** The amplitude and latency of the N170 was measured at eight occipito-temporal electrode positions (P7/P8, P9/P10, PO7/PO8 and TP7/TP8) (Eimer, 2000; Rossion et al., 2000). Individual peak amplitudes were measured for each subject and each condition using a ± 30 ms window, centred on the maximum of the grand-average. Amplitude values were entered into a three-way repeated measures ANOVA with adaptor (3; Control, OL, non-OL), recording hemisphere (left vs. right), and target face visual field (left vs. right) as within subject factors. Post-hoc t-statistics were performed by Fisher LSD tests.
Results

**Behavioural results**

Adaptation to an upright (Figure 2.12a) or upside-down (Figure 2.12b) female face resulted in a strong perceptual after-effect - test faces were perceived more masculine as compared to the Control – both, when the adapter and test faces were presented in the same retinal position (OL condition, main effect of adaptation for upright and upside-down faces: $F(1, 800)=275$, $P<0.00001$) and $F(1, 250)=126$, $P<0.00001$, respectively) as well as when they were presented in the opposite hemifields (non-OL condition, $F(1, 800)=121$, $P=0.00001$ for upright and $F(1, 250)=56$, $P<0.00001$ for upside-down faces).

However, the magnitude of the adaptation was significantly larger when the adaptor and test faces spatially overlapped as compared to when their location did not overlap (upright faces: $F(1, 520) =22.493$, $p=.000001$, upside-down faces: $F(1, 200) =13.7$, $p=.0003$), suggesting that the facial after-effects consist of a position invariant and non-invariant component.
ERP results

Similarly to our previous observations with centrally presented faces (Kovács et al., 2006) in the present study adaptation with a face stimulus had a significant main effect on the amplitude of the N170 component evoked by the test stimuli (upright faces (Figure 2.13a and b): F(2, 24)=12.4, P=0.0002; upside-down faces (Figure 2.14a and b): F(2, 22)=37.9, P=0.000001).
The N170 amplitudes in the Control condition were significantly larger than both, in the OL and in the non-OL conditions (P<0.05 for all comparisons). We also found that the overall adaptation effects on N170 amplitude measured over the left and right hemispheres did not differ in the case when adaptor and test stimuli spatially overlapped from that when they were presented in different hemifields (post hoc test for OL vs. non-OL adaptor conditions: upright: p<0.7; upside-down: 0.3). When the amplitude of the N170 was analysed separately for the two hemispheres, we found that the adaptation effects over the left hemisphere – but not over the right - were significantly larger in the OL than in the non-OL conditions (p<0.03 and p<0.05 for left and p<0.5 and p<0.8 for right hemisphere and upright and upside-down, respectively). There were no significant differences in N170 latency between Control, OL and non-OL (F(2, 48)=3, P=0.1 and F(2, 48)=0.9, P=0.4 for the main effect of adaptation of upright and upside-down faces, respectively).

We also tested whether the adaptation effects differ in the OL condition from the non-OL condition when the N170 components evoked by ipsi- and contralateral test stimuli are analysed separately. When the test stimuli were presented contralaterally, N170 adaptation effects were significantly larger both over the left (upright: p<0.004; upside-down: p<0.0001) and right hemisphere (upright: p<0.05; upside-down: p<0.005) in the OL condition than in the non-OL condition. However, in the case of ipsilateral test stimuli, N170 amplitudes did not differ in the OL and non-OL conditions over the left hemisphere (upright: p<0.23; upside-down:0.18), whereas over the right hemisphere, adaptation effects on N170 amplitudes were significantly larger in the non-OL than in the OL condition (upright: p<0.04; upside-down:0.0002).
Summary of the results

The results of the present experiment can be summarized as follows:

- Facial after-effects evoked by adaptation to both upright and inverted faces are significantly larger when the adaptor and test faces are presented on the same retinal position, thus we can say that this effect is only partially translation invariant.

- Adaptation effects on the N170 component measured over the hemisphere that was contralateral to the test face stimulus also show strong position-specificity.
2.3 Adaptation duration affects the spatial selectivity of facial after-effects

One of the important parameters affecting the site of adaptation is its timescale, i.e. the exposure time of the first (adaptor) stimulus (Krekelberg et al., 2006). In psychophysical experiments of mid-level adaptation it has been shown that shape contrast - which is a negative aftereffect for simple shape properties (i.e. convex or concave contours) - can also be induced by short (less than 1 sec) adaptation (Suzuki and Cavanagh, 1998). All other previous studies reporting shape-specific aftereffects used a prolonged adaptation period, lasting for several second or even for minutes. In fact, Leopold et al. (Leopold et al., 2001) reported that short adaptation failed to evoke identity specific facial adaptation. Using fMRI it was found that adaptation to basic visual features (such as orientation) at the early, retinotopically organized visual cortical areas – in particular in the primary visual cortex – can be revealed only after long-term exposure (several seconds) to the adapting stimulus (Fang et al., 2005). This is in sharp contrast to what was found in the later, more downstream visual cortical areas (i.e. V4, LO, IT) where not only long-term, but already very short adaptation (less than 1 sec, the so called fMRI adaptation procedure) to specific features or objects can evoke strong fMRI adaptation effects (Henson, 2003). These results suggest that different neuronal populations along the visual hierarchy might differ in their dynamics of adaptation. If this is so, it raises the intriguing possibility that by manipulating adaptation duration one can selectively adapt neuronal processes at different stages of visual processing and investigate their stimulus selectivity.

In the present study we investigated short- and long-term adaptation effects at the higher, shape-specific stages of visual processing using face adaptation (Leopold et al., 2001; Rhodes et al., 2003; Yamashita et al., 2005; Webster and MacLin, 1999; Webster et al., 2004). Recently, Leopold, Rhodes, Müller and Jeffrey (Leopold et al., 2005) has shown that varying
the duration of the adapter affects face adaptation; the strength of face identity after-effect increases as a function of adaptation time. Earlier fMRI (Andrews and Ewbank, 2004; Henson, 2003), ERP (Henson, 2003a; Schweinberger et al., 2002; Trenner et al., 2004 ) and MEG (Boehm et al., 2005; Harris and Nakayama, 2006) studies used short-term adaptation to investigate face specific processing and found strong adaptation effects. However, the neuroimaging results concerning the effect of adaptation duration on the strength of adaptation are less conclusive and suggest that it might depend on the type of the stimuli used as well as on the testing protocol. For examples, Sayres and Grill-Spector (Sayres and Grill-Spector, 2006) found, that the magnitude of repetition related BOLD response suppression increases gradually with the number of repetitions, whereas Zago, Fenske, Aminoff and Bar (Zago et al., 2005) found, that BOLD response reduction is maximal for relatively short (40-250 ms) exposition times, while longer (up to 1900 ms) presentations lead to the reduction of the effect. Importantly, Fang, Murray and He (Fang et al., 2006) have shown that adaptation duration can also affect the properties of fMRI adaptation effects: while long term (5000 ms) adaptation effects were found to be selective to the viewpoint of the face, short term (300 ms) adaptation effects were viewpoint independent along the ventral pathway, including the lateral occipital cortex and right fusiform area.

Here we used event related potentials (ERP) to investigate the effect of adaptation duration on the strength and properties of facial adaptation effects, as well as on their electrophysiological correlates. In a recent study, we have shown that long-term face adaptation has a strong effect on the ERP responses and that it is the N170 component that represents the primary ERP correlate of long-term face adaptation (Kovács et al., 2006). Moreover, we also have shown that both the behavioural face adaptation effects as well as their ERP correlates consist of position-specific and position-invariant components (Kovács et al., 2005), reflecting adaptation processes of neuronal populations with different positional sensitivity. The goal of the present study was to test whether adaptation duration differentially modulates the position-specific and position-invariant components of the behavioural face after-effects as well as the ERP correlates of face adaptation. Based on the fMRI results suggesting that adaptation of neuronal processes at the earlier stages of visual cortical processing need longer adaptation period than that at the higher stages of visual processing we predicted that: position-specific ERP face adaptation effects can be evoked only by long-term (5 sec) adaptation, whereas the position-invariant component will be present both after short- (500 ms) and long-term adaptation.
Materials and Methods

Subjects: 21 naïve, healthy volunteers (9 females) with normal or corrected-to-normal vision participated in the experiments (mean age: 22 ± 2 years). The procedures were approved by the Ethical Committee of the Budapest University of Technology and Economics.

Stimuli and procedures: The stimuli, the task, the long-term adaptation condition and the electrophysiological recording procedures were identical to those used and described in details in recent studies by Kovács et al. (Kovács et al., 2005, 2006) thus here just a brief description of these and additional details are given.

For the short-term adaptation condition we manipulated the shape information content of the image by randomising the Fourier phase of the images in the following ranges: 315, 270, 225, 180, 135, 90 and 45 deg (Nasanen, 1999). This results in eight images (see also in Figure 2.15), where the 0 deg randomised image is the original face, and the images with increasing randomisation-range contain gradually less and less of the original shape information. In the short-term adaptation condition 360 deg Fourier phase randomized adaptors were presented for 4500 ms, after which one of the images (chosen randomly either on the left or on the right side) was replaced consecutively by a 315, a 270, a 225, a 180, a 135, a 90, a 45 deg randomised images and finally by a face image which phase was not randomised (each image appeared for 100 ms).

This led to a percept of a face gradually emerging from a noise field. Pilot studies showed, that gender discrimination is at chance level for images above 180 deg phase randomisation, meaning a total time of gender-specific facial adaptation of approximately 500 ms.

In all three conditions (Control, short- and long-term adaptation), adaptors were followed by a 500 ms gap, and then by the test stimulus, presented for 200 ms (Figure 2.16). Subjects fixated a dot in the centre while their eye-movements were controlled by an infrared eye-tracking system (IView X RED, SMI, Germany). In the further analysis we only included those trials where the fixation stayed within a 2 deg radius circle of the fixation spot (the mean ratio of valid trials across our subjects was: 87.4 ± 5 %, 89.1± 5.4 and 87.9± 5.3 % for the Control, short and long-term adaptation conditions (mean±SE)). The various morph-levels were presented according to a method of constant stimuli. Subjects’ task was identical in...
all conditions: they had to perform a two-alternative forced choice gender discrimination task. We measured the ratio of female and male responses as dependent variable. Control, short-term and long-term adaptation trials were given in separate blocks (the order randomised across subjects). Within a block (consisting 120 trials), the different types of adaptors (SAME and DIFFERENT) and the hemifield of the test stimuli (left, right) were presented randomly.

Data analysis: Psychophysical data were modelled by a Weibull psychometric function, using the psignifit toolbox (ver. 2.5.6) for Matlab (http://bootstrap-software.org/psignifit/). First, to test whether adaptation to faces or to Fourier phase randomised images of faces results in differential effects on subsequent facial gender discrimination a three-way ANOVA was performed on the subjects’ female-male decisions with the adaptation condition (Control, SAME vs. DIFFERENT; 3), male-female pairs (3) and morph level (10) as within subject factors, separately for the short and long term adaptation blocks. Next, to test directly the effect of adaptation duration a four-way ANOVA was performed with adaptation duration (short, long; 2), position of the adaptor and test stimuli (SAME, DIFFERENT, 2), male-female pairs (3) and morph levels (10) as factors. To determine the onset of adaptation effects we made a point by point paired t-test on the ERP curves for each hemisphere (p<0.05 for at least 30 consecutive points), across subjects and electrodes. The amplitude and latency of the N170 was measured at six occipito-temporal electrode positions (P7/P8, PO7/PO8 and PO9/PO10) (Eimer, 2000, Rossion et al., 2000), while P100 was measured on O1 and O2 electrodes. In order to determine adaptation effects, we used adaptation condition (3; Control, SAME, DIFFERENT) as factors in a one-way ANOVA separately for each hemispheres and for short and long term adaptation conditions. Post-hoc t-statistics were performed by Fisher LSD tests. Finally we correlated the behaviourally and electrophysiologically observed adaptation effects. First, we determined the gender category border at 50% chance performance for each subject and each adaptation condition separately. Second we calculated the magnitude of adaptation by subtracting the stimulus morph level at the perceived category border in each adapted condition (SAME, DIFFERENT after long and short adaptations) from that of the Control condition. The electrophysiological adaptation effect was calculated by subtracting the N170 and P100 amplitude obtained during the adapted conditions from that of the Control condition. These adaptation effects were then correlated to each other by Pearsons’ tests.

Results

Behavioural results

After being adapted to a female face for 5 s the test faces were perceived more masculine as compared to the Control (Figure 2.17; main effect of adaptation condition: F(2, 1650)=135, P<0.00001). However, the magnitude of the adaptation, reflected in the shift of the psychophysical curves along the X axis, was significantly larger when the adaptor and the test faces were presented on the same side of the fixation as compared to when they were presented in the opposite hemifields (p=0.00003, post-hoc t-test). This is in accordance with our previous results (Kovács et al, 2005) and suggests that facial after-effects consist of a position invariant and non-invariant component after long adaptation durations.
Short-term adaptation also evoked significant after-effects (main effect of adaptation condition: $F(2, 1650)=29, P<0.009$).

However, in contrast to the long-term adaptation conditions in the short-term adaptation conditions the size of the adaptation effect did not differ significantly between the SAME and DIFFERENT conditions ($p=0.11$, post-hoc t-test). This suggests that the facial after-effects evoked by short-term adaptation are entirely position invariant. Adaptation duration had a significant effect on the strength of facial after-effects: comparison of the long term and short term adaptation conditions revealed that the test faces were judged more masculine after long term adaptation than after short term adaptation condition (main effect of adaptation duration ($F(1,1650)=87, P<0.00001$).

The difference in the strength of the facial after-effects between long term and short term adaptation was more pronounced in the SAME as compared to the DIFFERENT condition (interaction of adaptation condition and adaptation time: $F(1,1650)=4, P<0.05$).
**ERP results**

**ERP responses over the hemisphere contralateral to the test stimuli**

Figure 2.18: Grand average ERPs recorded from contralateral target stimuli displayed between -100 and 500 ms at occipito-temporal sites for the Control (thick line), DIFFERENT (thin line) and SAME (dashed line) conditions. Upper panel: 5000 ms adaptation time, lower panel: 500 ms adaptation time. Pooled channels data: PO7, P7, P9 – left hemisphere (LH), PO8, P8, P10 – right hemisphere (RH) registration.

Figure 2.18 shows the pooled grand average ERPs obtained on the occipito-temporal electrodes. We found strong adaptation effects for both short and long-term adaptation conditions (Figure 2.19a).
The onset of the adaptation effects - as measured by a point-by-point t-test on the grand average ERPs obtained in Control and the long term adaptation conditions - was very similar for each condition: 117, 119 and 109, 106 ms for the long term adapted SAME and DIFFERENT conditions over the right and left hemispheres, respectively.

The amplitude of the $N170$ component was smaller both in the long term and in the short term adaptation conditions than in the control condition for both hemispheres (main effect of adaptation condition: $F(2,120)>14; p<0.0001$).

In the case of short-term adaptation condition the onset of the adaptation effects measured on the $N170$ component was at 119, 117 and 111, 114 ms for the SAME and DIFFERENT conditions over the right and left hemispheres, respectively.

When comparing the magnitude of $N170$ adaptation effects in the SAME and DIFFERENT conditions after long-term adaptation we found significantly lower amplitudes in the SAME than in the DIFFERENT conditions (post-hoc t-tests: $p<0.00001$ and $p<0.005$ for the right and left hemispheres, respectively). However, in the case of short-term adaptation conditions, there was no such difference in $N170$ amplitudes between the SAME and DIFFERENT adaptor-target arrangements, in neither hemispheres (post-hoc t-tests: $p<0.08$ and $p<0.1$ for the right and left hemispheres, respectively).

The observed behavioural adaptation effect correlated significantly only with the adaptation effect on the $N170$ amplitude after long term adaptation ($r=0.36, p<0.05$).

The amplitude of the $P100$ component was larger after long-term adaptation over the left hemisphere (Figure 2.19b, main effect of adaptation condition: $F(2,38)=9; p<0.001$). Although short term adaptation caused some increase of the amplitude of the left $P100$ components as well, but this effect was not significant ($F(2,38)=2.5, p=0.09$).
More importantly, we have not observed significant differences in the adaptation effects on $P100$ amplitudes between the SAME and DIFFERENT conditions neither for the long-term nor for the short-term adaptation conditions over neither hemispheres (post-hoc t-tests: $p>0.2$ for all comparisons).

These results suggest, that although long-term adaptation produces some moderate adaptation effects on the $P100$ component of the ERP responses (in accordance with our previous results (Kovács et al, 2005) this effect is similar for position-specific and position-invariant situations, thus can not explain the observed differences in the $N170$ ERP component.

The fact, that the $N170$ and $P100$ adaptation effects reflect different adaptation mechanisms is supported further by the fact, that the $P100$ and $N170$ amplitudes showed no significant correlations neither on the left ($r=0.04$, nor on the right side ($r=-0.06$).

Regarding the $N170$ latencies (Table I.) in the long adapted condition the main effect of adaptation condition was significant over the right hemisphere ($F(2,120)=9$, $p<0.0002$) while no differences were found for the left hemisphere ($F(2,120)=0.6$, $p=0.55$).

After short-term adaptation the latencies were also significantly larger than in the Control condition (main effect of adaptation condition: $F(2,120)=36$, $p<0.0001$ and $F(2,120)=23$, $p<0.0001$ for right and left hemispheres, respectively).

<table>
<thead>
<tr>
<th>$N170$</th>
<th>Contralateral LH</th>
<th>Contralateral RH</th>
<th>Ipsilateral LH</th>
<th>Ipsilateral RH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>168 (2.3)</td>
<td>170 (1.9)</td>
<td>173 (1.7)</td>
<td>177 (1.6)</td>
</tr>
<tr>
<td>Long DIFFERENT</td>
<td>165 (1.9)</td>
<td>168 (1.3)</td>
<td>164 (1.7)*</td>
<td>170 (1.2)*</td>
</tr>
<tr>
<td>Long SAME</td>
<td>167 (2.1)</td>
<td>175 (2.2)*</td>
<td>165 (1.7)*</td>
<td>176 (1.6)</td>
</tr>
<tr>
<td>Short DIFFERENT</td>
<td>176 (2.4)*</td>
<td>186(2.2)*</td>
<td>176 (2.4)</td>
<td>188 (2.3)*</td>
</tr>
<tr>
<td>Short SAME</td>
<td>179 (2.2)*</td>
<td>183 (2.4)*</td>
<td>177 (2.0)</td>
<td>184 (2.5)*</td>
</tr>
</tbody>
</table>

Asterisks sign significant differences from the appropriate control condition by Fishers post-hoc tests ($p<0.01$).
Table II.
The latency values (mean and standard error) of \textit{P100} component for the left and right hemispheres in the different conditions

<table>
<thead>
<tr>
<th>P100</th>
<th>Contralateral</th>
<th>Ipsilateral</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>O2</td>
<td>O1</td>
</tr>
<tr>
<td>Control</td>
<td>114 (3.8)</td>
<td>113 (1.9)</td>
</tr>
<tr>
<td>Long DIFFERENT</td>
<td>118 (3.2)*</td>
<td>119 (2.4)*</td>
</tr>
<tr>
<td>Long SAME</td>
<td>121 (3.8)*</td>
<td>121 (2.1)*</td>
</tr>
<tr>
<td>Short DIFFERENT</td>
<td>123 (3.7)*</td>
<td>120(2.4)*</td>
</tr>
<tr>
<td>Short SAME</td>
<td>124 (3.7)*</td>
<td>120 (3.7)*</td>
</tr>
</tbody>
</table>

Asterisks sign significant differences from the appropriate control condition by Fishers post-hoc test (p<0.01).

The latency of the \textit{P100} component (Table II.) was significantly longer after both short and long term adaptation (main effects of adaptation condition F(2,38)>4.2 p<0.01 for all four comparisons).

As for SAME and DIFFERENT conditions the \textit{N170} latencies were longer in the SAME condition than in the DIFFERENT conditions (post-hoc t-tests: p<0.00004 for left and right hemispheres) after long-term adaptation, while no such differences were observed after short term adaptation (post-hoc t-tests: p=0.6 and p=0.1 for right and left hemispheres, respectively).
ERP responses over the hemisphere ipsilateral to the test stimuli

Figure 2.20 presents the pooled grand averages during ipsilateral stimulus presentations.

The point-by-point t-tests revealed that the onset of this effect was 134, 168 and 125, 116 ms for the long term adapted SAME and DIFFERENT conditions over the right and left hemispheres, respectively. In case of the short-term adaptation the onsets were 112, 116 and 125, 120 ms for the short term adapted SAME and DIFFERENT conditions over the right and left hemispheres, respectively.

The N170 amplitudes (Figure 2.21a) were lower in the adapted conditions than in the control condition for both hemispheres and for both adaptation durations (main effect of adaptation condition for all four comparisons: F (2,120)>9; p<0.0001).
No differences were found between the SAME and DIFFERENT adaptor-target arrangements, although in case of short-term adaptation there was a tendency for larger N170 amplitudes in case of DIFFERENT than in case of SAME condition (long term adaptation: post-hoc t-tests, p<0.7 and p<0.3 for the right and left hemispheres, respectively; short-term adaptation p<0.1 and p<0.07).

Long-term adaptation caused a significant increase on the P100 component in case of the ipsilateral stimuli (Figure 2.21b; main effects of adaptation: F (2,38)=7.9; p<0.001 and F(2,38)=6.8; p<0.003 for the left and right sides, respectively). SAME amplitudes were larger than DIFFERENT ones over the left hemisphere (post-hoc t-test: p<0.04).

After short term adaptation the P100 amplitude difference, was only significant over the right hemisphere (F (2,38)=1.3; p=0.28 and F(2,38)=4.8; p<0.01 for left and right, respectively) and no difference was observed between SAME and DIFFERENT (post-hoc t-tests: p=0.1 and p=0.8 for right and left hemispheres, respectively).

P100 and N170 amplitude values showed no significant correlations over the ipsilateral sites either (r= 0.07 and r=0.08 for the left and right hemispheres, respectively).

The N170 latencies (Table I.) in the long-term adaptation condition showed a significant main effect of adaptation condition because of the shorter latencies after adaptation compared to Control over both hemispheres (F(2,120)=9 p<0.00001) and (F(2,120)=11 p<0.0006) for left and right). SAME and DIFFERENT N170 latencies showed no differences (post-hoc t-tests: p=0.2 and p=0.7 for right and left hemispheres, respectively).
In case of short-term adaptation the latencies were longer than for control, but only over the right hemisphere (main effect of adaptation condition: F(2,120)=15 p<0.00001 for right and F(2,120)=2 p<0.2 for the left). Just like with long-term adaptation SAME and DIFFERENT latencies were similar (post-hoc t-tests: p=0.4 and p=0.3 for right and left hemispheres, respectively). SAME and DIFFERENT N170 latencies showed no differences (post-hoc t-tests: p=0.7 and p=0.5 for right and left hemispheres, respectively).

The P100 latencies (Table II.) showed no significant main effect of adaptation condition after long-term adaptation (F(2,38)=1.6 p=0.2 and F(2,38)=0.2 p=0.9 for left and right, respectively). Short-term adaptation, however, increased the P100 latencies (F(2,38)=7.5 p<0.003 and F(2,38)=8.6 p<0.002 for left and right, respectively). SAME and DIFFERENT were similar for both short and long-term adaptation, over both the left and right hemispheres (post-hoc t-tests: p>0.3 for all four comparisons).

---

**Summary of the results**

The results of the present study can be summarized as follows:

- Long-term (5 s) face adaptation evokes after-effects consisting a position-invariant as well as a position-specific component.

- When the adaptation time was reduced to 500 ms the resulting face after-effects were found to be entirely position-invariant and no position-specific adaptation effects were observed.

- In accordance with our behavioral results, the N170 ERP component - recorded over the hemisphere contralateral to the test stimulus - shows position-specific component only in the case of long-term adaptation: adaptation effects expressed in the N170 amplitude are larger for SAME than for DIFFERENT adaptor-test stimulus arrangement only in the long-term, but not in the short-term adaptation conditions.
2.4 Electrophysiological correlates of face distortion after-effects

In the last decades a number of studies have demonstrated that prior exposure to a face can bias our perception in a predictable manner, causing after-effects (Leopold et al., 2001; Webster and MacLin, 1999). For instance looking at a distorted face for a prolonged time the original, undistorted face will be perceived as distorted towards the opposite direction (see for an example earlier, in Chapter 1.2.2, Figure 1.10), a phenomenon called face distortion after-effect (FDAE) (Webster and MacLin, 1999). Such face adaptation related after-effects (FA) are present for various aspects of face processing (identity (Leopold et al., 2001; race, expression (Webster et al., 2004), eye-gaze direction (Schweinberger et al., 2007b), distortion (Webster and MacLin, 1999) and gender (Webster et al., 2004; Kovács et al., 2005) and imply, that adaptation is not a unique property of lower level visual cortical areas, but occurs also on higher processing levels.

Hierarchically higher processing steps in FA are suggested by evidences, such as the fact that FAs are, to a large extent, invariant to changes in size (up to a 4-fold difference in area; Anderson and Wilson, 2005; Kovács et al., 2006), orientation (adaptor and test face oriented ±45° from vertically upright; Rhodes et al., 2003; Watson and Clifford, 2003), colour and contrast (isoluminant red and green adaptor and test images and a contrast difference of 90% between adaptor and test faces; Yamashita et al., 2005) and partly in position (10° position difference between adaptor and test images, presented on the left and right of the fixation spot; Kovács et al., 2005, 2007, 2008). These findings support the idea that face selective neuronal processes at the higher stages of visual processing, tolerating the above changes of the stimuli, can adapt and might represent the neuronal basis of face-specific after-effects.

The goals of the present study were the following. First, we aimed at investigating the electrophysiological correlates of the FDAE. While most of the related previous ERP
(Schweinberger, Pfütze & Sommer, 1995, Schweinberger, Pickering, Jentzs ch, Burton & Kaufmann, 2002; Campanella, Quinet, Bruyer, Crommelinck & Guerit, 2002; Henson, Goshen-Gottstein, Ganel, Otten, Quayle & Rugg, 2003; Werheid, Alpay, Jentzsch & Sommer, 2005 Heisz, Watter, & Shedden, 2006) and MEG studies (Boehm, Sommer, & Lueschow, 2005; Harris & Nakayama, 2006) usually were testing how repeated short (< 1 s) presentation of a specific image affects the magnitude of neuronal responses to this stimulus in different visual cortical areas — studies reporting shape-specific aftereffects invariably used a prolonged adaptation period, lasting for several seconds or even for minutes (Leopold et al., 2001; Rhodes et al., 2003; Webster et al., 2004; Webster & MacLin, 1999; Yamashita et al., 2005). These, taken together with the fact, that different adaptation durations lead to altered neuronal responses at different steps of visual processing (Kovács et al., 2007, 2008) pose serious concerns regarding the possibility of making inferences about the neuronal mechanisms of face specific adaptation based on the results obtained with short adaptation using repetition suppression paradigms.

Previously, in ERP studies (Kovács et al., 2005, 2006, 2007) it has been shown, that adaptation to the gender of a human face or hand is reflected in the properties of the face-specific negative ERP component (Bentin et al., 1996; Rossion et al., 2000) – the so-called N170-, peaking at around 150-210 ms over the posterior occipito-temporal regions. Kovács et al. (Kovács et al., 2006) found, that the amplitude of the N170 decreases, while its latency increases after being adapted to a face or hand in a category-specific manner.

Since theories of face perception suggest, that the changeable aspects of faces (such as gender, age or expression) and facial identity are processed by separate neural systems (Haxby et al., 2000) it is possible that the previously observed N170 adaptation effects reflect specifically the processing of changing facial features. To study the generality of the effect here we measured ERPs associated with an after-effect, theoretically associated with the coding of facial configuration and with identity-specific information, the FDAEs of familiar faces (Webster and MacLin, 1999; Jeffery et al., 2006; Carbon et al., 2007). Indeed, recent results suggest, that the N170 to a repeated compared with an unrepeated face identity is reduced (Jacques, d’ Arripe & Rossion, 2007), even when the two faces are presented from different viewpoints (Caharel, d’ Arripe, Ramon, Jacques & Rossion, 2009). We hypothesized that if this N170 adaptation effects is due to the identity processing of faces than it should also manifest when the configuration of the stimuli is altered by distortions.
Second, previously the N170 adaptation effects were observed in comparison to a Fourier-spectrum randomized, noise image. Thus it is possible, that the changes of the N170 component reflect the adaptation of a general face sensitive mechanism (Schweinberger et al., 2007b), rather than reflecting the specific facial-feature related processing. Thus in our present experiments we used the original, undistorted faces as adaptors as well. Since the original and the distorted images only differ with respect to the relevant local feature information and feature configuration and previous behavioural studies showed that no aftereffects emerge to undistorted faces (Webster & MacLin, 1999) the comparison of the two adaptation conditions could reveal the specific FDAE related processing. Specifically, we hypothesized that if the neuronal processes, manifest in the early N170 ERP component, reflect information about facial configurations we should observe larger N170 adaptation effects after distorted faces as compared to the original faces.

Materials and Methods

Methods were similar to those of Chapter 2.1, thus here only the differences are mentioned.

Subjects: Twelve naïve, healthy subjects (5 females) participated in the study (mean age: 22±3). They all had normal or corrected-to-normal vision and gave written informed consent. We conform to the Declaration of Helsinki and the Ethical Committee of the Budapest University of Technology and Economics approved the study.

Stimuli: We used full-front grey-scale face images of two celebrities (namely Angelina Jolie and George Clooney2), presented centrally on uniform grey background (luminance: 1.3cd/m², viewing distance: 57 cm). We distorted the faces with the help of PhotoShop 6.0 „Pinch” option, which „pokes” the image inward or outward using a Gaussian model (Rhodes et al., 2003). So this distortion affected the internal face features but not the outer contour of the face or head shape. Test faces were presented at 8 distortion levels, (−20%, −15%, −10%, −5% expansions and 5%, 10%, 15%, 20% contractions; similarly to Jeffery et al. (Jeffery et al., 2006)).

Four different adaptation conditions were presented in separate blocks for both celebrities. In the Control condition the adaptor stimulus was the Fourier phase-randomised version of the original faces (Kovács et al., 2005). In the other three conditions we used the veridical faces (VERID) and their +25 or -25% distorted versions as adaptors (+DIST; -DIST). All software was written in MATLAB 6.5 (MathWorks Inc.) using PsychToolbox 2.45 for Windows.

2 First we performed a behavioural pilot experiment (n=15), where we tested the FDAE for 10 well known actors, actresses and politicians. Then we have chosen for the subsequent ERP recording expriments a female and a male celebrity face, which were reported the best known among our subjects.
**Task and electroencephalogram recording procedures:** In each trial – after a blank screen, presented for a random time between 500 and 700 ms – the adaptor was presented for 5000 ms, followed by a 500 ms gap, and then by the appropriate test stimulus, presented for 200 ms (Figure 2.22).

Subjects were instructed to fixate in the centre and after the disappearance of the test stimulus press a button when they perceived the test face expanded and another button when it appeared contracted when compared to the real, undistorted face of the given famous person. Control, +DIST and -DIST adaptor conditions of the two celebrities were given in separate blocks. The order of the adaptor conditions was randomised across subjects. In a given block we presented 40 trials (8-number of levels * 5-number of repetition of a given test stimulus) in a random order. Recording sessions last approximately 50 – 70 minutes. ERPs were recorded via 32 Ag/AgCl electrodes placed according to the international 10/20 system (Ground: FT9, Reference: AFz, impedance kept below 5 kΩ). EEG was digitized at 1024 Hz sampling rate. The EEG was segmented offline (Brain Vision Analyzer, Brain Products GmbH, Munich, Germany) into 1100 ms long trials, using a 100 ms prestimulus interval. After 30 Hz low-pass filtering of the EEG, trials containing blinks, movements, A/D saturation on EEG baseline drift were rejected on the basis of visual inspection of each recording by semi-automatic artefact detection (-35; +35 mV) standard deviation over 200 ms intervals on all electrodes with blink artefact correction.

After cleaning the ERP data we could use approximately 85% of the registered segments (on average 34 trials/conditions). ERPs were averaged separately for each subject, condition and channel. Averages were then digitally filtered (0.1 – 70 Hz) with a zero phase shift digital filter and average re-referencing was calculated.

**Behavioural data analysis:** Psychophysical data were modelled by a Weibull psychometric function, using the psignifit toolbox (ver.2.5.6.) for MATLAB (http://bootstrap-software.org/psignifit/). To test whether adaptation to the veridical face, to the distorted face or to the Fourier phase-randomised version of the faces results in different effects on subsequent face distortion discrimination a three-way ANOVA was performed (types of adaptor (4), identity (2) and distortion level (8)) as within-subject factors.

**ERP data analysis:** We measured the amplitude and latency of the P100 at 5 occipital (O1, PO3, Oz, O2, PO4) and of the N170 component at 8 occipito-temporal electrode sites (P7, P8, PO7, PO8, P9, P10, PO9, PO10) (Rossion et al., 2000, Caharel et al., 2009), respectively. After averaging, we searched for the individual
component peaks in a 40 ms window, centred on the respective ERP component maximum of the grand average of each electrode (P100: 130-170 ms, N170: 180-220 ms) for each subject and condition separately (Itie & Taylor, 2004). Both amplitude and latency values were measured here and entered into a two-way repeated measures ANOVA with adaptor type (4, Control, +DIST, -DIST, VERID) and hemisphere (2, two electrodes for each hemisphere for the P100 and four electrodes per hemisphere for the N170 analyses) as within-subject factors. In the amplitude analyses we compared the +DIST and -DIST conditions with a point by point t-test for the [100,300] ms time interval. Since the two distorted conditions were not statistically different we averaged them and entered the values in a two-way repeated measures ANOVA with adaptor type (3, Control, DIST, VERID) and electrode (see above) as factors. Since previous results regarding the electrophysiology of face perception (Rossion et al., 2003) suggest that the left hemisphere (LH) and the right hemisphere (RH) behave differently we analysed LH and RH electrodes in separate ANOVAs. Post-hoc t-statistics were performed by Fisher’s least significant difference tests.

Results

Behavioural results

Adaptation to a positively/negatively distorted face (Figure 2.23) led to a strong perceptual after-effect – test faces were judged more expanded in the +DIST (contracted) condition and vice versa, when compared to the Control condition (main effect of adaptor type: F(21,344)=114,2, p<0.0001, the interaction ‘adaptor × distortion level’: F(21,344)=8,2, p<0.0001).

However, we found no significant differences between the Control and the VERID conditions in the behavioural data (post-hoc test for conditions: p=0.06), suggesting that neither the veridical face nor its Fourier phase randomised version affects the perception of inner facial feature configuration. Due to the large number of trials required by the ERP recordings we had to limit our choice of stimuli to two.

Figure 2.23: Average (± standard error) ratio of negatively distorted face responses as a function of distortion level (% distorted). Results with control, positively (+DIST), negatively (-DIST) distorted and veridical (VERID) adaptors. The inset illustrates examples of the test stimulus displays.
This low number of exemplars may limit the generalization of our findings to general mechanisms of face perception. The fact, however, that we observed no main effect of identity (F(21,344)=0.4, p=0.5) suggests similar processing of the chosen female and male face stimulus. Thus we collapsed our ERP results of different identities.

**Event-related potential results**

Similarly to our previous studies (Kovács et al., 2005, 2006, 2007) we identified the two main early ERP components – *P100* and *N170* – and measured their parameters as a function of adaptation.

**Amplitude**

Since we found no significant differences between the amplitude of *+DIST* and *-DIST* conditions in the [100,300] ms time interval (point-by-point t-test, p>0.3 for the LH electrodes and p>0.5 for the RH electrodes for each time point) and the post-hoc analysis of *N170* latency was also not statistically different (post-hoc test for conditions: p>0.5 for the LH and p>0.1 for the RH), we averaged the two above-mentioned conditions. In the present experiment adaptation with a face stimulus - both veridical and distorted – had a significant effect on the amplitude of the *N170* over both hemispheres so that the amplitude in the *Control* condition was significantly larger than in the other conditions (Figure 2.24a, b and d; main effect of adaptor type for the LH electrodes and RH electrodes: F(9,132)=167.8, p<0.0001 and F(9,132)=136.2, p<0.0001, respectively and post-hoc test for *Control* vs *DIST* and *Control* vs *VERID*: p<0.001 and p<0.001, respectively).
In addition, we found significantly smaller N170 amplitudes in the DIST condition, when compared to the VERID condition (F(3,44)=9.02, p=0.004) over the RH. This difference did not reach the level of significance over the LH (F(3,44)=1.23, p=0.27) (Figure 2.24c). The lateralisation of the adaptation effect is clearly visible on the voltage maps of the ERP differences between VERID and DIST (Figure 2.24e). Given the small nature of the effect size over the RH, it is possible, that the differences between VERID and DIST are driven by one of the two distorted conditions alone. Thus we performed two additional tests comparing +DIST and –DIST separately with VERID. These tests suggest that the amplitude of the N170 component is significantly lower in both distorted conditions when compared to the VERID condition (+DIST vs VERID: (F(1,94)=3.79, p=0.048) and –DIST vs VERID: (F(1,94)=9.89, p=0.0022)).
The \textit{P100} component was significantly larger in \textit{VERID} and in \textit{DIST} conditions, when compared to \textit{Control} (main effect of adaptor type: F(4,104)=82.5, p<0.001, the interaction ‘adaptor × hemisphere’: F(4,104)=0.27, p=0.9 and post-hoc test for \textit{Control} vs \textit{DIST} and \textit{Control} vs \textit{VERID}: p<0.001 and p<0.001, respectively). However, unlike in case of \textit{N170}, this effect was similar for \textit{VERID} and \textit{DIST} for both left and right hemisphere electrodes (post-hoc test for conditions: p=0.48 for the left electrodes and p=0.66 for the right electrodes, respectively).

\textit{Latency}

The \textit{N170} was delayed after adaptation to \textit{VERID} and \textit{DIST}, when compared to \textit{Control} (main effect of adaptor type: F(2,188)=65.73, p<0.001, the interaction ‘adaptor × hemisphere’: F(2,188)=0.51, p=0.6 and post-hoc test for \textit{Control} vs \textit{DIST} and \textit{Control} vs \textit{VERID}: p<0.001 and p<0.001, respectively).

\textit{P100} latencies were also longer in the \textit{VERID} and \textit{DIST} adapted conditions than in the \textit{Control} condition for both hemispheres (main effect of adaptor type: F(4,114)=107.33, p<0.001, the interaction ‘adaptor × hemisphere’: F(4,114)=0.05, p=0.99 and post-hoc test for \textit{Control} vs \textit{DIST} and \textit{Control} vs \textit{VERID}: p<0.001 and p<0.001, respectively). No other latency differences were significant.

\textbf{Summary of the results}

The results of the present experiment can be summarized as follows:

- FDAE, similarly to gender adaptation, is also reflected in the amplitude and latency of \textit{P100} and \textit{N170} ERP components.
- The \textit{N170} adaptation effect over the right hemisphere is larger for distorted faces, than for the veridical faces, which suggests that the observed ERP adaptation effects are general for various steps of face processing and that the FDAEs similarly to gender after-effects, are related to the early face specific ERP components.
2.5 Modulation of the facial after-effects by transcranial direct current stimulation

Although functional brain imaging provides evidence for transient local activity changes during the execution of specific perceptual and cognitive processes, it does not permit inferences about causal structure-function relationships (Sack et al., 2007). Such inferences can only be made based on a controlled manipulaiton of brain activity whose impact on behaviour or cognition can be quantified. The technique of transcranial direct current stimulation (tDCS) is a promising tool for inducing so-called virtual lesions non-invasively and reversibly. This stimulation results in excitability changes of motor and visual cortices in humans, which evolve during tDCS but outlast the stimulation for up to an hour, given a sufficiently long duration of stimulation (Nitsche et al., 2004). Anodal tDCS increases, while cathodal tDCS increases excitability (Antal et al., 2003; Nitsche and Paulus, 2000, 2001). The initial effect of DC stimulation is accomplished by a hyper- or depolarization of neuronal membranes (Nitsche et al., 2003). Thus, external modulation of neuronal excitability by tDCS provides an opportunity to investigate how changes in the neural sensitivity in specific visual cortical areas affect visual perception in humans. It is known from animal studies (Creutzfeld, 1962; Bindman, 1964) that by using weak DC stimulation we are able to change the firing rate of the neurons in the motor and visual cortices – cathodal stimulation hyperpolarize the membrane potentials of cortical cells, whereas anodal stimulation results in a reversed effect (Nitsche and Paulus, 2000, 2001). Analogous effects have been observed recently in humans – using cathodal stimulation on the motor cortex we are able to reduce the amplitude of motor-evoked potentials (MEPs), whereas anodal stimulation increases them (Antal et al., 2004). This physiological effect is also mirrored in the behavioural performance in perceptual tasks. For example, it has been shown that cathodal stimulation of the primary visual cortex increases the contrast sensitivity threshold, whereas anodal stimulation does not affect it
tDCS can also manipulate the phosphene threshold (Antal et al., 2003). We can also observed effects on higher-level visual areas, as V5. The application of DC stimulation on the left V5 enhanced the performance of subjects in a visuo-motor coordination task (Antal et al., 2004a). Similar effects have been found in simple motion perception tasks (Antal et al., 2002). The stimulation affects not only the visual but also the motor cortex – tDCS also modifies the motion detection threshold (Antal et al., 2004a). Recently it was also applied in higher cognitive tasks – stimulation modified our decision criteria (Korsakov and Matveeva, 1984) and applying anodal stimulation on the left prefrontal cortex we can affect the performance of individuals in a working memory task (Fregni et al., 2005) or a probabilistic learning classification task (Kincses et al., 2003).

It can also be tested whether tDCS also alters our performance in an adaptation experiment. Since motion after-effect (MAE) is one of the most widely known and used perceptual manifestation of the neuronal adaptation process (Anstis et al., 1998) Antal et al., (Antal et al., 2004b) have tested if tDCS is able to change the magnitude of MAE. The authors have found that both anodal and cathodal stimulation of the human MT+/V5 decreased MAE duration.

In the mirror of these results we tested whether tDCS can have any effect on a facial after-effect.

Human face perception is one of the basic aspects of social interactions in humans from early childhood till elderly (Haxby et al., 2000, 2002, Haxby, 2006) and involves many different mechanisms and components. Haxby (2002), modifying the most influential face-perception model of Bruce and Young (Bruce and Young, 1986), emphasizes the difference between the changeable and invariant components of face perception. These components are probably linked to the primary somatosensory cortex (SI) and to the superior temporal sulcus (STS). The changeable aspects of the process (gaze direction, lip movement, emotional expressions, angle of profile) play a significant part in social communications, while the invariant components (identity) are the basis for recognizing individuals (Haxby et al., 2002, Pourtois et al., 2004). These later processes may take place in the inferior temporal (IT) cortex. Indeed, many functional imaging studies show the existence of a more-or-less specific face processing network in the human brain (Kanwisher, 2002, 2003). Furthermore, the existence of a special neural network for face perception in humans was also implied by the observation of patients with bilateral lesion of the ventral occipito-temporal cortex (Haxby et al., 2000). Such patients, suffering from prosopagnosia are unable to recognize familiar face but the recognition of other objects remains unaffected.
Neuroimaging (Haxby, 2006; Downing et al., 2001; Puce et al., 1998; Grill-Spector et al., 2006; Ng et al., 2006) and electrophysiological studies in humans (Kovács et al., 2006) tried to elucidate the precise location of visual recognition of distinct body parts. They implicated the role of STS, fusiform face area (FFA), ventral occipito-temporal cortex, cingulate gyrus (CG) and extrastriate body area (EBA)-mainly in the right hemisphere- in face recognition. Additionally, recent fMRI studies posed a new consideration of face recognition (Haxby, 2006; Grill-Spector et al., 2006). They suggest that the anatomical domains are morphologically interconnected cortical regions, called cortical spots. These spots responded selectively to different object categories, other than faces (Haxby, 2006; Grill-Spector et al., 2006). Concerning face recognition, different aspects of face recognition tend to pertain to distributed cortical networks (Ng et al., 2006; Webster et al., 2004). It was also observed in imaging and transcranial magnetic stimulation (TMS) studies that EBA regions activated by body parts do not overlap with those responding specifically to faces (Haxby, 2006; Urgesi et al., 2004).

Facial adaptation is a special case of the after-effects (AE): a prolonged exposure to an individual face will bias the perception of a subsequent face (Leopold et al., 2001). These AEs are shape-, orientation- and category-specific (Webster et al., 2004), depend on the duration of adaptor face (Rhodes et al., 2003; Watson and Clifford, 2003). Since these facial after-effects, to a certain extent, are invariant to the changes of the relative position, and size (Leopold et al., 2001; Rhodes et al., 2004) of adaptor and target faces, it was proposed that they involve higher-level, non-retinotopic mechanisms (Kovács et al., 2006; Niedeggen and Wiest, 1998). However, both changes of retinal position (Kovács et al., 2005) and size (Bruce and Young, 1986; Yamashita et al., 2005) reduce face after-effects to some extent, suggesting also the involvement of retinotopic mechanisms.

Transcranial direct current stimulation (tDCS) is able to influence the excitability of human visual and motor cortices focally, reversibly and non-invasively (Nitsche and Paulus, 2000; Antal et al., 2001). Early animal studies have shown that weak cathodal stimulation decreases cerebral excitability due to membrane hyperpolarization, while anodal stimulation increases it through membrane depolarization (Bindman et al., 1962; Purpura and McMurtry, 1965). The tDCS induced after-effects can be prolonged by enhancing the current intensity or duration (Nitsche and Paulus, 2000). Thus, the method is useful to reduce or increase the excitability of a particular cortical area reversibly and non-invasively. The relatively low cost and the ease of the technique made it as a popular tool of studying visual and motor cortical processing.
The goal of our study was to determine which cortical areas play a role in facial adaptation and the formation of face after-effects using tDCS. We were interested in whether anodal or cathodal tDCS affecting low-level, retinotopic steps of visual processing in the V1 or higher-level mechanisms in the right temporo-parietal cortex are responsible for the facial gender after-effect.

**Materials and Methods**

**Subjects:** 17 naive, healthy volunteers (8 females) participated in the studies (7 in the occipital (3 females) and 10 in the parietal (5 females) stimulation task) (age range: 18-30 years). They all had normal or corrected-to-normal vision, no metallic implants, no previous neurological, psychiatric or ophthalmologic disorders, drug abuse, or alcoholism and were not under medication at the time of the experiments. They gave written informed consent. We conform to the Declaration of Helsinki and the Ethics Committee of the University of Göttingen approved the study.

**Stimuli and conditions:** everything was identical to those of Chapter 2.1.

**tDCS:** tDCS was delivered by a battery-driven constant current-stimulator (see also in Figure 2.25; Neuro Conn GmbH, Ilmenau, Germany) using a pair of rubber electrodes in a 5 x 7 cm water-soaked synthetic sponge. For the occipital stimulation one electrode was placed over the Oz the other over Cz according to the international 10/20 system. During the temporo-parietal stimulation one electrode was over the P6-P8 and the other over Cz. The electrode was placed over Oz or P6-P8 and the other over Cz. For anodal stimulation the direction of the current was changed.

The current was applied for 10 minutes with an intensity of 1.0 mA. All subjects received both anodal and cathodal stimulation in randomized order, separated by at least 1 week to avoid interference effects. As a control we performed sham stimulation. For this stimulation, the electrodes were also placed over the same positions as in the stimulation conditions, but the current was switched off after a 5s stimulation. Subjects were blinded concerning the type of stimulation.
**Experimental procedures:** Subjects were seated in a dark room (background luminance: 3 cd/m2). They were asked to fixate a cross in the centre of the computer monitor. In each trial –after a randomized period between 500 and 700 ms – the adaptor was presented for 5000 ms followed by the test stimulus, presented for 200 ms. Subjects were instructed to perform a two-alternative forced choice gender discrimination task (left mouse button – female, right mouse button – male). The reaction times (period till pressing the mouse button) were measured and the ratio of female/male responses were counted during the experiments.

**Statistical analysis:** The magnitude of the adaptation was calculated as the difference of performance between the NOADA (without adaptation) and ADA (with adaptation) conditions for all of the 9 morph levels for the parietal and occipital tDCS stimulation separately. The performance and latency differences were entered into a 3 (stimulation: cathodal, anodal, sham) x 9 (difference at each morph level) ANOVA. For post-hoc test Tukeys HSD test was used.

**Results**

All of the subjects were able to complete the experiment. We observed that adaptation to a 100% female face stimulus causes strong perceptual after-effect: the test faces were perceived as more masculine, as compared to the control condition (main effect of adaptation: F(8,72)=2.53, p=0.017). Figure 2.26 illustrates the effect of adaptation on facial gender discrimination under sham stimulation condition.

![Figure 2.26: The graph shows a control (NOADA) condition, in which the adaptor stimulus was a grey circle, and an adaptation (ADA) condition, in which we used a prototypical 100% female face as adaptor. Note, that adaptation to a 100% female picture modifies gender discrimination. Bars show standard errors](image)

**Temporo-parietal stimulation**

The ANOVA revealed a significant interaction between the type of stimulation and the differences between the NOADA and ADA conditions at the 9 morph levels (F(2,36)=6.31, p=0.004).
According to the post-hoc test, cathodal stimulation decreased the magnitude of adaptation (p<0.05) (Figure 2.27), while neither the sham nor the anodal tDCS modified facial adaptation significantly (p>0.05). We also analyzed the effect of tDCS on the reaction times’ differences between the NOADA and ADA conditions. We found significant interaction between the type of stimulation and reaction times differences (F(2,52)=7.78, p=0.0001).

The post hoc test showed that cathodal stimulation could significantly decrease the reaction time difference between the NOADA and ADA conditions (p=0.009).

**VI stimulation**

Figure 2.28 shows the NOADA-ADA differences in performance after cathodal, anodal or sham stimulation. The stimulation had no significant effect on adaptation over the primary visual area (F(2,38)=0.20, p>0.05).
Summary of the results

The results of the present experiment can be summarized as follows:

- In the present study it was shown that cathodal transcranial weak direct current stimulation over the right temporo-parietal areas decreased the strength of facial adaptation effect.
- tDCS over the primary visual cortex resulted in no effect.
In a series of experiments that consisted of 5 studies we investigated the psychophysical and electrophysiological aspects of higher level configural after-effects, evoked by adaptation. We tested adaptation effects both to the gender of human faces and hand stimuli and distorted versions of familiar faces. After exposure to a related stimulus (a normal or distorted face or a human hand) we found strong perceptual bias (after-effect) in the performance of the subjects participating in our experiments. The electrophysiological data of these aforementioned experiments, suggest for the first time that this effect is already reflected in a very early ERP component, namely on the $N170$.

In the next section I will discuss our experimental data in the mirror of our main results which are the followings:

1. First, our electrophysiological and transcranial direct current stimulation data suggest that shape-selective after-effects mainly reflect higher-level adaptation processes and are not due to local adaptation to image elements.
2. Second, these effects are specific to a given category – no cross-category effects were found.
3. Third, we found hemispherical asymmetries in the strength of the adaptation effect, both in psychophysical and electrophysiological data.
4. Fourth, the duration of adaptation is an important factor in shape-selective after-effects – in a short adaptation paradigm no position-specific adaptation effects were found.
1. **Low-level versus high-level adaptation processes in shape-selective after-effects**

As seen before we can say that if the adaptor and test stimuli are different in a given property but the neurons are invariant to this altered stimulus and that’s why AE appears, we can guess that the neurons involved in the effect can be located in higher level visual areas (like IT). So in a psychophysical experiment, by manipulating well the size, orientation or retinal position of the adaptor and test stimuli we can argue that this effect is not due to the summation of low-level processes. By combining these measures (results) with electrophysiological recordings or transcranial DC stimulation we can get closer to the precise localization of the certain after-effect.

*Behavioural evidences*

Our findings that test hand images were judged more masculine or feminine as a result of adaptation to a female or a male hand, respectively, provides evidence for after-effects induced by adaptation to human body parts. Although, Kohler and Wallach (Kohler and Wallach, 1944) have reported that they could observe figural after-effects after “inspection of any specific entity in the visual field”, including human hands, the study presented first in my dissertation represents the first systematic investigation of the shape-selective adaptation mechanisms involved in the processing of human body parts. The existence of hand adaptation entailed an important question: what are the stages of visual processing where neuronal adaptation leading to hand after-effects could take place. It has been shown that facial after-effects generalize across differences in the size (Zhao and Chubb, 2001, Rhodes et al., 2004), retinal position (Leopold et al., 2001) and orientation (Rhodes et al., 2003) of faces, suggesting that the underlying neuronal adaptation takes place at the higher stages of face processing. We have shown that the newly described hand after-effects are also robust to changes in hand size or orientation, implying that these after-effects arise from the adaptation of the higher-level, non-retinotopic shape-specific mechanisms of object coding and cannot be based solely on combination of adaptation to low-level features.

To explain how adaptation can bias the perceived gender of face and hand stimuli would require that we understand: on the one side, what are the cues, what is the information – specific features or feature configurations - present in face and hand images that are used by humans to determine gender; and on the other side, which of these cues are affected by adaptation. As for the cues that observers might relay on during gender judgment, in the case
of faces, recent studies suggest that the region of the eyes and their second order relationships might be of primarily importance (Smith et al., 2004). In the case of hands, the question of relevant cues during hand gender judgments has not yet been investigated directly before. The results of the present study suggest that size might be one of the features, which is taken into account when the gender of a hand is determined; it was found that the same hand images were judged more masculine when their size was increased. However, our finding also clearly shows that gender judgment is based on more than one cue, since even though the size of the hands itself has a strong affect on gender judgment, adaptation-induced bias of the perceived gender was to a large extent size invariant. This suggests that adaptation affected the processing of those complex hand features, relevant for gender judgment, which are coded at the higher-level stages of object processing where neurons are already size-invariant. It is a challenge for further studies to determine the specific processing mechanisms that are modulated by shape-selective adaptation and the way they affect human judgments about different properties of visual objects.

According to the psychophysical results of the secondly presented experiment, facial after-effects following adaptation to upright or upside-down faces consist of both position-specific and position invariant components. The observer’s gender judgments were biased - i.e. test faces were judged more masculine – significantly more in the conditions where the adaptor and test faces were presented in the same retinal position as compared to when they were displayed in different visual hemifields. These results are in apparent conflict with a previous study (Leopold et al., 2001), where facial after-effects were almost completely translation invariant. However, the experimental designs of the two studies are very different, which might explain why differential results were obtained. First, to test for translation invariance of facial adaptation, Leopold et al. (Leopold et al., 2001) made observers to fixate at different points of the faces during adaptation and test, whereas in our study adaptor and test faces were presented in different visual hemifields. Second, the task, which was used to reveal facial after-effects, was also different in the two studies: Leopold et al. (Leopold et al., 2001) used a face identification task whereas we used a gender discrimination task. According to recent results of face processing (Haxby et al., 2000) facial identity and gender are processed in different sub-regions of the human temporal cortex, thus the different positional invariance of identity and gender specific after-effects might reflect separate neuronal processing of these attributes.
To sum up further behavioural results we can say that shape-selective adaptation effects not only based on local, low-level adaptation to simple pictorial cues but also on structural processing.

*Electrophysiological recordings and tDCS evidences*

We found that the early components of the ERP also reflect high-level adaptation processes.

*P100*

Based on recent findings it appears that *P100* may reflect more than simply the processing of low-level visual features. Relevant to the study presented first, recently it was found that *P100* shows some sensitivity to specific manipulations of visual faces – e.g. face inversion - and it was suggested that it might reflect the holistic processing of a face as a face (Itier and Taylor, 2004a, but see Rossion et al., 1999, 2000 for conflicting results). We found that in the conditions with face adaptors the amplitude of the *P100* evoked by the subsequent test stimulus was larger than that in the control conditions.

*N170*

*N170* is a negative ERP component peaking around 140-200 ms over the lateral occipito-temporo-parietal regions, which is believed to reflect the structural encoding of facial features. Although, other non-face objects can also evoke a negative ERP component with similar delay, the amplitude of the *N170* is larger and its latency is shorter for faces as compared to other objects (Rossion et al., 1999, Itier and Taylor, 2004a, 2004b). The selectivity of *N170* for faces is strongly supported by recent studies where analyses of scalp topographies of *N170* evoked by faces non-face object categories revealed an extra map for the face *N170* in the lateral temporal cortex, which was not present in the case of *N170* to objects (Itier and Taylor, 2004a).

The ERP results of our control conditions are consistent with previous findings. Face test stimuli – when presented following adaptation to Fourier-randomized face images - evoked strong *N170* over both hemispheres, the amplitude of which was larger over the right hemisphere than over the left. The human hand test stimuli also evoked bilateral *N170* and similarly to faces the *N170* over the right hemisphere was larger than that over the left. The
amplitude of $N170$ for faces at the right parieto-temporal electrode (P8) was significantly larger than the $N170$ for hands at the same site, which is in accordance with the results of numerous fMRI (for review see Haxby et al., 2000) as well as ERP studies (Itier and Taylor, 2004a; Rousselet et al. 2004) showing that face-specific neuronal responses are more pronounced in the right hemisphere than in the left. At the occipital electrodes (both O1 and O2), however, the $N170$s for hands were significantly larger than those for faces. As an explanation for the larger $N170$ for hands than for faces at the occipital electrodes, one could speculate that hands might have evoked stronger responses in the early visual cortical areas than faces. In fact, the average luminance of hand stimuli was slightly higher than that of the face stimuli (see Materials and Methods section of the firstly presented experiment), and hand stimuli contained more figure/background contour information than faces, which might lead to a stronger activation at the early stages of visual processing. This possibility is supported by our findings that the amplitude of $P100$ evoked by the hand test stimuli was significantly larger than that evoked by face test stimuli. The stronger responses to hands than to faces at the early cortical areas could, in turn, affect the occipital $N170$ component, rendering its amplitude larger in the case of hand tests compared to that for face tests.

Face adaptor followed by face test and hand adaptor followed by hand test resulted in strong reduction of the $N170$ and $VPP$ amplitudes and increase in their latency at all analyzed electrodes (O1, O2, P7, P8, Cz) both in the case of faces and hands. These adaptation effects on $N170$ are in agreement with the previous ERP results on visual adaptation (Hoffmann et al., 2001) as well as the neuroimaging studies using the fMRI adaptation paradigm in humans (for review see Henson, 2003), showing decreased and delayed neuronal responses as a result of adaptation.

Although the precise source of $N170$ is not known, previous studies analyzing the scalp topography of the $N170$ generators indicate that its main source is located in the lateral occipito-temporal cortex (Itier and Taylor, 2004b). This is in accordance with previous findings showing that specific regions of the lateral temporal cortex are involved in the processing of facial information: intracranial electrophysiological recordings suggested the involvement of the middle temporal gyri (Allison et al., 1999a); whereas neuroimaging studies have identified a face-responsive region in the superior temporal sulci (STS, for review see Haxby et al., 2000). In particular, it was suggested that lateral temporal cortex is primarily involved in the processing of the changing facial features conveying physiognomic information (say facial expressions, or information about gender or age) rather than analyzing facial identity (Allison et al., 2000; Haxby et al., 2000). Moreover, recently, visual processing
of the information about human body-parts was also attributed to a specific region of the lateral occipito-temporal cortex, the so-called extrastriate body area (EBA, Downing et al., 2001). Using fMRI, it was reported that EBA is selectively activated when static images of different non-facial parts of the human body are viewed. Importantly, there was no overlap between the region in the STS activated by faces and the EBA (Downing et al., 2001). Thus, based on the above results, it is reasonable to suppose that \( N170 \) both for faces and for hands might originate from the lateral occipito-temporal cortex but from different neural populations, located in the STS and EBA, respectively.

We also demonstrated that FDAE, similarly to gender adaptation, is also reflected in the amplitude and latency of \( P100 \) and \( N170 \) ERP components. As a result of adaptation to a distorted or to a veridical face \( P100 \) is larger and delayed, while the \( N170 \) is reduced and delayed, when compared to a Fourier-phase randomised noise image.

Although Bruce and Young (Bruce & Young, 1986) have argued that the processing level of gender and identity information is different, we found very similar adaptation effects for the early \( P100 \) and \( N170 \) components in the current experiment, using face distortion aftereffects, a phenomenon related to identity-specific information (Jeffery et al., 2006; Carbon et al., 2007), and in our previous studies using gender specific adaptation effects (Kovács et al., 2006). This would suggest that the observed ERP adaptation effects are general for any stimuli belonging to the category of faces.

This is not surprising if one considers how we measured adaptation: we compared the ERP components, obtained during \( DIST \) adaptation with those of the \( Control \) condition, where the adaptor image was a Fourier-phase randomised noise image. This noise image lacks any characteristic face features thus, in comparison with it adaptation to both a distorted or to a veridical face may desensitize general face processing mechanisms equally. This desensitization is then in turn manifested in reduced synchronous changes of neural activity, leading to ERP effects (Grill-Spector, Henson & Martin, 2006), that is, to elevated and delayed \( P100 \) and decreased and delayed \( N170 \).

The observed adaptation effects on the \( P100 \) are in line with our previous findings (Kovács et al., 2006; Kovács et al., 2007) in the sense, that we observed adaptation effects already at the time window of \( P100 \) component, when compared to Control. This suggests that this component might reflect the early detection of faces, a conclusion supported (Itier & Taylor, 2004) but also questioned (Rossion, Campanella, Gomez, Delinte, Debatisse, Liard et al., 1999, Rossion et al., 2000, Rossion & Jacques, 2008, Rousselet, Husk, Bennett & Sekuler, 2008) in the current literature. The results regarding \( N170 \) also suggest, that the observed
adaptation effects reflect the adaptation of a mechanism that is sensitive to any face-like stimuli, regardless of the specific information, present in the image. This conclusion is supported by recent work on gaze adaptation (Schweinberger et al., 2007), which observed reduced and delayed N170 without any gaze direction-specific adaptation effect as well.

1.3 Evidences from transcranial direct current stimulation

In our tDCS study we have observed that cathodal stimulation of the right temporo-parietal areas decreased the strength of facial adaptation while tDCS over the primary visual cortex resulted in no effect. The data imply that higher-level, non-retinotopic cortical areas play role in the creation of facial after-effects and probably also in gender discrimination.

Several previous studies have tried to reveal cortical regions participating in face recognition and gender discrimination. Urgesi et al. (Urgesi et al., 2004) could interrupt the visual processing of non-facial objects by repetitive TMS of the EBA. Pourtois (Pourtois et al., 2004), using TMS found that the SI and superior temporal regions play a role in the perception of facial expressions and gaze direction, in accordance with fMRI data (Haxby et al., 2002; Downing et al., 2001). Electrophysiological studies determined the role and specificity of right parieto-temporal regions (P8) in face recognition (Kovács et al., 2006; Mouchetant-Rostaing et al., 2000; Bentin et al., 2002; Mouchetant-Rostaing and Giard, 2003; Rossion et al., 1999; Itier and Taylor, 2004; Nitsche et al., 2003). Recent ERP results suggested that these areas play a role in facial gender discrimination as well (Kovács et al., 2006).

Here we have found, using the combined technique of adaptation and tDCS, that cathodal stimulation could modify higher-level facial adaptation over the right temporo-parietal areas. In agreement with previous studies, the inhibitory effect of cathodal tDCS on adaptation is possibly related to the focal diminishment of cortical excitability due to membrane hyperpolarisation (Nitsche and Paulus, 2000; Antal et al., 2001), however the exact molecular mechanism is not known yet. Pharmacological studies tried to reveal the possible involvement of different intracortical neuronal receptors in the mediation of both short- and long-term effect of tDCS. They showed the selective prevention of the anodal tDCS-induced membrane depolarization by the voltage-dependent sodium channel blocker carbamazepine (CBZ) and the calcium channel blocker flunarazine (Nitsche et al., 2003; Liebetanz et al., 2002). Nitsche et al observed that both anodal and cathodal tDCS after-effects are N-methyl-D-aspartate (NMDA) receptor dependent, which are also involved in neuroplastic changes (Nitsche et al.,
They showed in another study that the administration of an antiepileptic drug, lorazepam –which is a γ-amino-butyric acid (GABA)A receptor agonist-, caused a delayed, but then enhanced and prolonged anodal tDCS-induced excitability elevation (Nitsche et al., 2004b). On the other hand it was demonstrated that D2-receptor blocking by sulpiride abolishes the induction of long-term tDCS after-effects. The combined administration of cathodal tDSC and pergolide (D1-receptor agonist) enhanced and prolonged the tDCS-induced excitability-diminishment (Nitsche et al., 2006). Amphetamine, a catecholaminergic re-uptake-blocker, could significantly enhance and prolong the long-lasting excitability enhancement in human motor cortex after application of anodal tDCS (Nitsche et al., 2004c). These effects denote the importance of cathecolaminergic neuroplastic mechanisms. The fact, that tDCS had no effect on facial after-effect over the V1, but had a strong effect over temporo-parietal sites suggests the importance of higher-level, non-retinotopic regions in facial adaptation. This finding is in accordance with previous neuroimaging and electrophysiological studies (Downing et al., 2001; Ng et al., 2006; Kovács et al., 2006; Leopold et al., 2005). Our data also imply that high-level neural processing in extrastriate areas play a remarkable role in gender discrimination. The outcome of these studies support that tDCS can be reliably used in the modification of neuroplastic processes in an experimental clinical setting (e.g. epilepsy, migraine, pain, Parkinson’s disease, rehabilitation after stroke or traumatic brain injury), studying plastic neural processes.

2. The category-specificity of the shape-selective after-effects

2.1 Behavioural evidences

An important behavioural result of the study presented first in my dissertation is that - in sharp contrast to our within-category adaptation experiments – we found no adaptation effects in the cross-category experiments. Adapting to faces did not affect the perception of the subsequently presented hand test stimulus and similarly, hand adaptors did not bias the perception of the face test. These results appear interesting in the light of the fact that the task that the observers performed during testing both in the case of face and hand stimuli was identical, namely gender discrimination. Recently, Fang and He (Fang and He, 2005) reported analogue results in a viewpoint discrimination paradigm: viewpoint aftereffects showed no cross-adaptation between object categories such as faces, cars and wire-like objects. The lack of the cross-category adaptation effects thus suggest that adaptation takes place at the level of
sensory processing of the specific shape information associated with these two different visual categories and excludes the possibility that adaptation effects are due to an adaptation-induced bias in the information extraction strategy or at the stage of decision making.

2.2 Electrophysiological evidences

P100

Consistently with the behavioural data, only within-category but not cross-category adaptation leads to an increase of the P100 latency. Adaptation-induced response latency increase has been shown previously for motion adaptation by recording visual evoked potentials in humans (Hoffman et al., 2001). Thus, the fact that adaptation effects on the P100 latency were category-specific is in accordance with the previous findings suggesting that the P100, in addition to the early visual analyses (Gomez et al., 1994, Heinze et al., 1994), to some extent also reflects shape-selective neural processing and indicates that its latency might be sensitive to the state of adaptation at these higher processing stages. The adaptation-induced P100 amplitude enhancement was not category specific, since it was present both in the case of face and hand test stimuli. Further studies are needed to uncover the origin of this effect.

N170

Importantly, N170 adaptation effects, in accordance with our behavioural results, are category-specific, since cross-category adaptation (face adaptor followed by hand test and vice versa) had no effect on the N170; neither on its amplitude nor on its latency. These findings were also supported by a pilot study using houses as adaptors, while the experimental task was gender discrimination (Zimmer et al., 2006). These results altogether suggest that N170 reflects mechanisms of shape-selective adaptation acting at the higher-level stages of object processing.

Accordingly, the failure to find any cross-category adaptation in the study would suggest that N170 adaptation effects reflect shape-selective adaptation mechanisms acting in STS in the case of face tests and in EBA in the case of hands.

The analyses of the adaptation effects on ERP responses thus points to the conclusion that N170 might be regarded as the primary correlate of the neuronal adaptation at the higher-level shape-selective stages of visual processing. However, an important requirement for this conclusion is that the adaptation effects found on N170 should originate from the stage of
neural processing where \( N170 \) is generated; additionally, the possibility that the modulation of \( N170 \) amplitude is due to the fact that the input from earlier processing stages is changed as a result of adaptation should be excluded. In the case of adaptation effects on \( N170 \) amplitude this requirement appears to be fulfilled, since no other ERP components, measured over the occipito-parieto-temporal cortical areas, showed shape-selective modulation by adaptation. However, in the case of modulation of response latencies by adaptation, we found significant category-specific adaptation effects already on the earlier, \( P100 \) component of the ERP responses, in addition to the \( N170 \). Therefore, one might suppose that the adaptation effects on \( N170 \) latency are not due to the plastic changes at the processing stages where \( N170 \) is generated but rather are the consequence of adaptation processes acting at earlier levels of visual processing, which are reflected in the \( P100 \) component. However, this possibility seems unlikely, since it would predict a high correlation between the magnitude of adaptation effects on the \( P100 \) latency and those on the \( N170 \) latency, which is not supported by our results. In fact, we found that the adaptation-induced increase of the \( N170 \) latency does not correlate with that of the \( P100 \) latency, suggesting that adaptation effects on \( N170 \) are due to the adaptation mechanisms acting at the processing stages where \( N170 \) is generated.

In summary, the lack of cross-category adaptation effects, both at the perceptual and neuronal level supports the theory of modular, domain specific processing of faces and other objects (Downing et al., 2001; Kanwisher, 2000; Yovel and Kanwisher, 2004, Peelen and Downing, 2005). This theory postulates that processing of faces and human body parts, including hands, takes place on separate, specialized networks of the ventro-lateral temporal cortex. However, it should be noted that there might be an alternative account for our results based on a different view, which assumes that object processing is not domain specific. According to this theory the processing of different object categories might involve different mechanisms depending on how familiar is a given object category and whether we are experts in identifying specific prototypes of this category (Gauthier et al., 1999, 2000). It is assumed that the processing of objects that we have less experience with is based on their structural information whereas the processing of those objects that we developed expertise for – i.e. faces, which are abundant in our visual environment and are of high socio-psychological relevance – is more holistic and based on configural information. As a result of this changed processing, face-like ERP responses are evoked by non-face stimuli as well (Rossion et al., 2002, 2004). Thus, according to this view a possible explanation for the lack of cross-category adaptation effects in our experiments is that humans have more expertise in analyzing and identifying faces than hands. Our behavioural results, in fact, provide some
support for the difference in the expertise for determining specific characteristics based on faces and hands, since observers’ gender judgments were more accurate in the case of faces than in the case of hands. However, it remains to be directly tested whether the explanation based on the difference in expertise for faces and hands can indeed account for the category-specificity of the adaptation effects found in the present study. For this, observers should be trained on the hand gender categorization task to make performance comparable to that in the case of faces. This would serve as a critical test of expertise versus domain-specific processing explanations.

3. **Hemispherical asymmetries in shape-selective after-effects**

On the one hand, we can emphasize the differences in laterality in the perception and processing of face images. For example, in our face distortion after-effect (FDAE) experiment we have found that the $N170$ adaptation effect over the right hemisphere (RH) is larger for distorted ($DIST$) faces, than for the veridical ($VERID$) faces. The adaptors of $DIST$ and $VERID$ conditions are both faces of the same individuals and they differ with respect to the information, relevant to the after-effect (the configuration of the inner features of the faces) only. Thus, the lower $N170$ amplitude during $DIST$ than $VERID$ may reflect the adaptation at the level of detecting facial configurations.

Although our subjects were required to fixate in the centre of the screen, the long adaptation duration makes it possible, that they focused on various facial features, rather than on the configuration of the whole face. However, FDAE is mediated to a large degree by high-level, object-based processing (Watson & Clifford, 2003), that occurs even when subjects are encouraged to move their eyes freely (Morikawa, 2005). Thus we suggest, that the right lateralised adaptation effects on $N170$ are related to the representation of facial feature configuration. The fact that face identity related $N170$ adaptation effects also take place across the whole face (Jacques & Rossion, 2009) supports this conclusion.

However, the fact that we have observed similar $N170$ adaptation, irrespective of the direction of distortion (expansion or contraction), suggests that the $N170$ effect is not specific to the applied distortion.

Current models of face aftereffects (Robbins, McKone & Edwards, 2007) propose that two broadly tuned, overlapping neuronal populations code in an opponent fashion the stimulus space and adaptation causes each pool to decrease its activity in proportion to the strength of the unadapted response to the adapting stimulus, a model originally described for low- and
mid-level features (Regan and Hamstra, 1992). According to this model after adaptation to an expanded face, the response of the neural pool, sensitive to expanded stimuli will be reduced more than that of the neurons, sensitive to contracted faces, shifting the stimulus perceived as most normal toward the expanded direction. The fact, that we have not observed any differences in the N170 effects between +DIST and –DIST adaptation conditions suggests that the activity of the neural pools, sensitive to the direction of distortion, is not reflected in the N170 component.

The fact, that this ERP adaptation effect is lateralised to the RH suggests that there are differences in the face adaptation processes between the two hemispheres. Indeed, current ERP (Kovács et al., 2005, 2007) and imaging (Kovács et al., 2008) studies of face gender adaptation both support the idea of RH advantage in face processing (Rossion et al., 2003). In particular, in recent neuroimaging studies of face adaptation (Kovács et al., 2008) it was found, that the position-invariant components of face adaptation are lateralised towards the RH (in the right fusiform face area). Our results suggest that the identity-specific FDAEs are also based on these neural mechanisms of the RH.

The role of RH is also supported by recent identity adaptation studies. Jacques et al. (2007), using a similar adaptation paradigm, found that the N170 to a repeated compared with an unRepeated face identity is reduced, even when the two unfamiliar faces are presented from different viewpoints (Caharel et al., 2009) or when only half of the face changes in the context of whole faces (Jacques & Rossion, 2009). Previous results suggest that N170 seems to be largely unaffected by familiarity (Anaki, Zion-Golumbic & Bentin, 2007, Eimer, 2000). This and the fact, that face identity adaptation effects are observed on the N170 for unfamiliar faces as well (Jacques et al., 2007) suggests the present N170 adaptation effects are not due to the familiarity of the stimuli.

Whether these mechanisms are face specific or more general, shape coding in nature remains today unclear (for a discussion of this issue see Rhodes et al., 2003). However, in a functional magnetic resonance imaging (fMRI) study (Kovács et al, 2008), using similar parameters to the present study, the dorsal-caudal region of the lateral occipital area (LO), an area activated by complex object shapes (for reviews see Grill-Spector, Kourtzi & Kanwisher, 2001; Malach, Levy & Hasson, 2002), showed no adaptation of the blood oxygen level dependent (BOLD) signal while the occipital and fusiform face areas, known to be involved in the processing of facial features (Gauthier, Skudlarski, Gore & Anderson, 2000, Kanwisher & Yovel, 2006, Pitcher, Walsh, Yovel & Duchaine, 2007; Rossion et al., 2003; Schiltz & Rossion, 2006; Sorger, Goebel, Schlitz & Rossion, 2007) showed significant signal
adaptation. This supports the hypothesis that the face related after-effects reflect, at least partially, the adaptation of face-specific mechanisms.

On the other hand, the adaptation effects on the N170 component are also strongly affected by the relative retinal position of the adaptor and test stimuli. The largest and most consistent difference between the conditions with spatially OL and non-OL adaptor and test stimuli was found in the amplitudes of N170 components recorded over the hemisphere contralateral to the test stimuli. These results suggest that our behavioural facial adaptation effects might be mediated primarily by the adaptation of the shape-selective neural processes in the hemisphere contralateral to the test face stimulus. This conclusion is in line with recent results on visual categorization of natural scenes. It was found that observers are just as fast in determining whether two scenes presented simultaneously on the two sides of fixation contain an animal or not as when only one scene is presented (Rousselet et al., 2002, 2004c). Furthermore, as the ERP results suggest, the two scenes are processed independently in the occipito-temporal visual areas of the corresponding contralateral hemisphere (Rousselet et al., 2002, 2004c). These findings are in correspondence with macaque single cell recording results showing that inferior temporal neurons are strongly driven by the contralateral stimuli with very little influence from the ipsilateral ones (Chelazzi et al., 1998). These, together with the results of the present study strongly support the view that even higher, shape-selective processing of peripherally presented stimuli is strongly lateralized to the contralateral hemisphere and the outcome of visual processing in the contralateral hemisphere will primarily determine the perceptual judgments about the stimulus.

4. The temporal dynamics of facial adaptation effect

The results of the study presented third in my dissertation show that – in agreement with previous results (Kovács et al., 2005) - long-term (5000 ms) face adaptation evokes after-effects consisting of a position-invariant as well as a position-specific component. As a result of adaptation to a female face, test faces were judged more masculine when they were displayed in the same location as the female adaptor face, as compared to that when they were presented in the opposite visual hemifield. However, in conditions when the adaptation time was reduced to 500 ms the resulting face after-effects were found to be entirely position-invariant and no position-specific adaptation effects were observed. In accordance with our behavioural results, the N170 ERP component - recorded over the hemisphere contralateral to the test stimulus - shows position-specific component only in the case of long-term
adaptation: adaptation effects expressed in the *N170* amplitude are larger for SAME than for DIFFERENT adaptor-test stimulus arrangement only in the long-term, but not in the short-term adaptation conditions.

The fact, that the *N170* adaptation effects, recorded for ipsilaterally presented stimuli were always similar for SAME and DIFFERENT replicates our previous data (Kovács et al, 2005) and suggest, that the behavioural facial adaptation effects might be mediated primarily by the adaptation of the shape-selective neural processes in the hemisphere contralateral to the test face stimulus.

Recent fMRI studies suggest that adaptation processes may operate on different timescales in different cortical areas. It was found that obtaining orientation-tuned adaptation signal measurable with fMRI in visual cortical are V1 requires long-term (several seconds) adaptation (Fang et al., 2005), whereas, in the extrastriate cortical areas, adaptation effects were observed both after long-term as well as short-term adaptation (Henson, 2003). These results suggest that duration of adaptation is a critical factor, which can determine whether adaptation effects will be observed in a given cortical area. Moreover, Fang et al. (Fang et al., 2007a) has shown that adaptation duration can also affect the properties of fMRI adaptation effects. Only long term (5000 ms) adaptation effects were found to be selective to the viewpoint of the face, whereas short term (300 ms) adaptation effects were viewpoint independent. The results of the present study are in line with these previous findings, by showing that only the position-invariant face-selective processes are adapted after both short-term and long-term adaptation, whereas adaptation of the neurons responsible for position-specific coding of facial information requires long-term adaptation.

Although, our results do not allow us to determine the exact neural site of the position-specific and position-invariant component of face after-effects, it is important to point out that both components are reflected in the later, *N170* component of the face-evoked ERP responses. Since, the results of several previous studies point to the conclusion that *N170* reflects the structural coding of face stimuli (Bentin et al., 1996; Rossion et al., 1999; Rossion et al., 2000; Itier and Taylor, 2004a) and not the low level processing of their basic visual attributes (contrast, orientation etc), our findings suggest that even the position-specific component of the facial after-effects originate from adaptation processes at the higher, shape-specific stages of visual processing. Indeed, Melcher (Melcher, 2005) measured visual form after-effects across separate fixations and found large transfer of face adaptation, along with spatially non-specific effects, suggesting spatiotopic, rather than retinotopic coding. Furthermore, recent single-unit recording studies have provided evidence that some positional
information is preserved in the inferior temporal cortex of the macaque brain (for a review see Rousselet et al., 2004b). It was found that the receptive fields of inferior temporal neurons can differ in size, have “hot spots”, where they are most sensitive to stimulation, and are typically biased towards the contralateral hemifield (DiCarlo and Maunsell, 2003; OpDeBeek and Vogels, 2000; Rolls et al., 2003). Adaptation of high-level cortical neurons of the human brain, having similar properties could explain that facial after-effects are only partially position-invariant. However, further fMRI investigations are required to localize directly these position-specific and position invariant adaptation effects and to uncover whether they reflect adaptation of different face-selective neural populations of the same or different visual cortical areas.

We found some adaptation effects (enlarged amplitude and elongated latencies) on the earlier, P100 component of the ERP as well using 5000 ms adaptation time. This is in line with our previous findings (Kovács et al, 2006) and with the current literature (Itier and Taylor, 2004a), that suggest that this component might reflect the holistic processing of a face as a face. However, this effect was similar for SAME and DIFFERENT, suggesting that it reflects different mechanisms than the N170 component. After short-term adaptation no effect was observed on the amplitude of P100 while its latency was significantly longer than in the control condition. Such adaptation-related response latency increases have been shown previously for motion adaptation (Hoffman et al., 2001). This suggests that the P100, in addition to the early visual analyses (Gomez et al., 1994, Heinze et al., 1994), to some extent also reflects shape-selective neural processing and indicates that its latency might be sensitive to the state of adaptation at these higher processing stages.

Our results have important implications regarding the recently described face competition effect (Jacques and Rossion, 2004, 2006; Rossion et al., 2004). Jacques & Rossion (2004, 2006) found that neural responses (decreased N170 amplitude) to a face stimulus are reduced when it is displayed in the presence of another face stimulus, which has been on the screen for about 500-700 ms before the presentation of the test face. The main difference between the paradigm used by Jacques and Rossion (Jaques and Rossion, 2004, 2006) and our short-term adaptation condition is that in our experiments the first face stimulus is removed 500 ms prior to the presentation of the second (test) face stimulus, whereas in their experiments it stays on the screen during the presentation of the second (target) face. Since, in the present study we found strong position-invariant short-term adaptation effects on the amplitude of the N170 evoked by the test stimulus, it appears reasonable to suppose that such adaptation effects
might also have contributed to the observed $N170$ amplitude reduction in the experiments by Jacques and Rossion (2004, 2006).

In conclusion, our results suggest that long-term exposure to a face stimulus results in adaptation of a position specific population of face selective neurons in addition to the position invariant neural population, which is adapted both in the case of short and long term adaptation. These findings imply that adaptation duration is a critical factor, which should be taken into account during the interpretation of behavioural as well as neural effects of visual adaptation. Moreover, manipulating adaptation duration provides an opportunity to selectively adapt neural processes at different stages of visual processing and investigate their stimulus selectivity.
3. How can our perception be biased when a stimulus appears \emph{after} the target image? – backward masking results

3.1 Early correlates of face visibility

In the first (and larger) part of my dissertation I discussed our adaptational experimental results. As mentioned before, presenting a stimulus (a mask) \emph{after} the given target changes the perception of the earlier stimulus as well, a phenomenon called backward masking. Both the detection and identification of a stimulus can be impaired by the presentation of a subsequently presented image. The strength of backward masking effect depends on some stimulus properties (see for a review in Breitmeyer, 1984) as the kind of stimuli used in the experiment, the spatial overlap of target and mask, and the stimulus onset asynchrony (SOA, the interval between the onset of the target and the mask). These counterintuitive effects have given rise to considerable philosophical interest (Dennett, 1991), and several theories (Breitmeyer, 1984) have been proposed concerning their origin, ranging from neurophysiological models (Breitmeyer and Ganz, 1976) to more cognitive ones (Michaels and Turvey, 1979). Recent theories of backward masking suggest that the mask abbreviates the effective duration of the target by interrupting its neural processing (for review see Keysers and Perrett, 2002). Since the mask “erases the target from visual awareness” the technique also enjoyed utmost popularity among studies separating neural processes, correlating with visual awareness and unconscious processing (Kim and Blake, 2005). In terms of faces, there appeared surprisingly few studies of backward masking. Recent psychophysical results on face backward masking suggest that discrimination and recognition of facial identity as well as of emotions requires more time than was previously thought, in excess of 100 ms (Loffler et al., 2005b). However, in the last 3 or 4 years the number of backward masking studies using faces as targets is increasing. Here I will summarize the main results of recent psychophysical, psychophysiological and imaging studies.

In two psychophysical experiment Bachmann et al. (Bachmann et al., 2005a, b) have observed the stimulus dependence of masking effect. They explored the role of the relative spatial
positioning of targets and masks, in order to differentiate among different theories of backward masking (as local interaction model, attential model, transient-on-sustained inhibition theory and re-entrance-based substitution theory). In Experiment 1 the authors used stimuli consisting of low-level features (single letters targets – double letters masks). They found that the relative spatial arrangement of the letters, which was changed in order to vary the involvement of metacontrast-like processes, had an effect on shorter SOAs where strong masking effect still persisted. Bachmann et al. concluded that for spatially non-overlapping targets and masks both local metacontrast-like interactions and attential processes are involved in backward masking. In the second Experiment they used more complex visual stimuli, namely human faces as targets and they applied three types of masks – the same faces as targets, faces different from targets and gaussian noise images. Altogether they have found that different masking theories can not solve the problems of masking separately.

Schubö and her colleagues in 2006 (Schubö et al., 2006) also ran a series of experiments consisting of 2 experiments. In experiment 1 they instructed their subjects to detect facial expressions in two different conditions – one where emotional faces (threatening/friendly) were presented in a neutral-faces context and other, where neutral faces were presented in an emotional-faces context. As control they used stimuli which consist of all local features of emotional faces but without the original configurational information. The detection of threatening faces was faster, whereas no differences were found in the control condition. The authors increased the task difficulty by backward masking and got similar results. Interestingly, features were detected faster than faces when the task was hard. This effect did not appear in the threat-detection task.

In a psychophysiological study Tsunoda et al. (Tsunoda et al., 2008) selected from 125 healthy individuals according to the scores of Social Phobia Inventory and defined two groups with high and low social anxiety. Fearful and happy faces were presented subliminally using backward masking during registration of skin conductance responses (SCRs). The authors compared the groups and found that the group with high social anxiety showed greater differences in SCRs between masked fearful and happy faces than normals. Results from early functional neuroimaging studies indicate that the amygdala has a central role in the processing of threat-related facial expression, especially of fearful faces (Morris et al., 1996). Evidence was found that the amygdala responds to fearful faces even when processing them outside conscious awareness (Whalen et al., 1998; Williams et al., 2004). Numerous imaging studies have investigated the activation of amygdala by subliminally presented emotional faces. For example, Suslow et al. (Suslow et al., 2006) in an fMRI
experiment presented happy/angry/fearful and neutral faces in a Latin square design\(^3\) using a backward masking paradigm. The left and right amygdala activations in response to the masked fearful faces were correlated to the number of fearful faces detected. This effect was also observed for angry faces but only in the right amygdala. Killgore and Yurgelun-Todd (2008) have investigated the amygdala activations of adolescents evoked by presenting masked emotional faces. They found that sad faces activated the right amygdala, whereas masked happy faces failed to activate any of the regions of interest. Direct comparisons between masked happy and sad faces were mirrored in the anterior cingulate gyrus. Compared to the adult population adolescents showed higher activity in the right amygdala during the masked sad condition.

However, there are clear evidences that the processing of facial emotions is not located solely in the amygdala (see for a review Pessoa, 2008). Felmingham et al. (2008) examined activation of 23 post-traumatic stress disorder (PTSD) patients to fearful and neutral faces during consciously attended or non-conscious (rendered so via backward masking) conditions. Half of the subjects showed dissociation between conscious and automatic processing of threat stimuli. Dissociative PTSD was associated with enhanced activation in the ventral prefrontal cortex for conscious fear compared to non-dissociative PTSD. Comparatively reduced activation in the dissociative group was apparent in dorsomedial prefrontal regions for conscious fear faces. Pessoa et al. (2006) presented masked fearful faces at 33 or 67 ms durations and observed that, independently of presentation times, when fearful faces were not detected, no activation was found in the amygdala or in the fusiform gyrus (FG). By contrast, when subjects reported seeing a face (independently of real appearance) an increase of activation was observed in amygdala and FG as well.

Due to its excellent temporal resolution, recording event-related potentials (ERP) is a prominent tool for studying the timing of cortical information processing. So far only a few ERP (or magnetoencephalographical - MEG) experiments used backward masking paradigms to study the dynamics of neural processing of words (Holcomb et al., 2005; Kiefer and Spitzer, 2000; Kiefer and Brendel, 2006), simple geometrical figures (Hashimoto et al., 2006; Koivisto et al., 2006), objects (Wilenius-Emet et al., 2004) or natural scenes (Bacon-Macé et al., 2005; Rieger et al., 2005).

\(^3\) A type of statistical experimental design where the aim is to remove the error associated with variations in two non-interacting external variables in order to understand the effect of each variable in addition to the effect of manipulating an independent variable. Experimental units are allocated in such a way that variations in experimental treatments occur once in each row or columns (in a table of treatments).
As for faces, as we have seen before in behavioural and fMRI experiments, ERP studies, using a backward masking paradigm, usually tested the differential processing of fearful and neutral faces. Williams et al. (2004) presented fearful and neutral faces for 10, 30 or 170 ms, followed by a neutral face mask. While the stimuli were above threshold for the longest SOA, the mask rendered the faces below detection or discrimination threshold in the 10 or 30 ms conditions, respectively. The authors found that, relative to neutral faces, perception of fear is related to an enhanced ERP component at around 200 ms on the non-conscious and to an enhanced component at around 400 ms on the conscious levels. Using a similar paradigm with passive viewing Liddell et al. (2004) found that the differential processing of subliminally presented fearful and neutral faces dissociates from supraliminal processing at around 200 ms over central parietal and frontal sites. The aim of these studies was to describe the differences of subliminal and supraliminal fear perception thus both Williams et al. (2004) and Liddell et al. (2004) limited their analysis to the frontal, central and parietal midline electrodes. Consequently, while these studies offer insight into the differential processing of fear at different levels of awareness, they give no information regarding the electrophysiological correlates of earlier processing of faces, that involves ventral occipito-temporal areas (Kanwisher et al., 1997; Sergent and Signoret, 1992) and is reflected mainly in the ERP recorded over lateral occipito-temporal regions (Itier and Taylor, 2004b, 2004c; Rossion et al., 2003b; Schweinberger et al., 2002; Vuilleumier and Pourtois, 2007; Watanabe et al., 2003). Pegna and his colleagues (Pegna et al., 2008) were the first ones who investigated masking effect on emotional faces on the earlier posterior-occipito-temporal components, namely on the N170 and N2. ERPs were recorded during a fearful face detection task in which backward-masked fearful and non-fearful faces were presented ranging from 16 to 266 ms. The authors have found a significantly larger N170 both in the case of subliminal and supraliminal presentation of fearful faces compared to non-fearful faces, however there were no masking effect on this early face-related ERP component. A later negative component, the so-called N2 (~260-300 ms) was the first component related to the stimulus detectability, increasing with target duration. Source localization of the N170 showed that the presentation of a fearful face produce a greater activation of extrastriate visual areas, particularly on the right side. This latter result suggest that fearful faces (both subliminally and supraliminally presented) are processed at an early stage in the visual processing, giving rise to enhanced activation of right extrastriate temporal cortex as early as 170 ms post-stimulus onset. Later, the same group (Genetti et al., 2009) ran a familiar face detection experiment with similar experimental design. They got a large masking effect on N2 and
P300. The authors have found that the N2 shows a pattern of increase similar to the measure of visual awareness. Their source analysis suggests a distributed network of brain areas including frontal and temporo-occipital regions. The network will be activated parallel to the conscious process, and this conscious parallel processing begins from about 200 ms. However, in both of the latter mentioned experiments authors manipulated the visibility of the target images by varying the exposition time of them and in their experiments the mask followed the target immediately. However, different strength of backward masking can also be produced by holding the presentation time of the target fixed, whereas manipulating the stimulus onset asynchrony and/or the type of the mask.

The aim of the present study was to study the time course of face processing by recording ERPs in a facial backward masking paradigm. As seen before, one of the simplest way of altering the visibility of a target stimulus in a backward masking paradigm is to decrease the SOA. However, recording ERPs in such paradigms suffers from the fact that not only the target but also the mask evokes a brain response. Moreover, depending on the applied SOA, the interaction of the responses evoked by target and mask varies, causing the interpretation of the resulting complex response troublesome. To overcome this problem, in each of our conditions, we extracted the potentials specifically evoked by face targets (face) by subtracting the ERPs evoked by control, random noise target images (control). Additionally, recent theories hypothesize (Thorpe, xxxx) that the ERPs to the target and the mask sum linearly, which might not be the core. Furthermore, since recent studies revealed that the magnitude of masking depends on the similarity between target face and mask (Loffler et al., 2005b), we compared the ERP effects of a noise (FOUR), a 180 deg inverted face (INV) and an upright face (FACE) mask at the same SOA.

**Materials and Methods**

**Subjects:** Eleven naïve, healthy volunteers (2 female and 9 male) participated in the study (age: 20 - 25 years, mean 22 years). They all had normal or corrected-to-normal vision, no previous history of any neurological or ophthalmologic diseases and were not under medication. The procedures were approved by the Ethical Committee of the Budapest University of Technology and Economics.

**Stimuli:** Face stimuli were upright, grey-scale full-front digital images of six young females and six young males (chosen from a larger face database of our laboratory). Faces depicted 100% fearful and neutral expressions. They were fit behind an oval mask (fit into a square of 400 x 400 pixels, 7.3 deg) eliminating the outer contours of the faces. The mean luminance was 1.17 cd/m². Control stimuli were the Fourier phase randomized versions of the face stimuli, for details see Chapter 2.1. We used three different masks: (1) an upright neutral female face...
(FACE), (2) the same face but 180 deg inverted (INV) and (3) the Fourier randomized version of the original adaptor face (FOUR). To avoid possible apparent motion when replacing the stimulus with the mask and to maximize the masking effect, the mask images occupied approximately 50% larger area than the stimulus (600 x 600 pixels, 10.95 deg).

**Task, Procedure:** Stimuli were presented centrally (on a 19” monitor, 1024 x 768 pixel resolution, 85 Hz vertical refresh rate; with a viewing distance of 70 cm) on a uniform grey background (luminance 1.3 cd/m²). All software was written in MATLAB 6.5 (Mathworks Inc.) using Psychotoolbox 2.45 for Windows. Subjects were tested in a dimly lit room (average background luminance <1 cd/m²). They were instructed to fixate a spot in the centre of the monitor and to perform a two-alternative forced choice expression discrimination task by pressing the left mouse button when the stimulus was perceived as fearful and right button for neutral faces. To keep attentional level similar for the face and control stimuli subjects were not aware of the presence of control stimuli, they merely were informed that the stimuli, due to different SOAs and different levels of noise in the images, will be “more or less visible”.

In each trial (Figure 3.1), after a random interval (randomized between 500 and 700 ms), the target was presented for 23.5 ms, followed by one of three stimulus onset asynchrony values (SOA, 23.5, 117.5 and 470.5 ms) and finally one of the masks (FACE, INV or FOUR) was presented for 500 ms. To reduce inter-individual differences, 24 hours before the ERP recordings, we ran a pilot psychophysical test for each subject, in which we varied the SOA between 23.5, 58.8, 117.5 and 470.5 ms (10 repetition each SOA) in the same facial expression discrimination task with the upright face mask. In the subsequent ERP experiment we only included those subjects whose performance in the initial test was above chance level for the 23.5, 117.5 and 470.5 ms SOAs. Trials with different SOAs, target- and mask-types were intermixed randomly within one block. Each target stimulus was presented 20 times for each SOA within a block. Altogether ERPs were recorded for about 30 minutes. During the experiments subjects were asked to refrain from movements and blinking.

Electrophysiological recordings: ERPs were recorded via 25 Ag/AgCl electrodes placed according to the 10/20 system with the left earlobe as ground and F5 as the reference lead. Impedances were kept below 10 kOhm. The sampling rate was 1024 Hz. EEG was segmented offline (using BrainVision Analyzer (Brain Products GmbH; Munich, Germany), into 1100 ms long trials, using a 100 ms prestimulus interval. Trials containing blinks,
movements, A/D saturation or EEG baseline drift were rejected on the basis of visual inspection of each recording by semi-automatic artifact detection. ERPs were averaged separately for each subject, condition and channel. Averages were then digitally filtered (0.1-70 Hz) with a zero phase shift digital filter and average referenced.

**Psychophysical and ERP data analyses:** For analyzing the behavioral data we entered it into a three-way repeated measures analysis of variance (ANOVA) with SOA (3), mask-type (3, FACE, INV, FOUR) and emotion (2, Fearful, Neutral) as within subject factors for both stimulus types (face or control) separately (emotion effect were not analyzed for control stimuli). D-prime values were also calculated for each SOA and mask. We restricted our ERP analyses to occipito-temporal electrodes, corresponding to P7/P8, PO7/PO8 and P9/P10, selected on the basis of face-selective effects in previous studies of ourselves (Kovács et al., 2005, 2006) and others (Eimer, 2000; Rossion et al., 2000). In electrophysiological experiments of backward masking the biggest confound is that both target and mask stimuli evoke responses from the processing neurons (Kovács et al., 1995). In order to isolate the activity associated with the target faces we did the following analysis (Bacon-Mace et al., 2005; for a demonstration of the technique see later Figure 3.4, a): we subtracted the ERPs obtained with control targets from the ERPs obtained with fearful and neutral face targets for each subject, channel and SOA in the FACE, INV and FOUR mask conditions separately. Since both control and face conditions contain information about the target (a face or a noise image) and the mask (which was identical for the two targets), we could eliminate any confounding mask response from the ERPs. These differential activity curves (dERPs) are considered to reflect the emotional face vs. non-face (control) differences and therefore are used for any further analysis. To determine the onset of masking effect and the major dERP components, affected by backward-masking, for each hemisphere we made a point by point paired t-test on the differential curves, obtained in each SOA and mask-type condition (p<0.05 for at least 30 consecutive points) for fearful and neutral faces, across subjects and electrodes. Finally, within the previously determined time-periods, the individual amplitude of the main peak of the dERP was measured, for each subject and each condition using a ± 30 ms window, centered on the maximum of the grand-average dERP. For testing the effect of SOA we used only the data, obtained in the FACE mask condition by entering the amplitude values of the dERP components into a three-way ANOVA with SOA (3), emotion (2) and hemisphere (2) as within subject factors. The effect of different masks was only tested at the shortest SOA by a similar three-way ANOVA with mask-type (3), emotion (2) and hemisphere (2) as within subject factors. All analyses involved Greenhouse-Geisser adjusted degrees of freedom for correction for non-sphericity. Post-hoc t-statistics were performed by Fisher LSD tests. The correlations between the dERP data and behavioral data were performed with Pearson tests.

**Results**

**Behavioural results**

Upright female face masks resulted in a strong masking effect: the performance of the subjects was 60% ± 5.7% for fearful faces and 53.9% ± 6.6% for neutral faces for the shortest SOA, respectively and increased for the other SOAs (main effect of SOA: F(4,72)=3.28, p=0.043, Figure 3.2) in our FACE condition.
Not surprisingly, the detection of fearful faces was easier for subjects (main effect of EMOTION: $F(4,72)=11.67$, $p=0.001$). Post-hoc analyses revealed that for the shortest applied SOA the performance was significantly better for the FOUR condition than for FACE ($p=0.024$) and the performance was better for the longest SOA compared to the shortest one ($p=0.0056$). Interestingly, the pattern of performance is significantly different for fearful and neutral faces in the 117.5 ms SOA condition ($p=0.00456$). Both FOUR and INV masks affected differently performance on fearful and neutral faces ($p=0.0216$ and $p=0.0388$, respectively), but it was not true for FACE mask ($p=0.147$).

However, as one can expect for stimuli containing no shape-related information (as our control images did), subjects’ performance was at chance level for each SOA and mask-type (50.5% ± 1.5%).
We also calculated $d'$ values (see Figure 3.3) according to the various SOAs and different mask-types. We can see that the masking effect of the FOUR mask is independent from the applied SOA, whereas we can observe tendencies of SOA dependence in the case of FACE and INV masks – by increasing the delay between the target and the mask the performance of the subjects is better, they produce higher $d'$-prime values (main effect of SOA by a two-way ANOVA, $F(4,72)=2.3355$, n.s., post-hoc effects: 23.5 ms SOA vs 470.5 ms SOA, $p=0.034$, FACE-masked 23.5 ms SOA vs 470.5 ms SOA, $p=0.042$, tendencies: at the shortest SOA – FOUR vs FACE, $p=0.076$, INV vs FACE, $p=0.088$, at FACE mask – 117.5 ms SOA vs 470.5 ms SOA, $p=0.087$).

Figure 3.3: Mean emotion discrimination rate. The average $d'$ values computed from each subjects’ hit and false alarm rate. Left, middle and right panels represent the 23.5 ms, 117.5 ms and 470.5 ms SOA, respectively. Black bar: FOUR mask, middle gray bar: INV mask, light gray bar: FACE mask.

Electrophysiological results

Since data analyses are in preparation and we observed behavioural masking effects (i.e. SOA dependence) only in the case of neutral faces, in my dissertation I will only discuss data for neutral faces.

As mentioned before I will concentrate on two different cases – on the effects of different SOAs at FACE mask and on the effects of different masks at the shortest SOA. First we will focus on the earlier face-related components, as the P100 and the N170, than we discuss the results of point-by-point t-tests and the statistical analyses of later dERP components.
Masking effects on the earlier face-related ERP components

First, we concentrated on the masking effect mirrored on the earlier face-related ERP components, namely on the P100 and the N170.
Masking effect on the amplitude of N170

Interestingly, the amplitude of the face-related N170 component also shows masking effect: the amplitude is larger in the conditions leading to worse performance. Significant effect was observed both in the different masks at the shortest SOA conditions (main effect of masks: F(2,114)=19.48, p<0.000001) and in the various SOAs at the application of FACE mask (main effect of SOAs: F(2,116)=11.22, p=0.000035) (see Figure 3.5 a).

Masking effect on the latency of N170

At 23.5 ms SOA the effect of different masks was statistically significant (the main effect of various masks: F(2,114)=8.48, p=0.00037). We observed a delayed N170 in the INV masked condition. No significant laterality effects were found. We also got significant differences in the different SOA conditions at FACE mask condition (main effect of SOAs: F(2,116)=13.29, p=0.000006). We also observed a face×side interaction (F(2,116)=6.16, p=0.0029) (see Figure 3.5 b).
**Masking effect on the early P100 component**

No statistically significant masking effects were observed in the amplitude and latency values of the P100 component at the parieto-occipito-temporal sites (for amplitudes – main effect of masks: $F(2,34)=3.05$, $p=0.06$, n.s.; main effect of SOAs: $F(2,34)=1.647$, $p=0.208$, n.s. and for latencies - main effect of masks: $F(2,34)=0.31$, $p=0.74$, n.s.; main effect of SOAs: $F(2,34)=0.365$, $p=0.697$, n.s.).

**Correlation of the earlier dERP components and the behavioural performance**

By a subject-by-subject analysis we also correlated the behavioural performance (namely the relative frequency of hits for neutral faces) and the amplitude of N170 and P100 components. Surprisingly, the behavioural masking effect showed significant positive correlation (in other words an inverse effect) to the amplitude of the N170 both in the left and the right hemisphere ($r=0.4985$, $p=0.00023$ and $r=0.3692$, $p=0.0083$ for the left and right side, respectively), however no significant correlation values were observed in the case of P100.

**Masking effect on the dERP due to shorter SOAs**

Our primary question concerns how backward masking affects face processing. Since the dERP activity curves reflect the differential processing of the target faces, independently of the mask, here we only present the analysis of the dERPs. In the conditions, in which emotion discrimination performance was above chance level (e.g. FACE masked 470.5 ms SOA or FOUR masked 23.5 ms) they are dominated by a positive wave and two subsequent large negative waves at around 100, 150 and 230 ms, respectively. It is worth to note, that two large positive deflections are also observable on the dERPs at around 320 and 680 ms in the 117.5 and 470.5 ms SOA conditions, respectively, corresponding to 180-185 ms post mask-stimulus onset in both cases (Figure 3.4 b and c). These negative dERP waves suggest that the underlying ERPs, obtained for the FACE mask were more negative in the face than in the control condition. In other words the face target led to small adaptation effects, when compared to a noise image (control): the N170 evoked by the mask was reduced. This is in accordance with our previous results, showing that N170 ERP component is sensitive to adaptation (Kovács et al., 2005, 2006).

Comparing dERPs obtained at different SOAs revealed (Figure 3.4 b) that the FACE mask had a significant effect on the residual potentials: both dERPs, obtained with 117.5 and 470.5
ms SOAs were different from that of 23.5 ms SOA. After 117.5 ms SOA the onset of the difference was 192 ms and 182 ms for the left and right side, and remained significant until 284 ms and 286 ms, respectively. We also got a later time period in which the two above-mentioned conditions differed significantly – namely the 315 – 388 ms period for the left and the 304 – 382 ms period for the right side. After 470.5 ms SOA the onset of the difference was delayed compared to the further results (230 ms for the left and 196ms for the right side) and remained significant until 296 ms on the left and 397 ms on the right side.

Figure 3.6: Mean of the amplitudes (a and b) and latencies (c and d) of P2 and N2 components recorded for the three left and three right occipito-temporal channels in the 23.5 ms SOA condition with different masks (black, dark grey and light grey bars for the FOUR, INV and FACE mask, respectively) and in FACE masked conditions with different SOAs (light grey, red and green bars for 23.5 ms, 117.5 ms and 470.5 ms SOA, respectively).
A closer inspection of dERP reveals that a smaller and more positive (or rather less negative) (P2) and a large negative (N2) components are present in this time period over the occipito-temporal, occurring at around 208 ms and 246 ms, respectively. The amplitude of the P2 component shows significant differences in the SOA conditions at FACE mask (Figure 3.6 a; main effect of SOAs: F(2,108)=81.42, p<0.000001). This effect was also mirrored in the latency of the P2 (main effect of SOAs: F(2,108)=33.26, p<0.000001 and p values of the post-hoc tests: 23.5 ms vs 117.5 ms: p<0.000001 and 23.5 ms vs 470.5 ms: p<0.000001, see Figure 3.6 c). The amplitude of the N2 component also showed a dramatic SOA-effect (Figure 3.6 b; main effect of SOAs: F(2,92)=199.11, p<0.000001). The effect of different SOAs also mirrored in the latency data of this component (main effect of SOAs: F(2,92)=3.35, p=0.039, see Figure 3.6 d).

**Effects of different masks at 23.5 ms SOA**

The effect of mask-type on the dERP was tested at the shortest SOA, where the largest behavioral differences were observed (see Figure 3.4 c). In accordance with the perceptual differences of the masks, the dERP obtained in FOUR was more negative than in the INV and this difference occurred well after the first large negative wave, being significant between 276 – 322 ms on the left and 268-321 ms on the right side, as suggested by a point by point t-test, while the differences between FOUR and FACE did not reach the level of significance.

The amplitude of the P2 component shows significant differences in the different FACE masks conditions at 23.5 ms SOA (Figure 3.6 a; main effect of masks: F(2,90)=13.27, p=0.000009). This effect was also mirrored in the latency of the P2 (main effect of masks: F(2,90)=3.795, p=0.026 and p values of the post-hoc tests: FOUR vs FACE: p=0.033 and INV vs FACE: p=0.0083, see Figure 3.6 c). The N2 component also showed a dramatic mask-effect (Figure 3.6 b; main effect of masks: F(2,76)=10.4473, p=0.000098) and this component was lateralized: the amplitude of the N2 was larger (namely more negative) on the right side (main effect of side: F(2,76)=5.14, p=0.029). The effect of different masks also manifest in the latency data of this component (main effect of masks: F(2,76)=13.465, p=0.00001, see Figure 3.6 d).
Correlation of the later dERP components and the behavioral performance

By a subject-by-subject analysis we correlated the behavioural performance and the amplitude of the latter measured ERP components, namely the P2 and the N2. The behavioral masking effect showed significant correlation to the amplitude of the positive dERP component around 195 - 216 ms over the right hemisphere ($r=0.33$, $p<0.05$), while no such correlation was found over the left hemisphere. In other words the better the discriminability of the target face, the more positive the dERP was over the right hemisphere at the P2/N2 time window. The N2 component did not show any significant correlation with the performance of the individuals.

Summary of the results

The results of the present experiment can be summarized as follows:

- The better the performance the smaller the N170 amplitude
- Decreasing the SOA to a value where expression discrimination was hardly possible reduced the amplitude of a face-specific negative differential ERP component, peaking around 240 ms after stimulus onset.
- Inverted face and noise image masks evoked smaller behavioural masking effects and in the later case the amplitude of this differential ERP component was also reduced.
3.2 Discussion of our backward masking results

The aim of the present study was to investigate the electrophysiological correlates of facial backward masking.

*The early components of the dERP are independent of face visibility*

A typical face evoked ERP, recorded over occipito-temporal sites, consists of a large positive component, occurring at around 100 ms ($P100$) and a usually even larger negative component at around 150-200 ms ($N170$; Bentin et al., 1996; Eimer, 2000; Jeffreys, 1996; Rossion et al., 1999). These peaks show face specificity to different degrees. Although the $P100$ ERP component that is thought to be generated in early extrastriate visual areas (Linkenkaer-Hansen et al., 1998), is considered to reflect the processing of low-level visual features, it has been found recently to be larger for faces than for other objects and it was suggested that it might reflect the holistic processing of a face as a face (Itier and Taylor, 2004a). The earliest signal of the ERP, that is widely considered as specific to faces, is the so-called $N170$ (Bentin et al. 1996; Rossion et al. 1999; Itier and Taylor, 2004a), being always larger and earlier for faces than for other objects (e.g. Itier and Taylor, 2004b, Rossion et al., 2000). This component is considered to encode configural information of faces (Boutsen et al., 2006; Rossion et al., 2003b).

The $d$ERP curves of our study reflect the differential processing of face targets, compared to noise images with the same amplitude spectrum. Accordingly, similarly to the above described face related ERPs the early part of $d$ERPs is dominated by a positive wave and a subsequent large negative wave at around 100 and 150 ms, respectively. However, in our present experiments no masking effect was evident in the $P100$ time window. This implies that the processing steps, reflected by $P100$ are independent of the visibility of the target faces. Previously it has been suggested that both $P100$ and $N170$ reflect early, perceptual processing, involving local and global configural information (Boutsen et al., 2006; Itier and Taylor, 2004b; Rossion et al., 2003). Our present results support the idea, that the steps of $P100$, reflecting the processing of a face as a face, are relatively automatic, and take place independently of awareness (Maurer et al., 2002).
However, we observed an inverse relationship of N170 amplitude with visibility: it was larger after FACE mask than FOUR mask at 23.5 ms and than 470.5 ms FACE mask. This finding can be explained by two possible ways.

One explanation is that in the case of 23.5 ms FACE masked condition the responses evoked by the face target and FACE mask stimuli integrate in a non-linear fashion. This integration has to be non-linear since we extracted the Fourier target FACE masked condition from the ERPs. This integration of the stimulus and mask evoked ERP would be larger for the face – FACE 23.5 ms condition than for the Fourier - FACE 23.5 ms as well as the face - FACE 470.5 ms conditions. However, such an explanation is implausible, given the significant correlation of the observed changes of the \( \text{N170} \) to the behavioural performance of the subjects: we found, that the better the performance the smaller the N170 dERP amplitude. Indeed, Carlson et al. (Carlson et al., 2008) found that it is currently unclear, whether the effect of backward-masked fearful faces are due to the fearful nature of the face or perceptual inconsistencies between the emotional target and the neutral mask. Carlson and Reinke (2008) have shown that masked fearful modulate the orienting of spatial attention and it appears that this effect is due to the fearful nature of the face rather than perceptual inconsistencies between the initial faces and the masks.

Such attentional mechanisms could explain the observed changes of the N170 component. However, this explanation requires further studies to prove.

Bacon-Mace et al. (2005) recorded ERPs for complex scenes while subject’s had to detect the presence of animals in the pictures. Authors found that the amplitude of the differential ERP (derived from the subtraction of target and distractor scenes) at the first large negative wave, around 200 ms at occipito-temporal sites increases rapidly with SOA. Contrary to this, our dERPs show an “inverse masking effects” in the time window, corresponding to N170. This suggests that the differential masking effects of dERPs for human faces and for natural scenes, observed between 150-200 ms, reflect different sensory processes. Indeed, we know that non-face stimuli also evoke a negative ERP component in the 150-200 ms time-window, the N1. However, the amplitude of the face evoked N170 is larger and its latency is shorter as compared to the other objects evoked N1 (Rossion et al., 1999; Itier and Taylor, 2004a, b). Further studies analyzed the scalp topographies of N1/N170 evoked by faces and non-face object categories and revealed an extra map for the face N170 in the lateral temporal cortex which was not present in the case of N1 to objects (Itier and Taylor, 2004a). Comparison of
the results of Bacon-Mace et al. (2005) study and of ours raises the possibility that whereas the N1 evoked by non-face images is susceptible to masking effects, the face evoked N170 is not.

**The later dERP components are strongly affected by backward masking**

The effect of masking on the dERPs started at around 200 ms. On typical face related ERPs of the occipito-temporal cortex usually a small negative wave is observable between 250-300 ms post-stimulus time. This wave is enlarged for repeated (Begleiter et al., 1995, Schweinberger et al., 1995, 2002, 2004) or familiar faces (Herzmann et al., 2004; Itier and Taylor, 2004; Tanaka et al., 2005) and effects of task relevance (Martens et al., 2006) or attention (Lueschow et al., 2004) are also apparent in this time window. These data suggest that this component is related to higher face processing steps (Schweinberger et al., 2004). Our present results suggest that variations of stimulus visibility are also manifest in this time window. The fact that our masking related dERP amplitude reduction was correlated to behavioral performance only over the right hemisphere supports the right hemispheric dominance in face perception (Kanwisher et al., 1997; Sergent and Signoret, 1992).

Recording from central sites (Fz, Cz and Pz) Williams et al. (2004b) and Liddell et al. (2004), also revealed, that a positive component, peaking around 250-300 ms (N2) was significantly more negative for fearful than for neutral faces at subliminal SOAs, a difference absent at longer SOAs. This negative differential ERP component at centro-frontal sites might correspond to the more positive dERP curve of the present study at bilateral occipito-temporal sites after 23.5 ms SOA.

This time window corresponds well to the first ERP and MEG correlates of visibility for objects as well. In a series of studies (Koivisto et al., 2006; Vanni et al., 1996; Wilenius-Emet et al., 2004) authors compared the ERP/MEG signals, obtained for simple or complex line drawings when presented below or above detection threshold in a backward masking paradigm. They found differences of the electric or magnetic signals between 258-297 ms that correlate with the visibility of the object. Since the time window of this “visual awareness negativity” matches well our visibility effects, we propose that non-face objects and faces are processed in a similar fashion after 200-250 ms post-stimulus onset. This conclusion is supported by single-cell recordings from the inferior temporal cortex of the macaque brain, suggesting that faces (Rolls et al., 1999) and non-face shapes (Kovács et al., 1995) are similarly susceptible to backward masking.
The relatively slow processing of faces is supported further by recent psychophysical results of facial backward masking. Loffler et al. (2005b), using synthetic faces and a recognition task, found that face discrimination requires time in excess of 100-130 ms. Our results also show, that a SOA of 117.5 ms leads to above-threshold performance in an emotion discrimination task. Furthermore, while dERPs of this SOA are also significantly different from dERPs obtained with sub-threshold SOA, additional increase of the SOA (up to 470.5 ms) does not change the dERP further.

Finally, our results are also in agreement with previous findings that the magnitude of masking depends on the configural similarity between target and mask: in addition to providing support to the behavioral results of Loffler et al. (2005b), we also show that this similarity effect is related, to a certain extent, to the amplitude of the dERP amplitudes, recorded over the occipito-temporal sites: the more similar (and thus the more effective) the mask is to the target face, the smaller the dERP amplitude. Furthermore, the fact that upright and inverted face masks led to similar changes of the dERP suggests that the neuronal correlates of facial masking are insensitive to face-inversion. In behavioral experiments face-inversion invariably leads to a decrease in recognition performance and this effect is specific to faces (for a review see Rossion and Gauthier, 2002). However, a recent fMRI study (Yovel and Kanwisher, 2005) revealed, that not every face selective area is susceptible to the face inversion effect equally: while the activity of the fusiform face area and the superior temporal sulcus decreased for inverted faces, the occipital face selective area showed similar activation patterns for upright and inverted faces. Here we suggest that the neural mechanisms responsible for the electrophysiological masking effects are less sensitive to the inversion of faces.

Overall, our results suggest that regarding its susceptibility to backward masking, processing of facial emotions can be separated into two different stages. The first part of the processing, lasting up to 200 ms post-stimulus time, is related inversely to the visibility of the visual stimulus. From approximately 200 ms onwards, however, stimulus processing becomes strongly modulated by visibility and correlates with the conscious perception of faces.
4. General discussion

Many theorists have suggested that faces are represented mentally as a unique vector in a multidimensional face-space. The length and the direction of this vector show the deviation from the norm (or prototype). Since our external world changes rapidly and continuously, the representation of the world (as well as that of faces) is also an actively changing system in the human brain that integrates the inputs from one moment to another. The results of our experiments confirm the model that faces are coded in relation to face norms or prototypes, which are constantly calibrated to reflect the properties of the faces that surround us. However, it is not only the continuous changes of the external world that can alter the representation. Our results suggest that our previous and/or consecutive experiences, as well as artificial modulation of the neuronal activity can bias our perception and thereby the representation of a given object can be altered.

In five psychophysical, transcranial Direct Current Stimulation and electrophysiological experiments we studied high-level configural after-effects, evoked by adaptation. Principally we examined facial adaptation effects at the level of gender processing. In another experiment we tested the behavioural and electrophysiological correlates of face distortion after-effect. In all cases we found robust and long lasting effects in the performance of subjects. These effects were also reflected in the electrophysiological signals – we were the first to show adaptation effects on a very early ERP component, namely on the N170.

Our electrophysiological and transcranial Direct Current experiments suggest that shape-selective after-effects are the results of higher-level processing and not the results of the summation of lower level effects since their orientation-, position-, and size-invariance suggest that higher-level, shape-selective visual areas are involved in these after-effects.

After presenting human faces and hands as adaptors, we found that gender decision to the opponent category (i.e. if adapted to a face the target was a human hand), was not shifted compared to the control condition. Supporting the behavioural data the amplitude and latency of N170 did not show adaptation effect for different category adaptor and test stimuli. Summarizing the above it can be concluded that adaptation effects are category-specific – no cross-categorical effects were found.

It is widely known that the right hemisphere plays a crucial role in face perception. In our face distortion after-effect experiment we found that the amplitude of the N170 component decreased only in the right hemisphere. This effect was stronger in the case of distorted
adaptors compared to veridical ones. Right hemisphere dominance is also supported by our data adapting with overlapping and non-overlapping adaptors. The strongest adaptation effect in the amplitude of the N170 component could be observed when the adaptor was presented on the left side while registration was contralateral, namely on the right hemisphere.

By manipulating the duration of the adaptation we could also dissociate the position-specific and position-invariant elements that can be observed both in the performance and in the changes of the N170 component. We described that long term face adaptation (5000 ms) contains position variant and invariant components. Using a female adaptor androgenous faces were judged more masculine by participants if the adaptor and the test were presented in the same retinal position than when presented in non-overlapping positions. In case of short term adaptation (500 ms) we found position invariance in both overlapping and non-overlapping trials. In accordance with our behavioural data the amplitude of the N170 component has a position specific effect only after long term adaptation. These results suggest that both long and short term adaptation trigger the position-invariant face-selective mechanisms of adaptation. In contrast position-specific adaptation requires long adaptation time. Consequently, the manipulation of the adaptation time is a useful tool for selectively examining shape-specific adaptation and its stimulus selectivity.

Presenting a stimulus after the given target changes the perception of the earlier stimulus as well. For example the detection and identification of a stimulus can be impaired by the presentation of another stimulus (a mask). It is well known that this masking also occurs when the mask is not presented before but after the target. Several types of this backward masking exist (see also the introduction of Chapter 3.1), depending on the kind of stimuli used (light flashes versus patterns), the spatial overlap of target and mask, and the stimulus onset asynchrony (SOA) (the function relating the temporal interval separating target and mask onset).

Whereas in Chapter 2 we studied the adaptation mechanisms underlying high-level configural after-effects in human electrophysiological and psychophysical experiments, in Chapter 3 we tapped into the neuronal mechanisms of this phenomenon of backward masking.

The results of our backward masking experiment suggest that emotional face processing has two stages. The first step of processing - approximately the first 200 ms after stimulus presentation – is independent from the stimulus visibility. After 200 ms stimulus processing is strongly influenced by the stimulus becoming conscious. 220 ms after stimulus presenation
the amplitude of the negative ERP component (N2) (registered from the face processing areas) correlates with the conscious perception of the stimulus.
5. Acknowledgements

Most experiments of this dissertation were carried out at the Department of Cognitive Sciences, Budapest University of Technology and Economics, Budapest, Hungary. Many colleagues and friends provided great encouragement and help as I have prepared this work.

First of all, I would like to thank my supervisor, Dr. Gyula Kovács, who supported me since my first steps in cognitive neuroscience and who introduced me to the mysteries of human psychophysics and electrophysiology. It was from him that I learned not only the theoretical and practical background of my PhD topic but hopefully also such a research attitude that is maybe the most important thing to become a mature researcher.

Special thanks to Dr. Andrea Antal, who introduced me to the tDCS technique and during seven months I spent at the Department of Clinical Neurophysiology, Georg-August University, Göttingen, Germany she held fit my mind and my body.

I also wish to express my gratitude to Prof. Zoltán Vidnyánszky for the numerous enjoyable discussions.

Particular thanks must go to my collaborators, who are working and worked with me on many experiments. With Éva Bankó, Irén Harza, Patrícia Gerván, Mónika Albu, Anett Nagy, Krisztina Nagy we spent many exciting hours in the dark electrophysiological lab. I am also very grateful to my colleagues who introduced me to other cognitive research topics, as contour integration, perceptual learning, visual attention or memory inhibition: Prof. Ilona Kovács, Dave Melcher and Dr. Mihály Racsmány.

I owe special thanks to Rozália Ivády for her great assistance and friendly words, for correcting my manuscript and improving my English.

I also thank to my friends who always gave me a hand when I needed it: Edit Anna Garab, Gabriella Mari, Klára Boros, Csaba Poreisz, Petra Karácsinyi, Gábor Csifcsák, Eszter Tóth and Gábor Kocsis.

Last but not least I owe special thanks to my family for the emotional help and for the endless patience during my work on this dissertation.
6. Publications related to the thesis points


7. References


