Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. (2003) Mapping cortical change across the human life span. *Nat Neurosci* 6:309–315.

Spear PD. (1993) Neural bases of visual deficits during aging. *Vision Res* 33:2589–2609.

Spekreijse H, Van der Twell LH, Zuidema T. (1973) Contrast evoked responses in man. *Vision Res* 13:1577–1601.

Takahashi T. (2002) *Photosensitive epilepsy*. Igaku-Shoin, Tokyo.

Tobimatsu S, Zhang YM, Tomoda Y, Mitsudome A, Kato M. (1999) Chromatic sensitive epilepsy: a variant of photosensitive epilepsy. *Ann Neurol* 45:790–793.

Wilkins AJ, Darby CE, Binnie CD, Stefansson SB, Jeavons PM, Harding GFA. (1979) Television epilepsy-the role of pattern. *Electroencephalogr Clin Neurophysiol* 47:163–171.

ELSEVIER

Clinical Neurophysiology 119 (2008) 1803–1811

www.elsevier.com/locate/clinph

Electrophysiological evidence for sequential discrimination of positive and negative facial expressions

T. Nakashima[a,f,*], Y. Goto[a,d], T. Abe[b], K. Kaneko[c], T. Saito[a], A. Makinouchi[e], S. Tobimatsu[a]

[a] *Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Fukuoka 812-8582, Japan*

[b] *Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan*

[c] *Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan*

[d] *Department of Occupational Therapy, Faculty of Rehabilitation, International University of Health and Welfare, Okawa, Japan*

[e] *Department of Information and Network Engineering, Kurume Institute of Technology, Kurume, Japan*

[f] *Japan Society for the Promotion of Science (JSPS), Tokyo, Japan*

Accepted 17 April 2008

Available online 5 June 2008

ELSEVIER

Clinical Neurophysiology 119 (2008) 1803–1811

www.elsevier.com/locate/clinph

Electrophysiological evidence for sequential discrimination of positive and negative facial expressions

T. Nakashima[a,f,*], Y. Goto[a,d], T. Abe[b], K. Kaneko[c], T. Saito[a], A. Makinouchi[e], S. Tobimatsu[a]

[a] *Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Fukuoka 812-8582, Japan*

[b] *Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan*

[c] *Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan*

[d] *Department of Occupational Therapy, Faculty of Rehabilitation, International University of Health and Welfare, Okawa, Japan*

[e] *Department of Information and Network Engineering, Kurume Institute of Technology, Kurume, Japan*

[f] *Japan Society for the Promotion of Science (JSPS), Tokyo, Japan*

Accepted 17 April 2008
Available online 5 June 2008

Abstract

Objective: To elucidate the differences in temporal processing between positive and negative facial expressions by using event-related potentials (ERPs) with spatially filtered images.

Methods: Based on the traits of parallel visual pathways, four types of facial expression images (happiness, fear, anger and neutral) with low, high and broadband spatial frequencies (LSF, HSF and BSF, respectively) were carefully created with the consideration of luminance, contrast and emotional intensity. These images were pseudo-randomly presented to 13 healthy subjects to record ERPs. Twenty recording electrodes were placed over the scalp according to the International 10–20 system. For emotion-relevant late negative components with latencies of 190–390 ms, the amplitude differences among the four facial expressions were analyzed for sequential 20-ms time windows by ANOVA.

Results: There were significant amplitude differences between positive and negative LSF facial expressions in the early time windows of 270–310 ms at the occipitotemporal region. Subsequently, the amplitudes among negative HSF facial expressions differed significantly in the later time windows of 330–390 ms.

Conclusions: Discrimination between positive and negative facial expressions precedes discrimination among different negative expressions in a sequential manner based on parallel visual channels.

Significance: ERPs with spatially filtered images have provided the first evidence for sequential discrimination of positive and negative facial expressions.

Keywords: Event-related potentials; Positive and negative facial expressions; Parallel visual channels; High and low spatial frequencies

1. Introduction

Facial expression recognition is one of the most important skills in social communications. In everyday life, we can readily identify individuals, as well as the age and gender of faces, without difficulty. Humans are even able to guess someone's emotions from their facial expressions at a particular moment. However, patients with autism spectrum disorders and schizophrenia often show behavioral patterns that differ from the usual patterns for these cognitive activities, and thus have a disadvantage in daily social communications (Elger and Campbell, 2001). Therefore, elucidating the brain mechanisms involved in facial expression recognition may lead to the development of therapeutic interventions.

* Corresponding author. Address: Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine, Graduate School of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Fukuoka 812-8582, Japan. Tel.: +81 92 642 5543; fax: +81 92 642 5545.

E-mail address: taisuke@med.kyushu-u.ac.jp (T. Nakashima).

1388-2457/$34.00
doi:10.1016/j.clinph.2008.04.014

A number of studies have indicated that the amygdala is engaged in processing facial expressions, especially fearful expressions (Morris et al., 1998; Adolphs, 2002; Posamentier and Abdi, 2003). Several ERP studies have reported that the late components (180–750 ms) are related to facial expression processing (Streit et al., 2000; Sato et al., 2001; Krolak-Salmon et al., 2001; Eimer and Holmes, 2002; Balconi and Pozzoli, 2003; Schupp et al., 2004; Ashley et al., 2004). However, the scalp distributions and peak latencies of these slow late components are inconsistent among the studies. Moreover, direct functional correlations between late ERP components and the amygdala can be difficult to estimate. Therefore, the neural basis for the spatiotemporal characteristics of the late ERP components for discriminating various facial expressions has not yet been clarified.

A recent functional MRI study (Vuilleumier et al., 2003) revealed that low spatial frequency (LSF) images of a fearful expression activated the amygdala to a much greater extent than unfiltered images (broadband spatial frequency, BSF). In contrast, high spatial frequency (HSF) images of a neutral face mainly activated the fusiform gyrus compared with LSF faces. These observations are supported in part by the physiological traits of two parallel visual pathways. Specifically, the magnocellular (M) stream mainly analyzes coarse (or LSF) information and sends rapid visual signals to the amygdala via the tecto-pulvinar pathway (Schiller et al., 1979; Liddell et al., 2005), while slower visual signals with fine spatial (or HSF) information are chiefly projected to ventral cortical areas such as the fusiform gyrus via the parvocellular (P) stream (Livingstone and Hubel, 1988; Tobimatsu and Celesia, 2006). Since the study by Vuilleumier et al. (2003), several investigators have recorded ERPs to determine the effects of SF on face or facial expression perception by measuring the early P1 and face-sensitive N170 (Goffaux et al., 2003; Pourtois et al., 2005; Holmes et al., 2005; Halit et al., 2006). It has been suggested that LSF information plays an important role in the recognition of faces or facial expressions (especially fearful faces) in the early P1 and N170 (Goffaux et al., 2003; Pourtois et al., 2005; Halit et al., 2006). However, the importance of HSF information for M170, a counterpart of N170, was reported in a magnetoencephalographic study (Hsiao et al., 2005). It is also noteworthy that N170 was not affected by SF information (Holmes et al., 2005). To date, no reports have demonstrated how the late components behave differently and process HSF and LSF information of facial expressions.

The discrepant results for the P1 and N170 components among previous ERP studies may arise from methodological differences and/or differences between the SF cut-off values used to create the filtered faces. Briefly, the subjects' task-relevant attention to faces can be a critical factor that affects the ERP components, especially the late components (Krolak-Salmon et al., 2001). Nevertheless, face-relevant attentive tasks, such as the gender decision task and face categorization task, have often been employed (Goffaux et al., 2003; Pourtois et al., 2005; Halit et al., 2006). The definitions of LSF and HSF are quite different among previous studies, ranging from <5 to 8 cycles/face for LSF and >15 to 32 cycles/face for HSF (Goffaux et al., 2003; Pourtois et al., 2005; Holmes et al., 2005; Hsiao et al., 2005; Halit et al., 2006). These cut-off values did not take into account whether the subjects could precisely recognize the facial effects for each facial expression. In addition, the physical characteristics of the filtered photographs, such as luminance and contrast, were not strictly controlled.

On the other hand, psychological developmental studies have revealed that infants prefer looking at happy faces compared with angry or neutral faces (LaBarbera et al., 1976) and vice versa (Nelson and Dolgin, 1985). In addition, infants can distinguish happy faces from fearful faces (Nelson et al., 1979). These findings suggest that recognition of happiness and fear is essential for our early life. Infants have low vision compared with adults and rely more heavily on LSF information. Furthermore, the judgment speeds for happiness are faster than those for negative facial expressions in adults (Kirouac and Dore, 1983). Therefore, we hypothesized that discrimination between positive and negative expressions first occurs in the late components, even in healthy adults, via LSF information. Subsequently, more detailed analyses among other expressions occur via HSF information.

To test this hypothesis, we recorded ERPs to positive and negative expressions (anger, fear, happiness and neutral) using BSF, LSF and HSF facial expressions created by image engineering. Prior to the ERP study, we determined the psychophysical thresholds (cut-off values for SF) of facial expression recognition to equalize the subjective intensity for each expression (Fig. 1). Thus, our filtered face stimuli were uniquely created with consideration of the luminance contrast of the faces and facial expressions, and the physical factors and facial expression intensities were equated. We also employed a simple passive viewing task with a target stimulus (shoes) to maintain the subjects' vigilance and attentional direction away from the facial expressions. Using such a careful approach, we predicted that the late components would show characteristic changes in the amplitudes or latencies of the late components (e.g., an amplitude increase and a latency decrease) for each filtered positive and negative facial expression, and that such alterations would provide new insights into the neural mechanisms of social cognition in humans.

2. Methods

2.1. Subjects

The study participants were 13 healthy undergraduate students with normal or corrected to normal vision (6 females and 7 males; age range, 21–25 years; mean age, 22.4 years) who were naive to our experimental stimuli. They did not have any history of neurological or psychiatric disorders. The subjects were paid for their participation and signed informed consent after the nature of the exper-

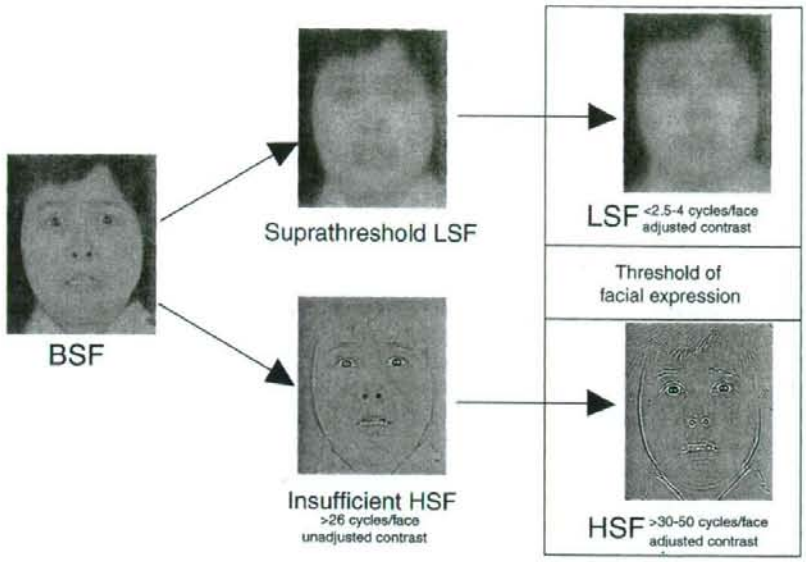

Fig. 1. Representative examples of the fearful stimuli used in the present study. BSF is an original non-filtered image (left), which contains broad spatial frequency components. LSF and HSF faces are filtered by fast Fourier transformation. The middle column shows inadequate stimuli. Insufficient HSF is caused by using a lower cut-off point and not adjusting the contrast, such that the HSF component of the image is not fully enhanced. Suprathreshold LSF is too easy to recognize and similar to the BSF condition, such that the effects of LSF on the ERPs are not disclosed by these stimuli. Therefore, these stimuli are not suitable for our ERP experiments. The right column shows our preferred stimuli. The threshold for recognizing facial expressions was measured by psychophysical experiments in advance, and the contrast of the images was adequately normalized. The LSF face consists of information with low spatial frequencies (<2.5–4 cycles/face width) and preserves the holistic facial image. The HSF face extracts information with high spatial frequencies (>30–50 cycles/face width) and emphasizes the detailed features of the facial components.

iments had been fully explained. The local Ethics Committee of Kyushu University approved the study.

2.2. Original stimuli

The original stimuli were 256-level grayscale photographs of emotional (angry, fearful and happy) and neutral faces taken from Japanese and Caucasian Facial Expressions of Emotion (JACFEE) and Neutral Faces (JACNeuF), respectively (Matsumoto and Ekman, 1988). Although JACFEE and JACNeuF consisted of four Japanese photos for each expression, we only used two photographs for each facial expression (1 male and 1 female) because the other photographs were not appropriately evaluated as portraying the intended emotion by our subjects in preliminary experiments.

To control the subjects' attention and vigilance, a target stimulus (shoes) was taken from our own 256-level grayscale photographs.

2.3. Manipulation of stimuli

LSF and HSF stimuli for faces and houses were created by image engineering techniques with two-dimensional fast Fourier transformation (one-order Gaussian window method for LSF; 35-order Hamming window method for HSF) using our own program written in C language and MATLAB ver. 7 (The MathWorks Inc.). The BSF stimuli were original photographs and left unfiltered. The mean cut-off frequencies (<2.5–4.0 cycles/face for LSF; >30.0–50.0 cycles/face for HSF) were determined by measuring the psychophysical thresholds for recognition of facial expressions using 30 other recruited subjects (10 females and 20 males; age range, 20–34 years; mean age, 25.7 years; unpublished data) prior to the ERP recordings (Table 1). We did not allow these subjects to participate in the ERP experiments because they were habituated by the repeated exposures of the filtered facial expressions and a repetition effect of upright faces on the brain activity at ~250 ms has been reported (Schweinberger et al., 2007).

In this experiment, we presented 20 levels of filtered photographs in descending series (LSF: from 1.0 to 17.0 cycles/face width; HSF: from 6.0 to 126.0 cycles/face width) and asked the subjects to choose the appropriate emotion from six basic emotions. The recognition thresholds for each category were determined as the first images when the subjects consecutively responded with successful answers. In general, none of the subjects recognized differences in the SF among the facial expressions despite the difference in threshold frequency for each facial expression condition. Although the cut-off values showed slight variability among the different facial expressions, they were approximately the same as those in a previous study using spatially filtered images (Vuilleumier et al., 2003). The images of all the faces were trimmed

Table 1
Mean (±SEM) recognition thresholds for each expression stimuli (cycles/face width) ($n = 30$)

SF	Gender	Anger	Fear	Happiness	Neutral
LSF	Male	3.8 ± 1.0	4.9 ± 1.7	2.4 ± 0.2	5.0 ± 0.9
	Female	4.4 ± 1.8	3.7 ± 0.7	2.5 ± 0.3	5.3 ± 1.0
HSF	Male	38.4 ± 4.6	30.8 ± 2.5	50.0 ± 3.7	38.4 ± 3.2
	Female	32.8 ± 3.7	36.9 ± 4.7	50.8 ± 4.1	41.2 ± 3.1

into rectangles in accordance with the face width, and adjusted to 512 × 614 pixels. If the contrast of HSF images was not adjusted, these images became relatively low-contrast stimuli because the original power of the HSF information of the natural images was low (Fig. 1, lower middle) and the evoked responses would therefore be reduced. Our HSF stimuli were strictly adjusted to ensure that the LSF and HSF images had the same extents of contrast (Fig. 1, right column). We also used high-order filtering (35 orders) to enhance the HSF information and exclude the LSF information more clearly. In other words, the high-order window method helped us to extract the HSF components more strictly. The mean luminance and contrast were controlled by normalizing the mean and standard deviation (SD) of the gray values of all the stimuli using our own program written in C language (mean luminance, 48 cd/m^2; mean gray value ±SD, 128 ± 40). Representative examples of the stimuli (fearful expression) are shown in Fig. 1. Moreover, images of gray value 128 (1024 × 768) were added to all the stimuli as a background. Representative examples of the other facial expressions under each SF condition are shown in Fig. 2.

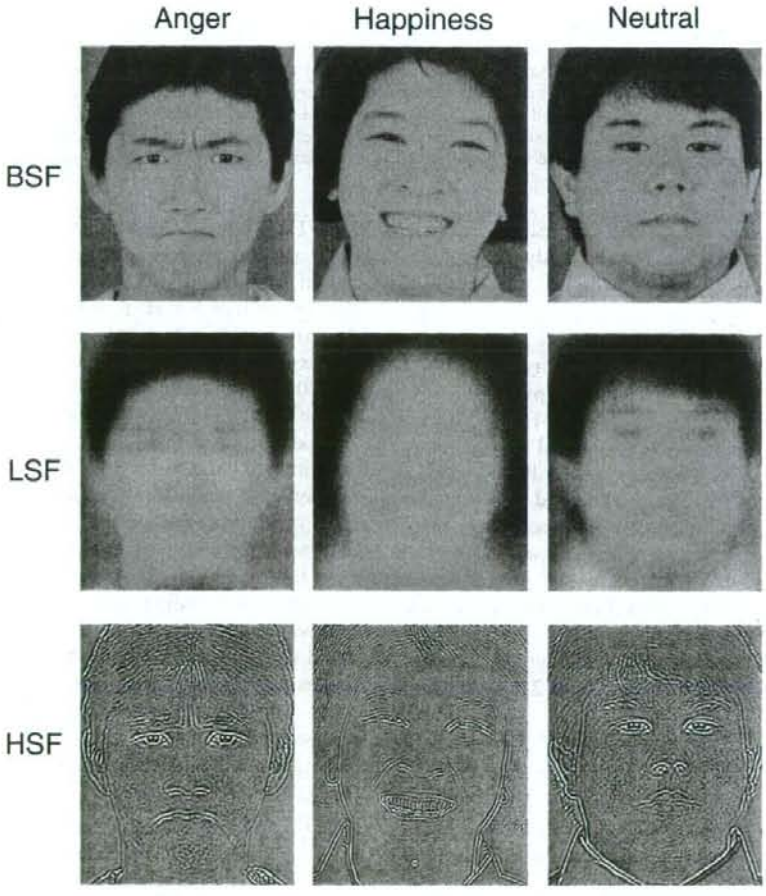

Fig. 2. Representative examples of the angry, happy and neutral stimuli used in the present study. The details for the BSF, LSF and HSF images are described in the legend to Fig. 1.

2.4. Task

The subjects viewed all the stimuli presented on a 17-in. computer screen (Sony Trinitron GT200; refresh rate, 100 Hz; 1024 × 768 pixels) in a dark room. Each facial emotion, except for the target stimulus (shoes), was presented pseudo-randomly under each SF condition for 300 ms with an inter-stimulus interval (ISI) of 700 ms on the computer screen at a viewing distance of 114 cm (visual angle, 8.3 × 10.0 degrees) under the control of VSG 2/5 (Cambridge Research System Co.). The subjects were instructed to fixate on a point in the center of the screen and push a button when a target stimulus (shoes, 10% probability) appeared but ignore other stimuli. The target stimulus was used for controlling the attention and maintaining the vigilance of the subjects. The visual stimuli were replaced by the background image (1024 × 768 pixels) of gray value 128 during the ISI. The reaction time (RT) and accuracy for the target stimulus were also measured while recording the ERPs.

2.5. ERP recording

Silver–silver chloride electrodes were applied to 20 scalp sites on the basis of the International 10–20 system (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, Pz, P3, P4, T5, T6, Oz, O1 and O2) with collodion. All the recording electrodes were referred to an electrode at the nose-tip, and the impedance was kept below 5 kΩ. Vertical and horizontal electrooculograms (V-EOG and H-EOG, respectively) were also recorded in order to reject artifacts caused by blinking and eye movements. A trigger pulse was generated at the onset of each stimulus. A total of 220 responses were averaged for each stimulus with a band-pass filter between 0.05 and 200 Hz. For visualization of the late components, a 70-Hz high-cut filter was applied. The sampling rate was 1333 Hz, and the recording time for each stimulus was 768 ms. A baseline was set from −72 ms to the onset of the stimuli (0 ms). Artifact rejections were conducted manually on the basis of deflections above 30 μV from the baseline. As a result, the mean rejection rate was 30.8%.

After the ERP recordings, the subjects participated in an unstructured interview with one of the authors (TN) and were asked to state as many as possible of the kinds of facial images they recognized. However, we did not measure the psychophysical thresholds for the recognition of each facial stimulus or the difference in valence between positive and negative expressions before the experiment in order to avoid the repetition effect of upright faces on the brain activity at ~250 ms (Schweinberger et al., 2007) and the task-relevant attention to faces (Krolak-Salmon et al., 2001).

2.6. Data analysis and statistics

The amplitudes of the late components were calculated by averaging the values of 20-ms time windows from 190 to 390 ms at the occipitotemporal regions (T5, T6), since previous studies detected different ERP activation patterns between facial identification and facial expression recognition tasks with unfiltered photographs at these electrodes (Bentin et al., 1996; Sato et al., 2001; Krolak-Salmon et al., 2001).

Statistical analysis was conducted by four-way repeated-measures analysis of variance (ANOVA) (spatial frequency (BSF, LSF or HSF) × 20-ms time windows (190–210, 210–230, 230–250, 250–270, 270–290, 290–310, 310–330, 330–350, 350–370 or 370–390 ms) × laterality (T5 or T6) × facial expression (anger, fear, happiness or neutral)) with the Huynh–Feldt correction for degrees of freedom. As a post hoc analysis, multiple paired comparisons (Bonferroni) were conducted for each time window under each SF condition to focus on the effects of SF on facial expressions in the late components in each time window. All statistical analyses were performed using SPSS 11.01 (SPSS Inc.).

3. Results

3.1. Preliminary psychological experiment

The mean (±SEM) SF thresholds for recognition of facial expressions are shown in Table 1. The mean cut-off frequency for each filtered facial expression was used as the visual stimulus for the ERP recordings in the separate group.

3.2. Psychological evaluation after ERP recordings

All the subjects reported that they perceived at least two kinds of negative expressions (i.e., anger and fear), a positive expression (happy) and neutral faces for all of the filtered conditions.

3.3. Performance of target detection during ERP recordings

The mean (±SD) reaction time for the target stimulus (shoes) was 386.8 ± 31.5 ms, and the correct detection rate (±SD) was 99.9 ± 0.1%. The P300 ERP component was recorded in all subjects with this high performance. N170, a hallmark of face identification, was predominantly recorded over the occipitotemporal sites (T5 and T6) (Fig. 3). However, we did not find any significant effects of SF on this component. The high detection rate indicates that the subjects were alert and that their attention was well controlled during the ERP recordings.

3.4. Late components differentiate positive–negative and negative–negative expressions

Late components were found at the T5' and T6 electrodes. Fig. 4 shows enlarged waveforms of the late components for LSF, BSF and HSF faces at T5 and T6 for the time window of 200–450 ms. Significant main effects of SF, time window and laterality were observed

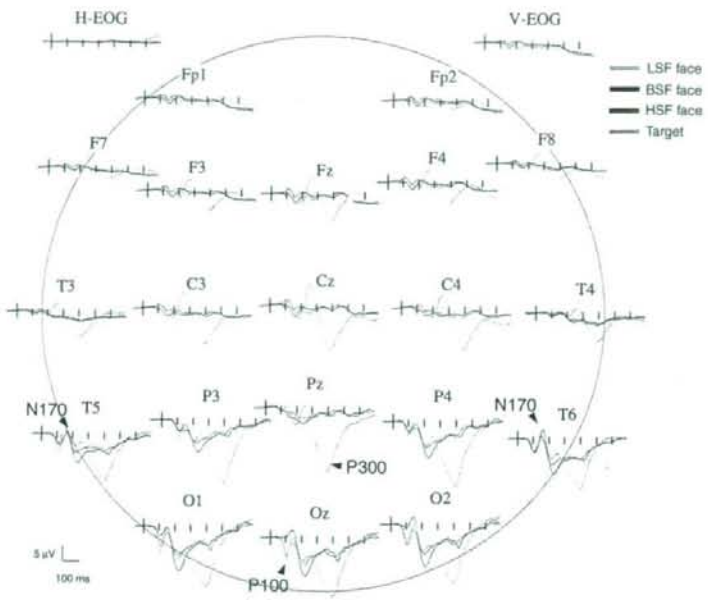

Fig. 3. Topographical mapping of grand-averaged ERPs to the face and target stimuli. Note that P300 is only clearly recorded for the target stimulus.

($F(1.615, 19.375) = 4.501$, $p < 0.05$, $F(2.500, 30.004) = 7.253$, $p < 0.005$ and $F(1,12) = 8.513$, $p < 0.05$, respectively). A significant three-way interaction of SF × time window × facial expression was also found ($F(15.401, 184.812) = 5.121$, $p < 0.001$). The colored boxes in Fig. 3 illustrate significant amplitude differences among facial expressions following the post hoc test of the three-way interaction. The LSF happy faces induced significant negative shifts of the potential compared to negative expressions (anger, fear) in the early time windows (270–290 and 290–310 ms) ($p < 0.05$). In contrast, the HSF fearful faces elicited significant negative shifts of the potential compared to angry faces in the later time windows (330–350, 350–370 and 370–390 ms) ($p < 0.05$). No significant amplitude differences were observed for neutral faces (Fig. 4). It is important to note that we did not find any significant effects for facial expressions under the BSF condition.

In summary, LSF images induced different responses between positive and negative expressions (anger/fear vs. happiness) in the relatively early phase of the late components, while HSF images elicited significant differences in the amplitudes among different negative expressions (anger vs. fear) in the late phase of the late components.

4. Discussion

4.1. Importance of SF for discrimination between positive and negative expressions

The present study examined the influence of SF filtering on brain processing of affective facial photographs. To manipulate the SF characteristics of the stimuli, the photographs were presented without filtering (BSF), after low-pass filtering to retain only LSF components (<2.5–4.0 cycles/face width) or after high-pass filtering to retain only HSF components (>30.0–50.0 cycles/face width). As a result, different effects of LSF and HSF information on the late components were clearly demonstrated, while BSF had no significant effects on the late components. These results suggest that modulation of the SF information is a critical factor that influences the behaviors of the late components.

The amplitudes within 270–310 ms (N270–310) differed significantly among happiness and fear/anger in the LSF condition, while the amplitudes within 330–390 ms (N330–390) differed significantly between anger and fear in the HSF condition. These findings suggest that discrimination of a positive expression from a negative expression occurs in the early phase of the late components, while discrimination among negative expressions is achieved in a much later phase of the late components. These interpretations are supported by a behavioral study in which the latency of judgment of a happy expression was earliest, while those of angry and fearful expressions were last (Kirouac and Dore, 1983). The early components (from P1 to 150 ms) were significantly modulated by LSF fearful faces (Pourtois et al., 2005), which may be related to early detection of a threat. However, we did not find any effects of SF on facial expressions in these time windows. Therefore, our late components may not reflect the first detection of emotional content but rather the processing of detailed emotional contents.

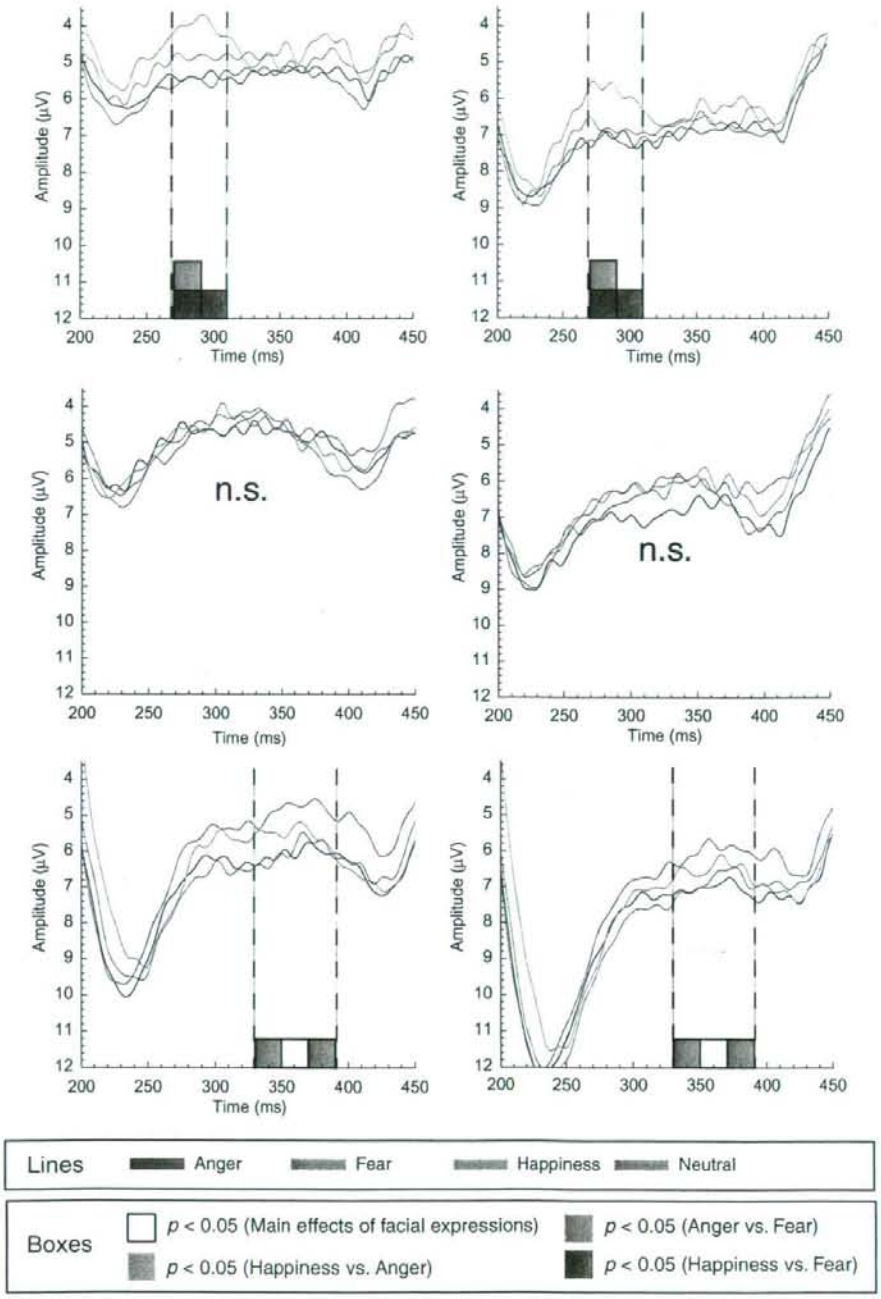

Fig. 4. Waveforms of the late components for the facial expressions of each stimulus. The preceding N170 is omitted while the time window of 200–450 ms is enlarged for comparison. The white square boxes on the abscissa indicate main effects of facial expressions, while the colored boxes show statistically significant differences revealed by paired comparisons (Bonferroni correction). Under the LSF condition, there are significant differences in the amplitudes between positive (happiness) and negative (anger or fear) expressions during the time window of 270–310 ms, regardless of the hemisphere. In contrast, a significant difference is found among negative expressions (anger vs. fear) during the time window of 330–390 ms under the HSF condition.

4.2. Insignificant contribution of unfiltered images to the late components

Although naturalistic (BSF) images of faces provide both LSF and HSF information simultaneously, there were no significant differences in the late components among the facial expressions under the BSF condition. In the present study, the attention of the subjects was directed toward the target stimulus (shoes), and the influence of attention to the facial ERPs was therefore relatively small compared with direct attention tasks to facial stimuli. Krolak-Salmon et al. (2001) reported that there were no significant differences in the amplitudes among facial expressions when the subjects did not pay direct attention to the facial expressions themselves. Similarly, Vuilleumier et al. (2003) did not find any significant effects for BSF stimuli in a gender identity task using functional MRI. In addition, we determined the psychophysical thresholds (cut-off values for SF) of facial expression recognition in a separate group of subjects to equalize the subjective intensity for each facial expression, since the psychophysical threshold of mosaic faces was found to have a significant effect on face recognition in our previous study (Goto et al., 2005). Overall, the expression intensity of our stimuli and attention had no significant influences on the late components under the BSF condition. Therefore, only the SF information is an important factor that affected the late components.

4.3. Neural basis for sequential processing of facial expressions

Our results provide evidence that processing of facial expressions is enhanced by SF modulation and processed sequentially over time by using LSF and HSF information in conjunction. We hypothesize that enhancing either the LSF or HSF components of faces can alert the M and P pathways contingently, which in turn facilitate information processing in the human brain to discriminate among facial expressions. Indeed, brain structures sensitive to emotionally relevant information such as the amygdala, superior colliculus and pulvinar, part of a phylogenetically old route specialized for rapid processing of fear-related stimuli, appear to be preferentially activated by LSF information, but not HSF information, via the fast M pathway (Vuilleumier et al., 2003; Winston et al., 2003; Liddell et al., 2005; Eimer and Holmes, 2007). In contrast, the slow P pathway, which is more responsive to HSF stimuli, provides input to the ventral visual cortex such as the fusiform gyrus, but not to subcortical areas, and is crucial for detailed processing of the shape of the visual object (Livingstone and Hubel, 1988; Tobimatsu and Celesia, 2006). Therefore, it is plausible that the N270–310 component uses LSF information via the M pathway and rapidly segregates positive and negative expressions. Subsequently, the N330–390 component utilizes HSF information via the P pathway and differentiates among negative expressions in more detail. Interestingly, Goren and Wilson (2006) recently demonstrated that recognition of happiness was robust even if happy faces were presented in the peripheral vision. Since the M pathway handles information processing from the peripheral visual field, their finding is consistent with our result that the behavior of the N270–310 component under the LSF condition reflects a discriminative stage between positive and negative expressions.

To date, the neural generators of the late components have remained controversial. Although it is possible that many cortical structures contribute to the late components observed in our study, especially for happy expressions, previous reports had suggested that the visual cortices are activated by the distant influence of the amygdala when subjects recognize facial expressions (Morris et al., 1998; Vuilleumier et al., 2004). In addition, Amaral et al. (2003) proved the existence of direct connections from the amygdala to the visual cortices. Therefore, it is possible that the positive shifts in the early late components for LSF fearful faces compared to LSF happy faces may result from a re-entrance projection from the amygdala or various other cortico-subcortical regions related to emotional processing (Sato et al., 2001).

4.4. Methodological considerations

Care was taken to create the spatially filtered images in this study. However, because we did not measure the psychophysical data for each stimulus before the ERP experiment, which subjects actually recognized the facial valence (negative, positive) on a single trial basis might be debatable. We sought to avoid the repetition effects of facial expressions on ERPs (Schweinberger et al., 2007); however, such effects can arise while measuring ERPs. Future research needs to include concurrent measures of psychophysical data and ERP recordings in the same subjects. This should provide additional information on SF/valence interaction effects in the late components.

5. Conclusions

We have found for the first time that positive and negative facial expressions are processed sequentially over time based on the traits of parallel visual pathways. Specifically, the early late component within 270–310 ms reflects the discrimination between positive and negative expressions depending on LSF information, while the subsequent late component within 330–390 ms reflects the discrimination among different negative expressions based on HSF information.

Acknowledgements

We thank Professor G.G. Celesia (Department of Neurology, Loyola University Medical Center, Maywood, IL) for his critical comments regarding this manuscript. We thank Professor H. Sakamoto (Department of Visual Com-

munication Design, Faculty of Design, Kyushu University, Fukuoka, Japan) for helpful advice regarding the image processing using MATLAB. We also thank Dr. N. Kinukawa (Department of Medical Informatics, Kyushu University, Fukuoka, Japan) for advice regarding the statistical analysis.

This study was supported in part by The 2004 Satow's Research Fund for Behavioral Science, Kyushu University Foundation, Grants-in-Aid for Scientists (Nos. 16390253 and 16200005) from the Ministry of Education, Culture, Sports, Science and Technology in Japan and Grants-in-Aid for the 21st Century COE Program and Kyushu University COE program.

References

Adolphs R. Neural systems for recognizing emotion. Curr Opin Neurobiol 2002;12:169–77.

Amaral DG, Behniea H, Kelly JL. Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. Neuroscience 2003;118:1099–120.

Ashley V, Vuilleumier P, Swick D. Time course and specificity of event-related potentials to emotional expressions. Neuroreport 2004;15:211–6.

Balconi M, Pozzoli U. Face-selective processing and the effect of pleasant and unpleasant emotional expressions on ERP correlates. Int J Psychophysiol 2003;49:67–74.

Bentin S, Allison T, Pace A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. Cogn Neurosci 1996;8:551–65.

Eimer M, Holmes A. An ERP study on the time course of emotional face processing. Neuroreport 2002;13:427–31.

Eimer M, Holmes A. Event-related brain potential correlates of emotional face processing. Neuropsychologia 2007;45:15–31.

Elger K, Campbell R. Annotation: the cognitive neuroscience of face recognition: implications for developmental disorders. J Child Psychol Psychiat 2001;42:705–17.

Goffaux V, Gauthier I, Rossion B. Spatial scale contribution to early visual differences between face and object processing. Brain Res Cogn Brain Res 2003;16:416–24.

Goren D, Wilson HR. Quantifying facial expression recognition across viewing conditions. Vis Res 2006;46:1253–62.

Goto Y, Kinoe H, Nakashima T, Tobimatsu S. Familiarity facilitates the corticocortical processing of face perception. Neuroreport 2005;16:1329–34.

Halit H, de Hann M, Schyns PG, Johnson MH. Is high-spatial frequency information used in the early stages of face detection? Brain Res 2006;117:154–61.

Holmes A, Winston JS, Eimer M. The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. Cogn Brain Res 2005;25:508–20.

Hsiao FJ, Hsieh JC, Lin YY, Chang Y. The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. Neurosci Lett 2005;380:54–9.

Kirouac G, Dore FY. Accuracy and latency of judgment of facial expressions of emotions. Percept Mot Skills 1983;57:683–6.

Krolak-Salmon P, Fischer C, Vighetto A, Mauguière F. Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. Eur J Neurosci 2001;13:987–94.

LaBarbera D, Izard E, Vietze P, Parisi A. Four-and six-month-old infants' visual responses to joy, anger and neutral expressions. Child Dev 1976;46:535–8.

Liddell BJ, Brown KJ, Kemp AH, Barton MJ, Das P, Peduto A, et al. A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. Neuroimage 2005;24:235–43.

Livingstone M, Hubel D. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 1988;240:740–9.

Matsumoto D, Ekman D. 1988. Japanese and Caucasian Facial Expression of Emotion (JACFEE) and Neutral Faces (JACNeuF) [slides]. Human Interaction Laboratory, University of California, San Francisco, 401 Parmassus Avenue, San Francisco, CA, 94143.

Morris JS, Friston KJ, Buchel C, Frith CD, Young AW, Calder AJ, et al. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 1998;121:47–57.

Nelson A, Dolgin K. The generalised discrimination of facial expressions by 7-month-old infants. Child Dev 1985;56:58–61.

Nelson A, Morese A, Leavitt A. Recognition of facial expressions by 7-month-old infants. Child Dev 1979;50:1239–42.

Posamentier MT, Abdi H. Processing faces and facial expressions. Neuropsychol Rev 2003;13:113–43.

Pourtois G, Dan ES, Grandjean D, Sander D, Vuilleumier P. Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. Hum Brain Mapp 2005;26:65–79.

Sato W, Kochiyama T, Yoshikawa S, Matsumura M. Emotional expression boosts early visual processing of the face: ERP recording and its decomposition by independent component analysis. Neuroreport 2001;12:709–14.

Schiller PH, Malpeli JG, Schein SJ. Composition of geniculostriate input to superior colliculus of the rhesus monkey. J Neurophysiol 1979;42:1124–33.

Schupp HT, Ohman A, Junghofer M, Weike AI, Stockburger J, Hamm AO. The facilitated processing of threatening faces: an ERP analysis. Emotion 2004;4:189–200.

Schweinberger SR, Kaufmann JM, Moratti S, Keil A, Burton AM. Brain responses to repetitions of human and animal faces, inverted faces, and objects – an MEG study. Brain Res 2007;1184:226–33.

Streit M, Wolwer W, Brinkmeyer J, Ihl R, Gaebel W. Electrophysiological correlates of emotional and structural face processing in humans. Neurosci Lett 2000;278:13–6.

Tobimatsu S, Celesia GG. Studies of human visual pathophysiology with visual evoked potentials. Clin Neurophysiol 2006;117:1414–33.

Vuilleumier P, Armony JL, Driver J, Dolan RJ. Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nat Neurosci 2003;6:624–31.

Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat Neurosci 2004;7:1271–8.

Winston JS, Vuilleumier P, Dolan RJ. Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. Curr Biol 2003;13:1824–9.

Neuroscience Research 62 (2008) 225–235

Contents lists available at ScienceDirect

Neuroscience Research

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

Early ERP components differentially extract facial features: Evidence for spatial frequency-and-contrast detectors

Taisuke Nakashima [a,f,*], Kunihiko Kaneko [b], Yoshinobu Goto [a,d], Tomotaka Abe [c], Takako Mitsudo [a,f], Katsuya Ogata [a], Akifumi Makinouchi [e], Shozo Tobimatsu [a]

[a] Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
[b] Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan
[c] Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan
[d] Department of Occupational Therapy, Faculty of Rehabilitation, International University of Health and Welfare, Okawa, Japan
[e] Department of Information and Network Engineering, Kurume Institute of Technology, Kurume, Japan
[f] Japan Society for the Promotion of Science (JSPS), Tokyo, Japan

ARTICLE INFO

Article history:
Received 1 May 2008
Received in revised form 15 August 2008
Accepted 19 August 2008
Available online 2 September 2008

Keywords:
Spatial frequency
Equiluminance
Contrast
P1
N170
N2

ABSTRACT

It is generally accepted that the N170 component of an event-related potential (ERP) reflects the structural encoding of faces and is specialized for face processing. Recent neuroimaging and ERP studies have demonstrated that spatial frequency is a crucial factor for face recognition. To clarify which early ERP components reflect either coarse (low spatial frequency, LSF) or fine (high spatial frequency, HSF) processing of faces, we recorded ERPs induced by manipulated face stimuli. By filtering the original grayscale faces (broadband spatial frequency) spatially, we created LSF and HSF face stimuli. Next, we created physically equiluminant (PEL) face stimuli to eliminate the effects of lower order information, such as luminance and contrast. The P1 amplitude at the occipital region was augmented by LSF faces, while the N170 amplitude increased for HSF faces. The occipital P1 amplitude for PEL faces was relatively unaffected compared with that for PEL houses. In addition, the occipital N2 for PEL faces was spatiotemporally separable from N170 in a time-window between P1 and N170. These results indicate that P1 reflects coarse processing of faces, and that the face robustness further assures face-specific processing in the early component. Moreover, N2 reflects the early contrast processing of faces whereas N170 analyzes the fine facial features. Our findings suggest the presence of spatial frequency-and-contrast detectors for face processing.

1. Introduction

Over the past decade, a number of electrophysiological and neuroimaging studies have stressed the specialized visual processing of faces, since the discovery of the intracranial N200 (Allison et al., 1994) and the fusiform face area (Kanwisher et al., 1997; Haxby et al., 2000). Electrophysiological studies have revealed face-selective event-related potential (ERP) components from the scalp, such as the vertex-positive potential (Jeffreys and Tukmachi, 1992) and N170 at the temporo-occipital region (Bentin et al., 1996). In particular, N170 has been considered to reflect a structural encoding stage in the psychological model of face processing (Bruce and Young, 1986; Bentin et al., 1996; Eimer, 2000a). However, it remains unclear whether N170 reflects all the aspects of facial encoding. Itier and Taylor (2002) proposed that P1 (P100) at the occipital site reflects the coarse processing of faces or the 'faceness', while N170 at the posterior temporal site plays a role in featural processing of faces or facial identification. It is therefore essential to determine the functional roles of the early ERP components, such as P1 and N170, in the detection of facial features.

Visual information is processed in parallel via distinct visual channels (Livingstone and Hubel, 1988). Specifically, the magnocellular stream mainly analyzes coarse (low spatial frequency, LSF) information and sends rapid visual signals to the amygdala via the tecto-pulvinar pathway (Schiller et al., 1979; Liddell et al., 2005), while slower visual signals with fine contour (high spatial frequency, HSF) information are chiefly projected to ventral cortical areas such as the fusiform gyrus (FG) via the parvocellular stream (Livingstone

* Corresponding author at: Department of Clinical Neurophysiology, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. Tel.: +81 92 642 5543; fax: +81 92 642 5545.
E-mail address: taisuke@med.kyushu-u.ac.jp (T. Nakashima).

0168-0102/$ – see front matter © 2008 Elsevier Ireland Ltd and the Japan Neuroscience Society. All rights reserved.
doi:10.1016/j.neures.2008.08.009

and Hubel, 1988; Tobimatsu and Celesia, 2006). Vuilleumier et al. (2003a) used filtered faces to explore the underlying mechanisms of facial identification and facial expression recognition using functional magnetic resonance imaging (fMRI). They clearly demonstrated that HSF faces induced more activation of the FG than unfiltered (broad spatial frequency, BSF) and LSF faces, regardless of the emotional expression. In contrast, LSF fearful faces induced more activation of the amygdala than BSF and HSF faces. These results suggest that LSF and HSF stimuli are not simple constituents of BSF, but can differentially emphasize the function of parallel facial processing. Since then, several studies have recorded ERPs to determine the effects of SF on face perception (Goffaux et al., 2003; Holmes et al., 2005; Pourtois et al., 2005; Hsiao et al., 2005; Halit et al., 2006; Nakashima et al., 2008). However, it remains controversial how the early ERP components behave differently and process HSF and LSF information due to methodological differences and/or differences between the SF cut-off values used to create the filtered faces.

In addition, previous ERP/MEG studies have shown discrepant results for early ERP components using filtered face stimuli (Goffaux et al., 2003; Holmes et al., 2005; Hsiao et al., 2005; Pourtois et al., 2005). This discrepancy probably arises through differences in stimulus factors, such as controlling the SF (Herrmann et al., 2004), unification of luminance (Goffaux et al., 2003; Pourtois et al., 2005) and contrast (Goffaux et al., 2003), because P1 is easily affected by these physical traits of visual stimuli (Tanskanen et al., 2005; Tobimatsu and Celesia, 2006).

The first objective of the present study was to systematically explore how differences in the SF information of faces affect the early ERP components, thereby leading to better understanding of the neural representations of faces in Experiment 1. However, even if differential effects of SF on face perception are proved, it will remain unclear whether the early P1 component has a specific role in face recognition, because it is almost always evoked by any kind of stimulus and is seriously affected by physical luminance and contrast (Tobimatsu and Celesia, 2006). Therefore, unless we perfectly control the luminance and contrast information of the face stimuli (Nakashima et al., 2008), it is difficult to determine whether P1 is related to the early stages of face processing per se. To solve this issue, we made physically equiluminant (PEL) face stimuli for the first time using image engineering techniques in Experiment 2. By using PEL face and house stimuli, we can eliminate confounding luminance information in the visual stimuli that obscures the role of P1 in face recognition. Therefore, the purpose of Experiment 2 was to provide further evidence that P1 reflects early face perception stages per se. Furthermore, we expected to separate the other ERP components related to early face processing from the ERPs influenced by physical factors.

2. Experimental procedures

2.1. Subjects

A total of 12 healthy right-handed adults with normal vision (7 females and 5 males; age range, 21–25 years; mean age, 22.3 years) who were naive to our visual stimuli participated in both experiments. The Edinburgh inventory was used for the assessment of handedness (Oldfield, 1971). The subjects had no history of neurological disorders and provided written informed consent to participate in the study after the nature of the experiments, which were in accordance with the Declaration of Helsinki, had been fully explained. The local Ethics Committee of Kyushu University approved this study.

2.2. Stimuli

The original stimuli were 256-level grayscale photographs of neutral faces taken from Japanese and Caucasian Neutral Faces (JACNeuF) (Matsumoto and Ekman, 1988). In Experiment 1, the BSF stimuli were original photographs. LSF and HSF stimuli of faces and houses were created in line with our previous study (Nakashima et al., 2008). We presented 20 levels of filtered photographs in descending series (LSF: from 1.0 to 17.0 cycles/face width; HSF: from 6.0 to 126.0 cycles/face width) and instructed the subjects to choose the appropriate emotion from six basic emotions. The recognition thresholds for each category were defined as the first SF of each facial expression when the subjects consecutively responded with the correct answers. The cut-off frequencies (<4.0 cycles/face for LSF; >30.0 cycles/face for HSF) were the same as those in our previous ERP study. According to a method for two-dimensional SF power spectra (Tsurusawa et al., 2008), the picture of the house contained more HSF information than the picture of the photographed face, while the opposite was true for LSF information (Fig. 1). Next, we normalized the contrast of all the images through unification of the standard deviation (S.D.) of all the stimuli. If the contrast of HSF images was not adjusted, they became relatively low-contrast stimuli because the HSF power of the natural images was originally low (Fig. 2B) and the evoked responses would therefore be reduced. Our HSF stimuli were strictly adjusted to ensure that the LSF and HSF images had the same contrast (Fig. 2A, right column). We also used high-order filters to allow us to extract HSF information of the face stimuli more strictly. The mean luminance and contrast of three SF images were equated by normalizing the mean and S.D. of the gray values of all the stimuli using our own program written in C language (mean luminance, 48 cd/m^2; mean gray value ± S.D., 128 ± 40). Representative examples of the face stimuli are shown in Fig. 2A.

In Experiment 2, we created PEL stimuli using our own program written in C language by one of the authors (K.K.). The color images consisted of pixels, and each pixel had red (R), green (G) and blue (B) values. We defined L as representing the luminance of any pixel, with R', G' and B' representing the red, green and blue values of the PEL images, respectively. L was calculated using the following equation:

$$L = 0.299R + 0.587G + 0.114B$$

This equation shows that the luminance of any pixel can be calculated by linear superposition of its R, G and B values. The luminance of pixels in an original image varied pixel by pixel. Thus, these variations in luminance among the pixels needed to be equalized to generate a PEL image. The R', G' and B' values were calculated from the R, G and B values at the same position of the original image using the following equations:

$$R' = \frac{R}{L}$$

$$G' = \frac{G}{L}$$

$$B' = \frac{B}{L}$$

Substituting these equations into the previous equation gives the following equation:

$$0.299R' + 0.587G' + 0.114B' = 0.299\left(\frac{R}{L}\right) + 0.587\left(\frac{G}{L}\right) + 0.114\left(\frac{B}{L}\right) = \frac{(0.299R + 0.587G + 0.114B)}{L} = 1$$

Hence, the luminance of each pixel of all PEL images can be completely equated (48 cd/m^2; 128 levels of bitmap scale), thereby resulting in no information regarding contrast. A diagram explaining the PEL surface is shown in Fig. 2C as a reference. Theoretically, a PEL surface is three-dimensional, and defined by three scales (R, G and B). Any pixel on this surface has the same physical luminance (R', G' and B'). Finally, the color (R, G and B) in the original facial photographs was projected onto the PEL surface (R', G' and B') using linear transformation, such that all the luminance in each pixel was equated (PEL image). Representative examples of the stimuli are shown in Fig. 2D.

2.3. Tasks

In Experiment 1, we employed a simple passive viewing task with a target stimulus (shoes) to maintain the subjects' vigilance and direct their attention away from the nature of the visual stimuli (Nakashima et al., 2008). The subjects passively viewed BSF, LSF and HSF stimuli that were pseudo-randomly presented under each SF condition for 300 ms with inter-stimulus intervals of 700 ms on a 17-in. CRT screen at a viewing distance of 114 cm (visual angle, 8.3 × 10.0°) in a dark room. The refresh rate of the CRT monitor was 100 Hz. The subjects were instructed to fix their eyes on a point at the center of the screen, push a button when a target stimulus (shoes; probability 10%) appeared and ignore other stimuli.

In Experiment 2, the subjects also passively viewed BSF and PEL stimuli that were presented in pseudo-random orders. The other procedures were the same as those in Experiment 1.

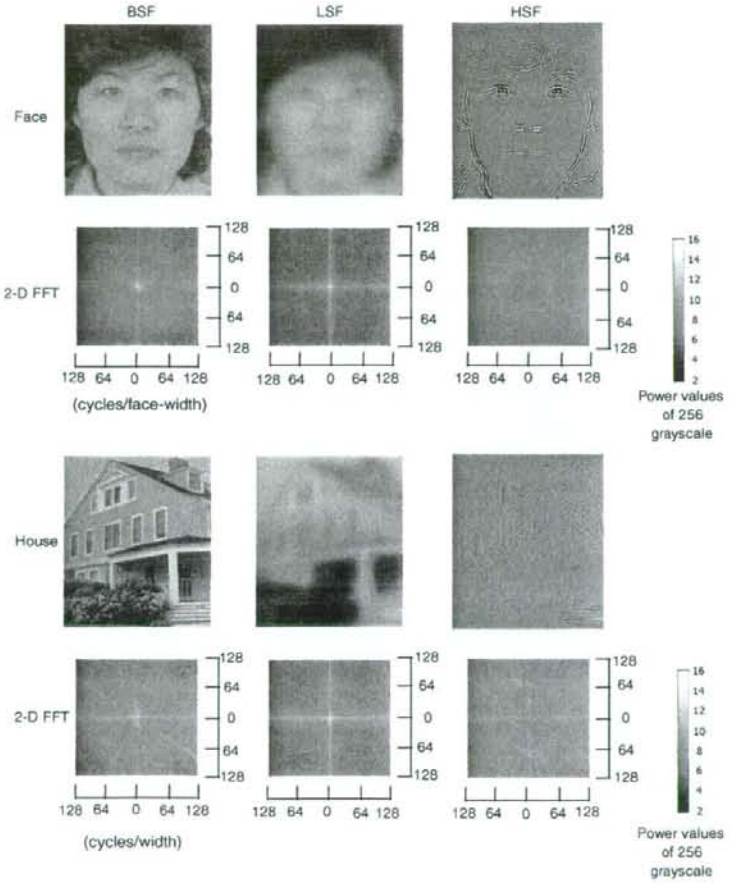

Fig. 1. Two-dimensional spatial frequency (SF) power spectra of pictures of a neutral face and a house. Fast Fourier transformation (FFT) was applied to these pictures. The SF (cycles/face) ranges from 0 to 128, while the power spectra vary from 2 to 16. The power values for the BSF house spread to peripheral parts that include high SF components. However, the major signals for the BSF face are found around the center, in which low SF components are present. Note that there are differences in the distributions of the power spectra between the LSF face and LSF house, and also between the HSF face and HSF house. These effects result from equating the recognition thresholds (cut-off frequencies) between the face and the house by psychophysical experiments.

2.4. ERP recordings

In both experiments, silver–silver chloride electrodes were applied to 20 scalp sites based on the International 10-20 method. All the recording electrodes were referred to an electrode at the nose-tip, and the impedance was kept below 5 kΩ. Vertical and horizontal electrooculograms (V-EOG and H-EOG, respectively) were also recorded to eliminate artifacts caused by blinking and eye movements in the off-line analysis. In total, 220 responses were averaged for each stimulus with a bandpass filter between 0.05 and 200 Hz. The sampling rate was 1333 Hz. The baseline was set from −72 ms to the stimulus onset (0 ms). Artifact rejections were conducted manually on the basis of deflections above 30 μV from the baseline. The mean rejection rate was 30.8%.

2.5. Data analysis and statistics

In Experiment 1, P1 and N170 were analyzed separately (mean latencies of 111 and 155 ms, respectively). The P1 amplitudes were measured from the baseline to the first positive peak at the Oz electrode because we found no preceding component at Oz. The peaks of N170 were measured as relative amplitudes from the first positive peaks to the first negative peaks at the T5 and T6 electrodes to eliminate the effects of the preceding component (P1) at the temporo-occipital region. Statistical analyses for each component were conducted by two-way repeated-measures analysis of variance (ANOVA) for SF (LSF, BSF or HSF) × stimulus category (faces or houses). As post hoc analyses, interaction contrast methods were employed for P1 and N170 to detect the direct effects of LSF and HSF stimuli by using the responses to BSF stimuli as a benchmark.

In Experiment 2, we intentionally focused on the spatiotemporal behavior of the early components, which were P1 and the following negativity at the occipital region and the face-specific N170 at occipito-temporal sites (mean latencies for all category; 114, 158 and 162 ms, respectively) with inclusion of the laterality and the topographical distribution. The latency of the following negativity at the occipital region for PEL faces was around 170 ms (Table 1), and this negativity was therefore designated N2 to avoid confusion with the face-specific N170. The occipital P1 amplitudes were measured from the baseline at O1 and O2, while those of N2 and N170 were measured as relative amplitudes from the preceding peaks at the O1, O2, T5 and T6 electrodes. P1 was analyzed by three-way repeated-measures ANOVA for stimulus category (faces or houses) × image condition (BSF or PEL) × laterality (O1 or O2). N2 and N170 were analyzed by four-way repeated-measures ANOVA for stimulus category (faces or houses) × image condition (BSF or PEL) × laterality (O1, T5 or O2, T6) × component (N2 at O1, O2 or N170 at T5, T6). Multiple comparisons were performed with a paired *t*-test (Bonferroni correction). Topographical maps of the ERPs were created using a source derivation method (Hjorth, 1975).

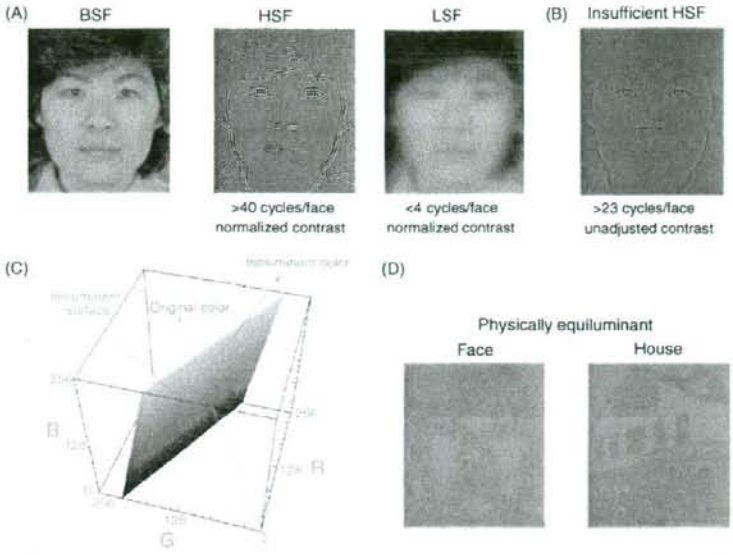

Fig. 2. (A) Representative examples of the stimuli used in Experiment 1. BSF is an original non-filtered image that contains broad spatial frequency components. HSF and LSF faces are filtered by fast Fourier transformation. The HSF face extracts information with high spatial frequencies (>40 cycles/face width) and emphasizes the fine features of the facial components. The LSF face consists of information with low spatial frequencies (<4 cycles/face width) and preserves the coarse facial image. (B) Insufficient HSF is caused by using a lower cut-off point and not adjusting the contrast, such that the HSF component of the image is not fully enhanced. Therefore, these stimuli are inadequate for our ERP experiments. The threshold for recognizing faces was measured by psychophysical experiments in advance, and the contrast of the images was adequately normalized. (C) Physically equiluminant surface in the RGB three-dimensional space. By putting all the pixels of the facial images onto an equiluminant surface with linear transformation, the luminance of all the pixels is equated. (D) Representative examples of the stimuli used in Experiment 2. BSF faces (A) retain contrast while physically equiluminant (PEL) faces consist of the same luminance pixels without contrast.

In both experiments, the latencies and amplitudes of the P1, N170 and N2 components used for statistical analyses were evaluated from individual data, rather than from grand-averaged data. Grand-averaged waveforms were created by simply averaging all the subjects' ERPs. Thus, it should be noted that the mean amplitudes of the P1, N170 and N2 components do not necessarily correspond to the grand-averaged waveforms.

All statistical analyses in both experiments were performed using SPSS 11.01 (SPSS Inc., Chicago, IL).

3. Results

3.1. Experiment 1

3.1.1. Performance of target detection and grand-averaged ERPs to face and target stimuli

The mean (±S.D.) reaction time for the target stimuli was 386.8 ± 31.5 ms, and the correct detection rate (±S.D.) was 99.9 ± 0.1%. After the ERP recordings, all the subjects verbally reported in short interviews that they could recognize faces and houses under all the stimulus conditions (BSF, LSF, HSF and PEL).

Table 1
P1 and N170 latencies under each condition

Component	Category	Image condition	Mean latency ± SEM (ms)
P1	Face	BSF	107.6 ± 8.0
		LSF	114.3 ± 5.3
		HSF	105.3 ± 24.0
	House	BSF	110.8 ± 16.1
		LSF	110.6 ± 16.5
		HSF	112.5 ± 19.2
N170 (T6)	Face	BSF	145.7 ± 13.5
		LSF	154.8 ± 13.2
		HSF	168.1 ± 16.5
	House	BSF	146.5 ± 13.8
		LSF	143.3 ± 16.1
		HSF	169.1 ± 16.4

As shown in Fig. 3, the first positive component (P1) was predominant over the occipital region, and was larger for LSF faces. The second negative component (N170) was recorded from the temporo-occipital areas. N170 was clearly larger for HSF faces than for the other conditions in the right hemisphere. In contrast, the P1 and N170 latencies for each filtered image did not differ between face and house stimuli, although the latency of N170 tended to be prolonged for HSF images regardless of the stimulus category (Table 1). P300 was largest at the parietal region and only recorded for target stimuli, indicating that the subject attention was well controlled.

3.1.2. P1 component

Fig. 4A shows the grand-averaged P1 waveforms at Oz for face and house stimuli under the three SF conditions. The mean amplitudes ± SEM are shown in Fig. 4B. The P1 amplitude was greater for LSF faces than for BSF faces (Figs. 4A and B, left). However, the amplitude for LSF houses was not augmented. ANOVA supported this tendency, showing an interaction of SF (BSF, LSF or HSF) × stimulus category (faces or houses) ($F(2,22) = 3.550$, $P < 0.05$). A post hoc analysis revealed an interaction contrast between LSF and BSF ($F = 10.45$, $P < 0.01$) (Fig. 4B, right). In contrast, no significant difference was found for the P1 latencies. Overall, the P1 amplitudes were significantly enhanced by LSF faces relative to BSF faces, but not by LSF houses relative to BSF houses.

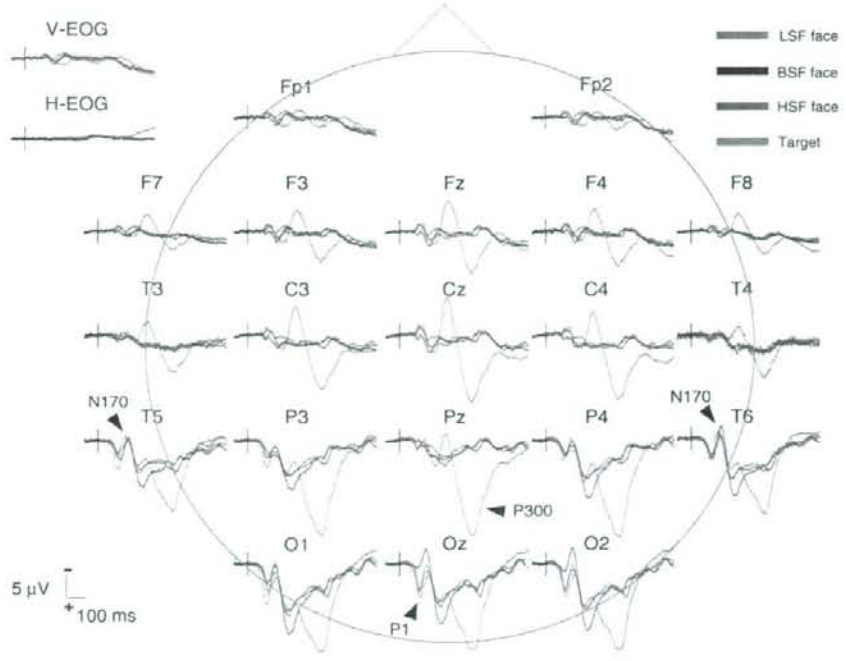

Fig. 3. Topographical mapping of grand-averaged ERPs to the face and target stimuli. The P1 amplitude increases for the LSF condition, while the N170 amplitude is augmented by the HSF condition in the right hemisphere. Note that P300 is only clearly recorded for the target stimuli.

3.1.3. N170 component

The grand-averaged N170 waveforms at the T5 and T6 electrodes are shown in Fig. 5A. The mean (±SEM) values of the N170 amplitudes for all the stimuli are shown in Fig. 5B. It is apparent that the N170 amplitudes were much larger for HSF faces than for BSF faces in the right hemisphere (Fig. 5A and B, right). In contrast, the N170 amplitudes for HSF houses were not augmented. ANOVA confirmed a significant interaction of SF (BSF, LSF or HSF) × stimulus category (faces or houses) (T5: $F(2,22) = 3.99$, $P < 0.05$; T6: $F(2,22) = 4.83$, $P < 0.05$), similar to the case for the P1 amplitude. A post hoc test clarified that an interaction contrast of BSF × HSF was only present at T6 (T5: $F = 5.75$, $P > 0.05$; T6: $F = 8.12$, $P < 0.05$) (Fig. 5B, lower). However, we found no interaction of the latency. In summary, the N170 amplitudes in the right hemisphere were significantly augmented by HSF faces relative to BSF faces, but not by HSF houses relative to BSF houses.

3.2. Experiment 2

3.2.1. Grand-averaged ERP waveforms to PEL and BSF stimuli

As shown in Fig. 6, BSF and PEL faces clearly elicited N170 while BSF and PEL houses did not. These ERP responses confirmed that our PEL faces were recognized as faces similar to the BSF faces (Fig. 6, left). Although the P1 amplitude for PEL houses was notably diminished compared with that for BSF houses (Fig. 6, right), the P1 amplitude for PEL faces was rather retained relative to that for PEL houses (Fig. 6, lower). In addition, the N170 latency was almost the same as the occipital N2 latency for BSF faces (Fig. 6, left upper), but much later than the N2 latency for PEL faces (Fig. 6, left lower).

3.2.2. P1 component

With regard to the P1 amplitude, we found a significant interaction of stimulus category (faces or houses) × image condition (BSF or PEL) ($F(1,11) = 24.051$, $P < 0.001$). A post hoc analysis revealed significant differences for the P1 variations induced by both PEL faces and houses (Table 2). As shown in Table 2 and Fig. 6 (right lower), P1 for houses was prominently decreased by the PEL image. In contrast, the P1 decrease in response to PEL faces was relatively smaller than that in response to PEL houses (Table 2 and Fig. 6, left lower). Topographical mapping confirmed these results (Fig. 7). The P1 amplitudes for faces were even robust for the PEL images. We did not find any significant effect on the P1 latency.

3.2.3. N2 and N170 components

We found a significant interaction of stimulus category (faces or houses) × image condition (BSF or PEL) × component (N2 at occipital or N170) concerning the latency. A post hoc analysis revealed that N2 was evoked about 20 ms earlier than N170 for PEL faces but not for faces in the other conditions (Table 3 and Fig. 6, left lower). This result was confirmed by topographical mapping of N2 and N170 (Fig. 7). Regarding the amplitude, no interaction by the image condition was observed. The occipital N2 was separable from N170 according to their latencies.

3.2.4. Occipital N2 component

We further analyzed the laterality of the occipital N2 and N170 components under all the conditions with a paired *t*-test. As a result, we found that N2 at the left hemisphere was significantly smaller than N2 at the right hemisphere only for PEL faces ($t = 3.379$, $P < 0.05$). In contrast, there was no significant laterality

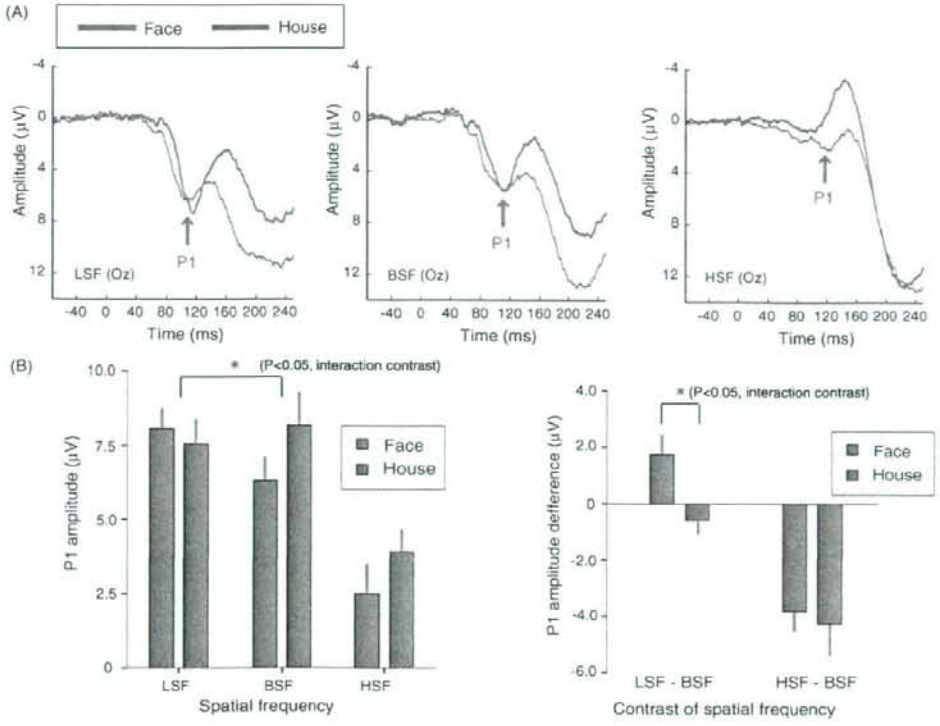

Fig. 4. (A) P1 responses to each stimulus at the mid-occipital region (Oz). The grand-averaged P1 is specifically larger for LSF faces (left) than for BSF (middle) and HSF (right) faces. (B) Mean amplitudes (±SEM) of P1 for each stimulus (left) and the interaction contrasts (right). There is a significant difference between faces and houses for each image (left). P1 for faces is significantly enhanced by the LSF condition. There is a significant interaction contrast between LSF and BSF faces (right), suggesting that the P1 component is highly sensitive to LSF faces. It should be noted that the amplitudes of the P1 responses to BSF faces and house appear to be similar in (A), but different in (B). This discrepancy is probably caused by the methods used to measure the amplitudes. Specifically, the waveforms of all subjects were averaged and affected by the P1 latencies in (A), resulting in the same amplitudes, whereas the amplitudes were measured in each individual independently of the latencies in (B), resulting in different amplitudes. Therefore, the discrepancy in the amplitudes is probably caused by the large variation in the P1 latency as well as the following negativity for BSF houses compared with BSF faces (see Table 1).

under the other conditions. These results were again consistent with topographical mapping of N2 and N170 (Fig. 7), and the occipital N2 was separable from N170 with left-dominant attenuation.

4. Discussion

4.1. Face robustness of the P1 component assures holistic face processing

We have demonstrated quantitatively that P1 was only augmented by LSF faces and retained by PEL faces, suggesting that P1 is involved in the early holistic processing of face perception. A number of psychological studies using inverted faces or mosaic faces have reported the importance of coarse information for facial recognition (Yin, 1969; Harmon and Julesz, 1973; Ginsburg, 1986; Farah et al., 1998; Calder et al., 2000; Eimer, 2000a; Itier and Taylor, 2002). Sagiv and Bentin (2001) suggested that a face-specific visual mechanism is triggered whenever a stimulus contains sufficient information to generate the concept of a face (e.g., gestalt-based). A single unit study also suggested that coarse facial information plays a key role in recognizing faces (Sugase et al., 1999). Recently, the face specificity of the early ERP components, such as P1 (P100), has been focused upon (Linkenkaer-Hansen et al., 1998; Halit et al., 2006; Itier and Taylor, 2002; Liu et al., 2002; Herrmann et al., 2004; Goto et al., 2005). Nevertheless, it has remained uncertain whether the early components reflect face-specific processing and what aspect of spatial information in faces is the key for early face perception. We have provided further evidence that P1 reflects early face selectivity and its holistic process per se, because coarse LSF faces elicited a large P1 and our PEL faces also clearly evoked P1 (P1 face robustness), whereas PEL houses did not. Theoretically, filtering out the HSF leaves faces with a higher energy of LSF, while subtracting the LSF gives a higher power of HSF to houses. Thus, the P1 amplitude increases for LSF faces relative to LSF houses and vice versa for HSF houses. However, this was not the case in our study for the P1 amplitude to HSF houses (see Fig. 3). Therefore, P1 could reflect the early processing of face robustness by using LSF information.

4.2. Functional role of N2 in face recognition

To the best of our knowledge, no previous studies have focused on separating the other facial components confounded by various physical factors at the occipital region. In the present study, N2 was

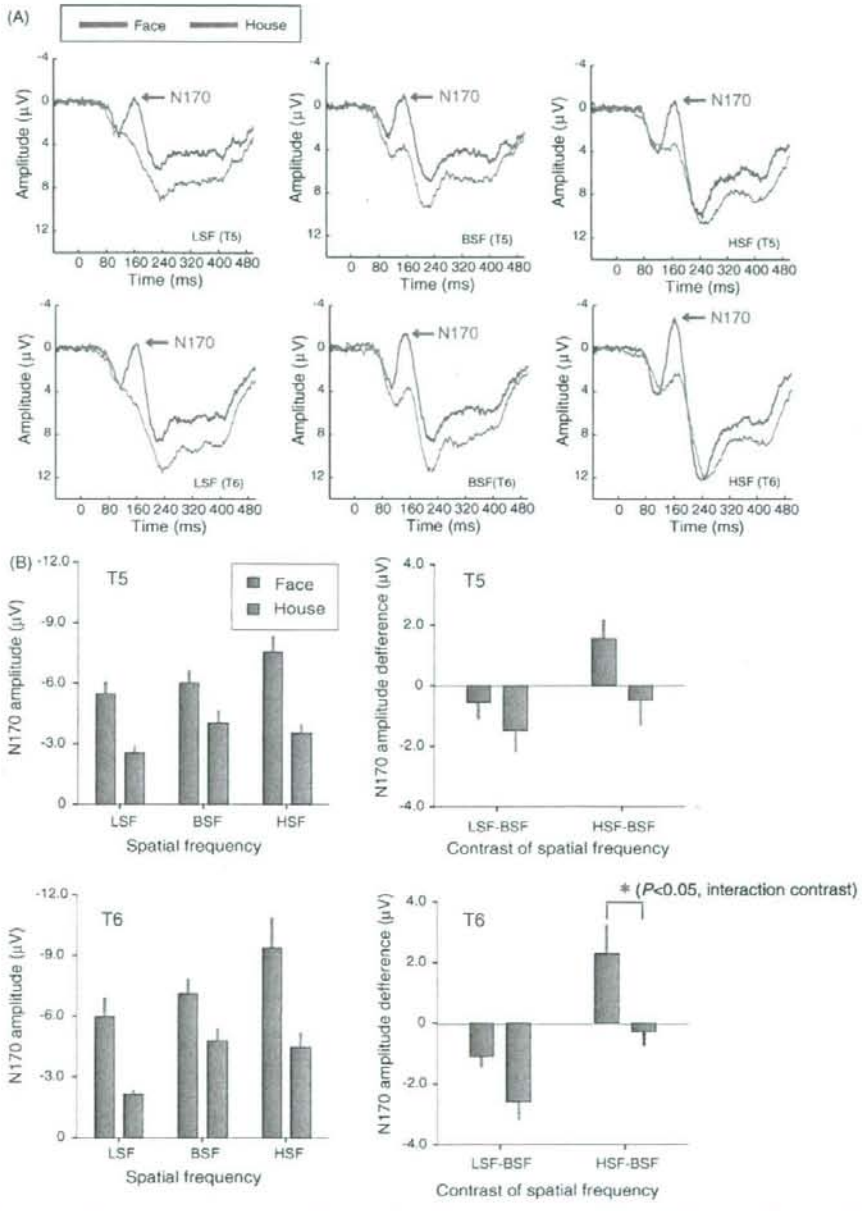

Fig. 5. (A) N170 responses to each stimulus at the temporo-occipital electrodes (T5 and T6). The grand-averaged N170 for HSF faces is clearly larger than those for BSF and LSF faces with right hemisphere predominance. (B) Mean amplitudes (±SEM) of N170 for each stimulus (left) and the interaction contrasts (right) at the T5 and T6 electrodes. There is a significant difference between faces and houses for each image. In contrast to the behavior of P1, a significant interaction contrast is only found between the HSF and BSF conditions at T6 (right lower).

only separable from P1 or N170 when PEL stimuli were used. PEL faces evoked N2 earlier than PEL houses and significantly reduced the amplitudes of N2 at the left occipital region. These findings clearly indicate that the left occipital N2 was evoked by a distinct source independently of N170 and contributed to PEL face processing.

Human faces possess rather complex information of contrast compared with other creatures. For example, human eyes have a

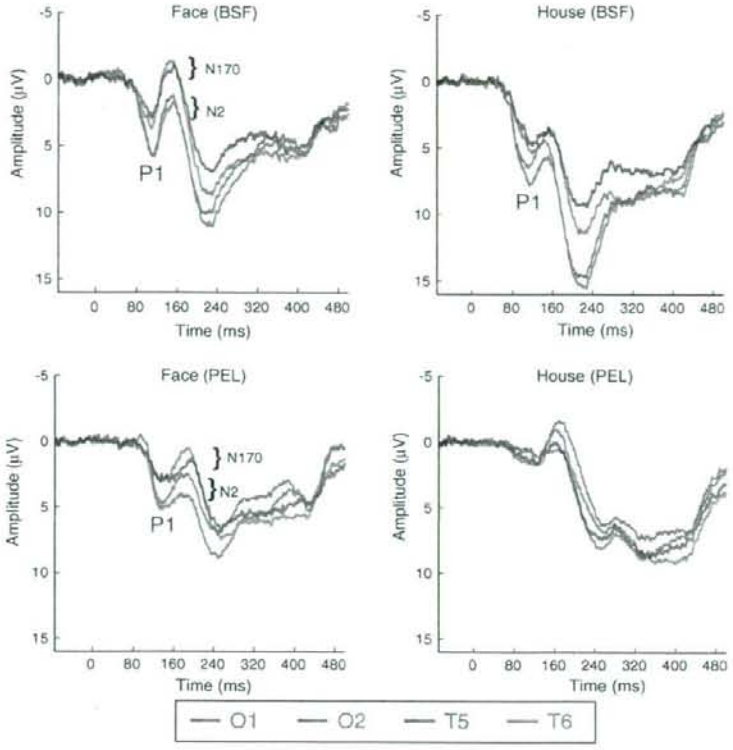

Fig. 6. Grand-averaged ERP waveforms to all the stimulus conditions at the occipital (blue: O1; gray: O2) and temporo-occipital (red: T5; green: T6) regions. The P1 component is recorded for BSF faces and houses. It is also evident for PEL faces. However, it almost disappears for PEL houses. Although N2 and N170 are elicited almost simultaneously for BSF faces, N2 is significantly earlier than N170 for PEL faces with left predominant attenuation. N170 is clearly seen for both BSF and PEL faces, but not for BSF and PEL houses.

distinct white sclera and a dark iris in a narrower space than other primates (Kobayashi and Kohshima, 1997). In addition, the teeth are lighter and the eyebrows are darker than other facial parts. These facial components thus have high contrast information. On the other hand, the nose and lips are curvilinear, and have little contrast but much complexity. However, houses tend to have a simple rectangular type of contrast. Hence, the complex contrast information based on complicated luminance can play a key role in face recognition. Thus, we infer that N2 may be related to processing such complex contrast features in faces.

This result may represent a novel finding, because this characteristic change of the occipital N2 has not previously been reported. This component may represent a link between P1 at the occipital region and N170 at the temporo-occipital region. We therefore consider that N170 does not solely integrate all the aspects of facial encoding and that the occipital N2 may be an electrophysiological component of the "facial contrast detector" or an "intermediator" before N170 and after P1/N100.

Table 2
P1 amplitudes under each condition

Category	Image condition	Mean amplitude ± SEM (μV)	*P* value
Face	BSF	6.7 ± 0.7	<0.05
	PEL	5.2 ± 0.9	
House	BSF	9.8 ± 1.3	<0.001
	PEL	2.4 ± 0.3	

From a comprehensive standpoint of neurocognitive models for face perception (Bruce and Young, 1986; Haxby et al., 2000; Adolphs, 2002) intermixed with our results for Experiment 1, faces are roughly processed at the occipital region as if we perceive faces as faces based on the laws of gestalt at the P1 (P100)/N100 stages (Itier and Taylor, 2002; Liu et al., 2002; Herrmann et al., 2004; Goto et al., 2005). At the N170/M170 level, facial features are analyzed in detail at the temporo-occipital region (Itier and Taylor, 2002; Liu et al., 2002). In this case, the human visual system needs to acquire detailed contrast information in advance in order to analyze the facial features at the N170 level. Unless we use contrast information, we cannot analyze various facial features. Because the left occipital N2 was selectively reduced by PEL faces, we infer that N2 is essential before analysis of the fine information of faces.

This idea is consistent with the accumulated evidence that holistic facial information is mainly analyzed by the right hemisphere and not the left hemisphere (Tanaka and Farah, 1993). Interestingly, Vuilleumier et al. (2003b) reported a female patient with a focal left lateral temporo-occipital venous infarct who misidentified unknown faces to her as familiar, but not well-known, faces. The authors suggested that part of the difficulty for this patient was having to place more reliance on global features

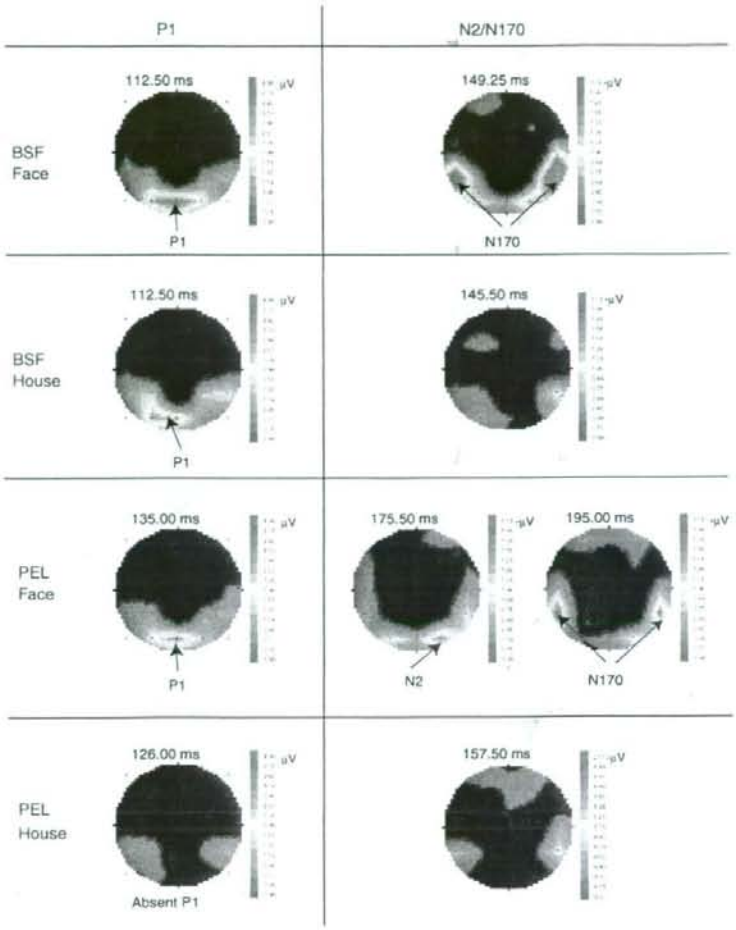

Fig. 7. Topographical maps of P1, N2 and N170 for each condition. For graphical presentation, these maps are created by adopting the relative amplitudes. Thus, both the positive and negative components are displayed in red. P1 disappears for PEL houses but is clearly preserved for PEL faces. Interestingly, N2 is separable from N170 for PEL faces. N170 is clearly seen for faces but not for houses, regardless of the image condition.

processed chiefly by the right hemisphere and less reliance on internal parts of faces whose processing depends on the left hemisphere. This greater reliance on the left hemisphere for analyzing facial features is compatible with our experimental results. We assume that the activation of the left occipital N2 was reduced due to the lack of information of contrast features of facial parts in the PEL images. As a result, the left occipital N2 amplitude may be decreased.

Table 3
N2 and N170 latencies under each condition

Category	Image condition	Component	Mean latency ± SEM (ms)	*P* value
Face	BSF	N2 (O1, O2)	147.8 ± 4.4	n.s.
		N170 (T5, T6)	147.8 ± 4.2	
	PEL	N2 (O1, O2)	176.4 ± 4.6	<0.05
		N170 (T5, T6)	195.2 ± 4.8	
House	BSF	N2 (O1, O2)	140.7 ± 4.9	n.s.
		N170 (T5, T6)	144.9 ± 4.7	
	PEL	N2 (O1, O2)	166.2 ± 2.9	n.s.
		N170 (T5, T6)	158.8 ± 6.6	

4.3. N170 reflects the encoding of face configurations

Fiorentini et al. (1983) demonstrated that HSF information for feature processing was not redundant, but in fact sufficient for facial identification. However, electrophysiological studies have not proved this important role of HSF information for face recognition. In the present study, N170 was only enhanced by HSF faces with right hemisphere predominance. Although it has been presumed that N170 reflects the structural encoding of faces (Bentin et al., 1996; Eimer and McCarthy, 1999; Sagiv and Bentin, 2001), it remains uncertain which aspects of facial information are encoded by N170.