

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, Behnia 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, Puce 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 Psychiatr* 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, Kinoo 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, Manguière 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 Parmanus 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, Ceesia 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.



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

Taisuke Nakashima<sup>a,f,\*</sup>, Kunihiro Kaneko<sup>b</sup>, Yoshinobu Goto<sup>a,d</sup>, Tomotaka Abe<sup>c</sup>, Takako Mitsudo<sup>a,f</sup>, Katsuya Ogata<sup>a</sup>, Akifumi Makinouchi<sup>c</sup>, Shozo Tobimatsu<sup>a</sup>

<sup>a</sup> Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

<sup>b</sup> Department of Intelligent Systems, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

<sup>c</sup> Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan

<sup>d</sup> Department of Occupational Therapy, Faculty of Rehabilitation, International University of Health and Welfare, Okawa, Japan

<sup>e</sup> Department of Information and Network Engineering, Kurume Institute of Technology, Kurume, Japan

<sup>f</sup> 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.

© 2008 Elsevier Ireland Ltd and the Japan Neuroscience Society. All rights reserved.

### 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 Tuckmachi, 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](mailto:taisuke@med.kyushu-u.ac.jp) (T. Nakashima).



and Hubel, 1988; Tobimatsu and Clesia, 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 (Hermann 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 Clesia, 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 Clesia, 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<sup>2</sup>; 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:

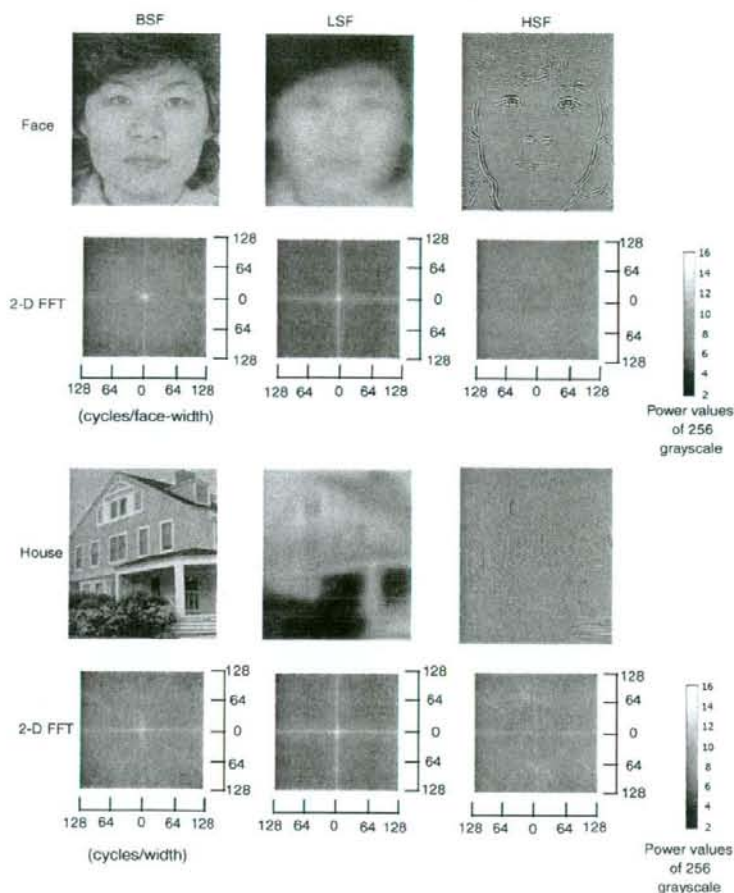
$$0.299R' + 0.587G' + 0.114B' = 0.299\left(\frac{R}{L}\right) + 0.587\left(\frac{G}{L}\right) + 0.114\left(\frac{B}{L}\right) \\ = \frac{(0.299R + 0.587G + 0.114B)}{L} = 1$$

Hence, the luminance of each pixel of all PEL images can be completely equated (48 cd/m<sup>2</sup>; 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 k $\Omega$ . 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  $\mu$ V from the baseline. The mean rejection rate was 30.8%.

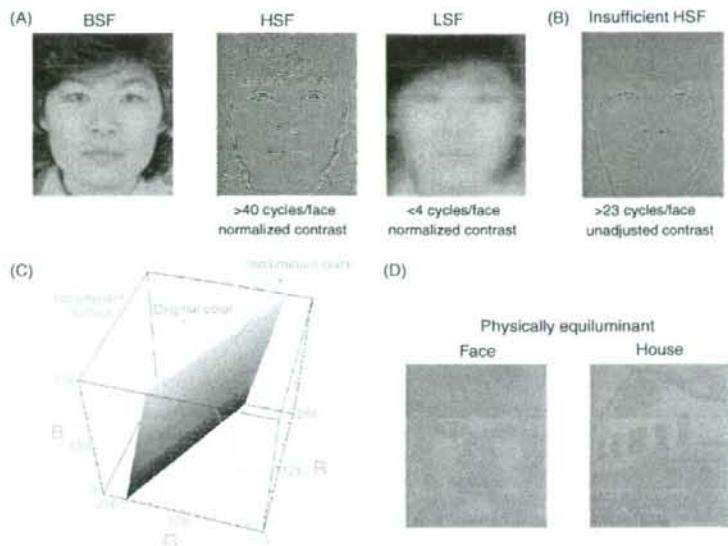
#### 2.5. Data analysis and statistics

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)  $\times$  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)  $\times$  image condition (BSF or PEL)  $\times$  laterality (O1 or O2). N2 and N170 were analyzed by four-way repeated-measures ANOVA for stimulus category (faces or houses)  $\times$  image condition (BSF or PEL)  $\times$  laterality (O1, T5 or O2, T6)  $\times$  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 ( $\pm$ S.D.) reaction time for the target stimuli was  $386.8 \pm 31.5$  ms, and the correct detection rate ( $\pm$ S.D.) was

$99.9 \pm 0.1\%$ . 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).

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  $\pm$  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)  $\times$  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.

**Table 1**  
P1 and N170 latencies under each condition

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

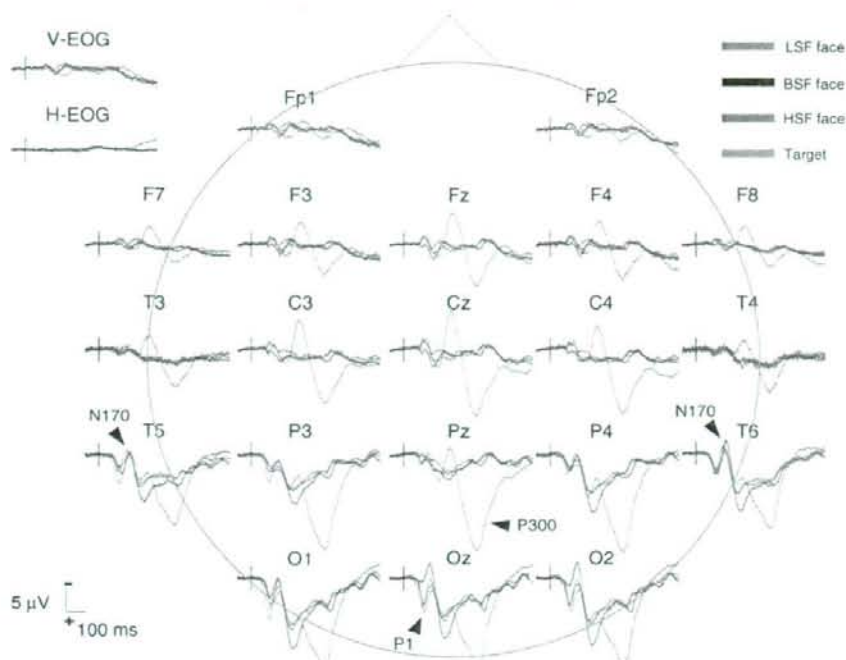


Fig. 3. Topographical mapping of grand-averaged ERPs to the face and target stimuli. The P1 amplitude increases for the LSF condition, while the N170 amplitude is augmented by the HSF condition in the right hemisphere. Note that P300 is only clearly recorded for the target stimuli.

### 3.1.3. N170 component

The grand-averaged N170 waveforms at the T5 and T6 electrodes are shown in Fig. 5A. The mean ( $\pm$ 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)  $\times$  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  $\times$  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)  $\times$  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.

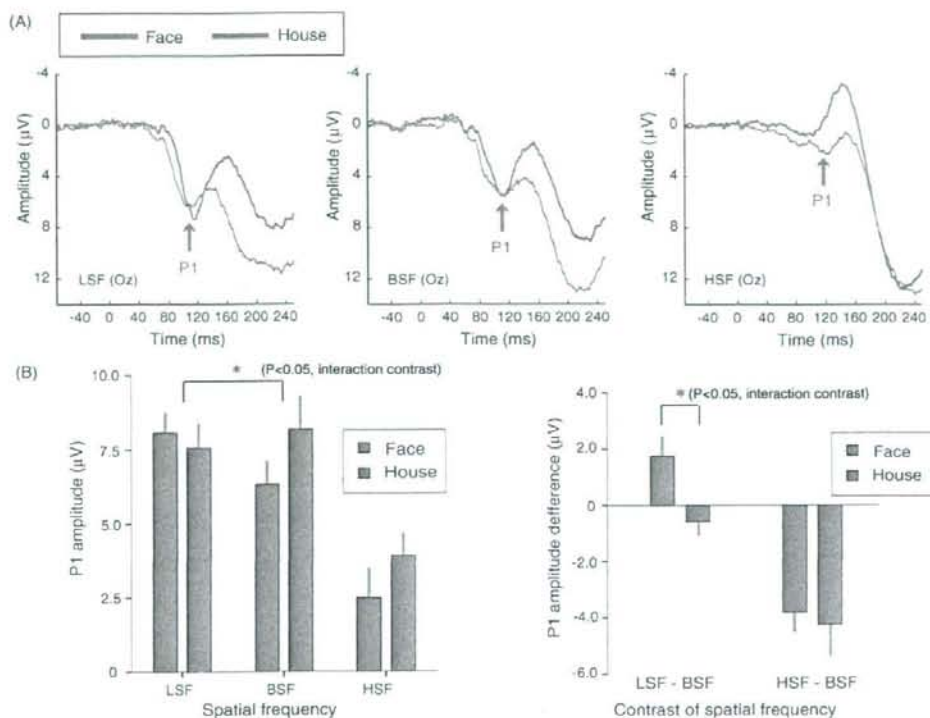
### 3.2.3. N2 and N170 components

We found a significant interaction of stimulus category (faces or houses)  $\times$  image condition (BSF or PEL)  $\times$  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 ( $\pm$ 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

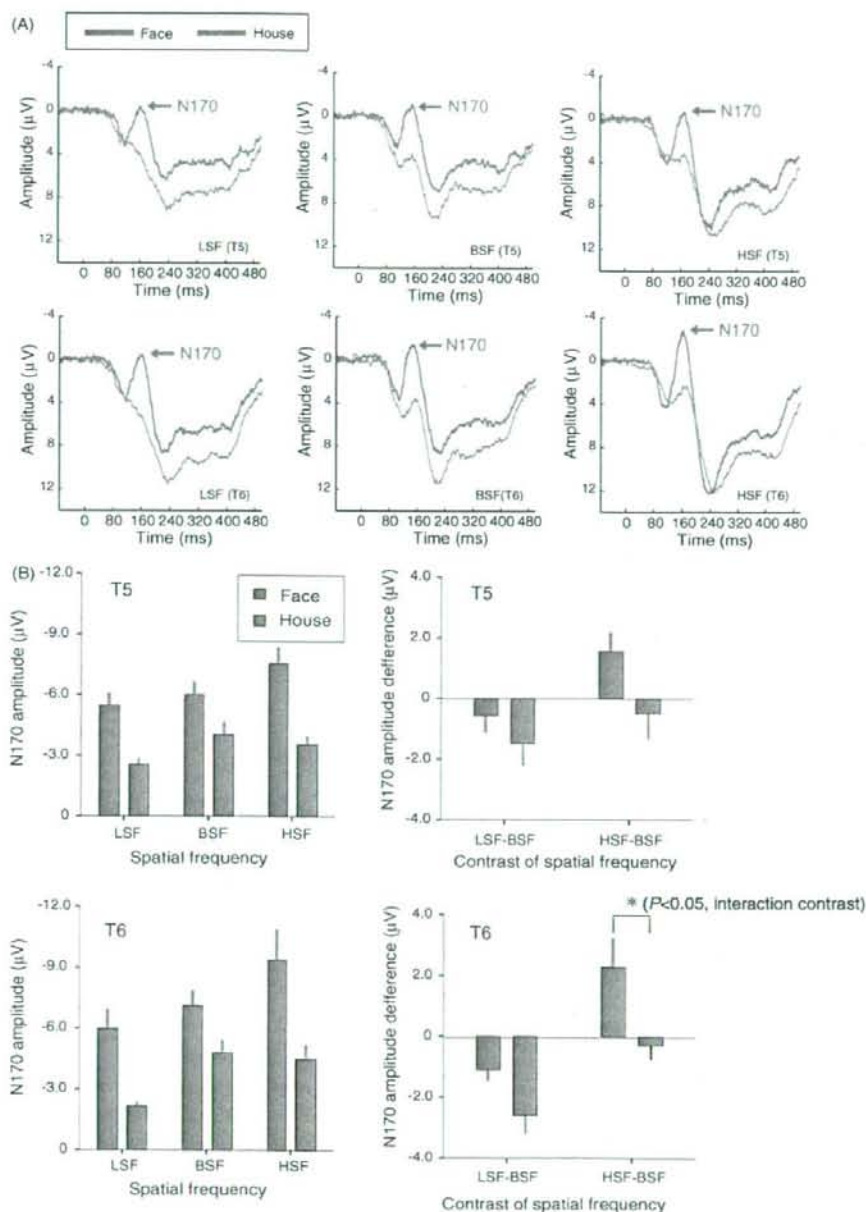
##### 4.1. Face robustness of the P1 component assures holistic face processing

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 Ben-Tin (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



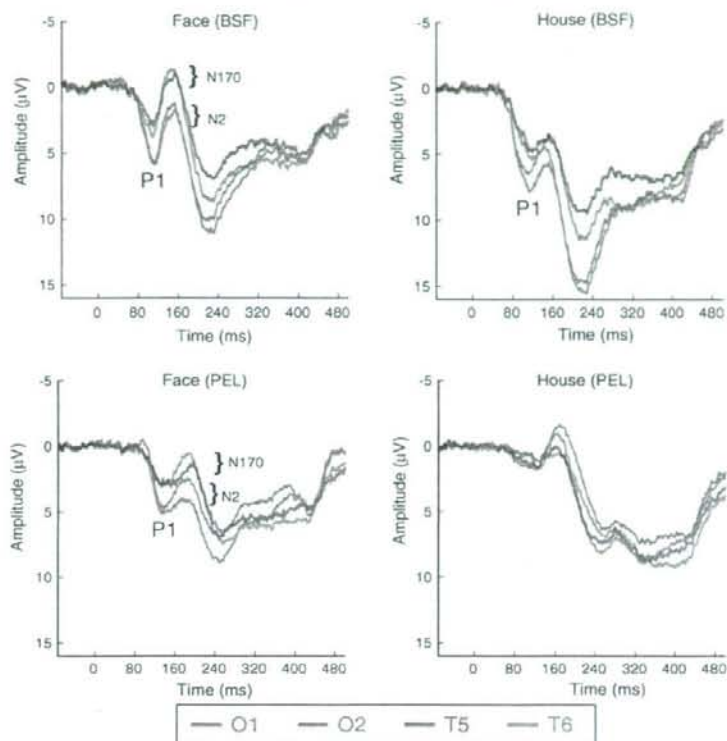
**Fig. 5.** (A) N170 responses to each stimulus at the temporo-occipital electrodes (T5 and T6). The grand-averaged N170 for HSF faces is clearly larger than those for BSF and LSF faces with right hemisphere predominance. (B) Mean amplitudes ( $\pm$ SEM) of N170 for each stimulus (left) and the interaction 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.

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

**Table 2**  
P1 amplitudes under each condition

Category	Image condition	Mean amplitude $\pm$ SEM ( $\mu$ V)	P value
Face	BSF	6.7 $\pm$ 0.7	<0.05
	PEL	5.2 $\pm$ 0.9	
House	BSF	9.8 $\pm$ 1.3	<0.001
	PEL	2.4 $\pm$ 0.3	

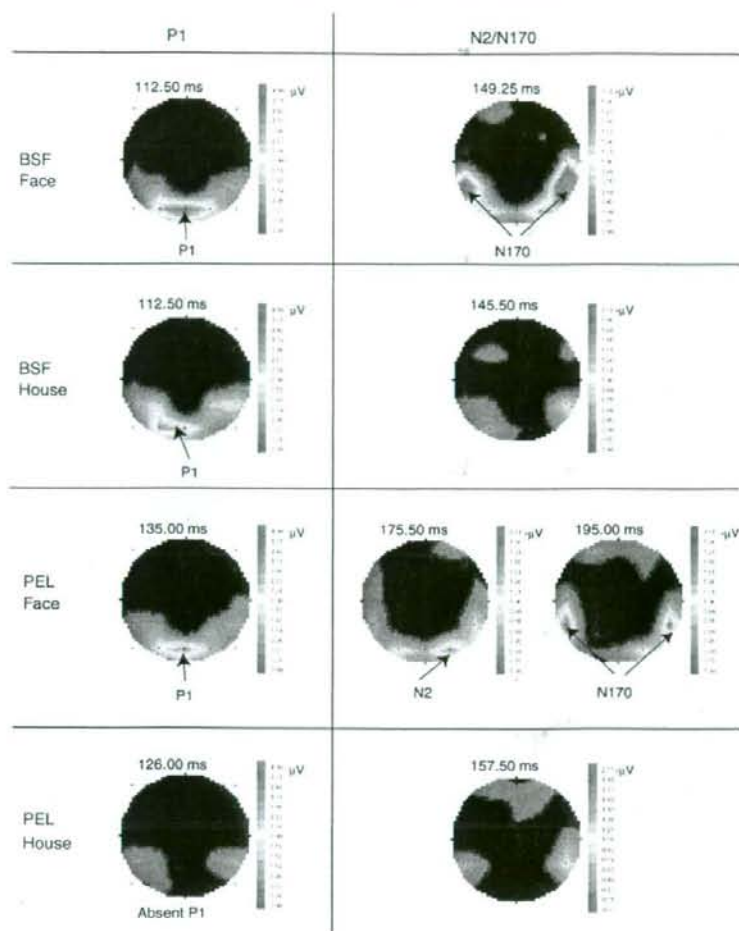


Fig. 7. Topographical maps of P1, N2 and N170 for each condition. For graphical presentation, these maps are created by adopting the relative amplitudes. Thus, both the positive and negative components are displayed in red. P1 disappears for PEL houses but is clearly preserved for PEL faces. Interestingly, 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 $\pm$ SEM (ms)	P value
Face	BSF	N2 (O1, O2)	147.8 $\pm$ 4.4	n.s.
		N170 (T5, T6)	147.8 $\pm$ 4.2	
	PEL	N2 (O1, O2)	176.4 $\pm$ 4.6	
		N170 (T5, T6)	195.2 $\pm$ 4.8	
House	BSF	N2 (O1, O2)	140.7 $\pm$ 4.9	n.s.
		N170 (T5, T6)	144.9 $\pm$ 4.7	
	PEL	N2 (O1, O2)	166.2 $\pm$ 2.9	
		N170 (T5, T6)	158.8 $\pm$ 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.



Our findings indicate that N170 reflects the facial encoding of feature (HSF) information. Previous studies have reported face inversion effects, enlargement of N170 amplitudes and delay of N170 latencies (Eimer, 2000b; Rossion et al., 2000). These authors suggested that the face inversion effects were caused by loss of holistic and configural information of faces. If this is the case, humans would need to make an effort to recognize inverted faces with feature information and the N170 amplitude would be augmented. In accord with this assumption, the face inversion effects for N170 disappeared for HSF faces (Goffaux et al., 2003).

To date, previous studies have reported that HSF faces relatively attenuated, did not affect or increased the N170 amplitude compared with BSF faces (Goffaux et al., 2003; Holmes et al., 2005; Pourtois et al., 2005; Hsiao et al., 2005). In the present study, the N170 amplitude was clearly augmented by HSF faces. The discrepant results for N170 components among the present and previous ERP studies may arise from the SF cut-off values used to create the filtered faces or tasks during ERP recordings. First, the physical characteristics of the filtered photographs, such as luminance and contrast, were not strictly controlled in the previous studies. The HSF stimuli used in those studies seem to be lower contrast images than our stimuli (see Fig. 2C). Second, the subjects' task-relevant attention to faces can be a critical factor that affects the ERP 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). Therefore, we employed a simple passive viewing task with a target stimulus to avoid task-relevant attention to faces. Third, the definitions of LSF and HSF have been 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 have not taken into account whether the subjects can precisely distinguish between faces and objects under each SF condition. Therefore, we infer that the N170 amplitude may be decreased in the previous studies.

A recent fMRI study showed that HSF information of faces activated the FG to a much greater extent than BSF information (Vuilleumier et al., 2003a). Given that N170 is generated near the FG (Horowitz et al., 2004; Iidaka et al., 2006), these fMRI findings are consistent and linked with augmentation of the N170 amplitude. In addition, the amplitudes of N170 for HSF faces were much larger than those for HSF houses despite the fact that HSF faces contained less HSF information than HSF houses in the present study (Fig. 1). Taken together, we conclude that N170 reflects the encoding of facial features with right hemisphere predominance.

## 5. Conclusions

We found two face processing stages preceding the face-encoding N170 stage. P1 reflects face-selective holistic processing based on LSF information. The "P1 robustness" observed for PEL faces further assures this interpretation. N2, the "intermediator", relays facial information from P1 at the occipital region to N170 at the temporo-occipital region based on contrast information. N170 is related to the encoding of feature information of faces on the basis of HSF information. Spatially filtered and PEL images can be useful for studying the mechanisms of early processing for complex visual stimuli.

## 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 Communication 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., 2002. Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177.
- Allison, T., Ginter, H., McCarthy, G., Nobre, A.C., Puce, A., Luby, M., Spencer, D.D., 1994. Face recognition in human extrastriate cortex. *J. Neurophysiol.* 71, 821–825.
- Bentin, S., Allison, T., Puce, A., Perez, E., McCarthy, G., 1996. Electrophysiological studies of face perception in humans. *Cogn. Neurosci.* 8, 551–565.
- Bruce, V., Young, A.W., 1986. Understanding face recognition. *Br. J. Psychol.* 77 (Pt 3), 305–327.
- Calder, A.J., Young, A.W., Keane, J., Dean, M., 2000. Configural information in facial expression perception. *Exp. Psychol. Hum. Percept. Perform.* 26, 527–551.
- Eimer, M., 2000a. The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11, 2319–2324.
- Eimer, M., 2000b. Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Res. Cogn. Brain Res.* 10, 145–158.
- Eimer, M., McCarthy, R.A., 1999. Prosopagnosia and structural encoding of faces: evidence from event-related potentials. *Neuroreport* 10, 255–259.
- Farah, M.J., Wilson, K.D., Drain, M., Tanaka, J.N., 1998. What is "special" about face perception? *Psychol. Rev.* 105, 482–498.
- Fiorentini, A., Maffei, L., Sandini, G., 1983. The role of high spatial frequencies in face perception. *Perception* 12, 195–201.
- Ginsburg, A.P., 1988. Spatial filtering and visual form perception. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), *Handbook of Perception and Human Performance*. Wiley, New York.
- Goffaux, V., Gauthier, I., Rossion, B., 2003. Spatial scale contribution to early visual differences between face and object processing. *Cogn. Brain Res.* 16, 416–424.
- Goto, Y., Kinoc, H., Nakashima, T., Tobimatsu, S., 2005. Familiarity facilitates the corticocortical processing of face perception. *Neuroreport* 16, 1329–1334.
- Halit, H., de Haan, M., Schyns, P.G., Johnson, M.H., 2006. Is high-spatial frequency information used in the early stages of face detection? *Brain Res.* 1117, 154–161.
- Harmon, L.D., Julesz, B., 1973. Masking in visual recognition: effects of two-dimensional filtered noise. *Science* 180, 1194–1197.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Herrmann, M.J., Ehlis, A.C., Elgting, H., Fallgatter, A.J., 2004. Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *J. Neural Transm.* 112, 1073–1081.
- Hjorth, B., 1975. An on-line transformation of EEG scalp potentials into orthogonal source derivations. *Electroencephalogr. Clin. Neurophysiol.* 39, 526–530.
- Holmes, A., Winston, J.S., Eimer, M., 2005. The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Cogn. Brain Res.* 25, 508–520.
- Horowitz, S.G., Rossion, B., Skudlarski, P., Gore, J.C., 2004. Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage* 22, 1587–1595.
- Hsiao, F.J., Hsieh, J.C., Lin, Y.Y., Chang, Y., 2005. The effects of face spatial frequencies on cortical processing revealed by magnetoencephalography. *Neurosci. Lett.* 380, 54–59.
- Iidaka, T., Matsumoto, A., Haneda, K., Okada, T., Sadato, N., 2006. Hemodynamic and electrophysiological relationship involved in human face processing: evidence from a combined fMRI-ERP study. *Brain Cogn.* 60, 176–186.
- Itier, R.J., Taylor, M.J., 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *Neuroimage* 15, 353–372.
- Jeffreys, D.A., Tukmachi, E.S., 1992. The vertex-positive scalp potential evoked by faces and by objects. *Exp. Brain Res.* 91, 340–350.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Kobayashi, H., Kohshima, S., 1997. Unique morphology of the human eye. *Nature* 387, 767–768.
- Krolak-Salmon, P., Fischer, C., Vighetto, A., Mauguier, F., 2001. Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. *Eur. J. Neurosci.* 13, 987–994.

- Liddell, B.J., Brown, K.J., Kemp, A.H., Barton, M.J., Das, P., Peduto, A., Gordone, E., Williams, L.M., 2005. A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. *Neuroimage* 24, 235–243.
- Linkenkaer-Hansen, K., Palva, J.M., Sams, M., Hietanen, J.K., Aronen, H.J., Ilmoniemi, R.J., 1998. Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neurosci. Lett.* 253, 147–150.
- Liu, J., Harris, A., Kanwisher, N., 2002. Stages of processing in face perception: an MEG study. *Nat. Neurosci.* 5, 910–916.
- Livingstone, M., Hubel, D., 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749.
- 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 Parman Avenue, San Francisco, CA, 94143.
- Nakashima, T., Goto, Y., Abe, T., Kaneko, K., Saito, T., Makinouchi, A., Tobimatsu, S., 2008. Electrophysiological evidence for sequential discrimination of positive and negative facial expressions. *Clin. Neurophysiol.* 119, 1803–1811.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Pourtois, G., Dan, E.S., Grandjean, D., Sander, D., Vuilleumier, P., 2005. Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: time course and topographic evoked-potentials mapping. *Hum. Brain Mapp.* 26, 65–79.
- Rosson, B., Gauthier, I., Tarr, M.J., Despland, P., Bruyer, R., Linotte, S., Crommelinck, M., 2000. The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 11, 69–74.
- Sagiv, N., Bentin, S., 2001. Structural encoding of human and schematic faces: holistic and part-based processes. *J. Cog. Neurosci.* 13, 937–951.
- Schiller, P.H., Malpeli, J.G., Schein, S.J., 1979. Composition of geniculostriate input to superior colliculus of the rhesus monkey. *J. Neurophysiol.* 42, 1124–1133.
- Sugase, Y., Yamane, S., Ueno, S., Kawano, K., 1999. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400, 869–873.
- Tanaka, J.W., Fauth, M.J., 1993. Parts and wholes in face recognition. *Q. J. Exp. Psychol.* A 46, 225–245.
- Tanskanen, T., Nasanen, R., Montez, T., Paallysaho, J., Hari, R., 2005. Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cereb. Cortex* 15, 526–534.
- Tsurusawa, R., Goto, Y., Mitsudome, A., Nakashima, T., Tobimatsu, S., 2008. Different perceptual sensitivities for Chernoff's face between child and adults. *Neurosci. Res.* 60, 176–183.
- Tobimatsu, S., Celesia, G.G., 2006. Studies of human visual pathophysiology with visual evoked potentials. *Clin. Neurophysiol.* 117, 1414–1433.
- Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2003a. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat. Neurosci.* 6, 624–631.
- Vuilleumier, P., Mohr, C., Valenza, N., Wetzell, C., Landis, T., 2003b. Hyperfamiliarity for unknown faces after left lateral temporo-occipital venous infarction: a double dissociation with prosopagnosia. *Brain* 126, 889–907.
- Yin, R.K., 1969. Looking at upside-down faces. *J. Exp. Psychol.* 81, 141–145.



## Psychophysical and electrophysiological study on auditory temporal assimilation

Takako Mitsudo<sup>#1</sup>, Yoshitaka Nakajima<sup>†2</sup>, Gerard B. Remijn<sup>‡3</sup>, Hiroshige Takeichi<sup>§4</sup>,

Yoshinobu Goto<sup>¶5</sup>, Shozo Tobimatsu<sup>#6</sup>

<sup>#</sup>Department of Clinical Neurophysiology, Neurological Institute, Faculty of Medicine, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8581, Japan.

<sup>†</sup>Department of Acoustic Design, Faculty of Design, Kyushu University, Shiobaru, Minami-ku, Fukuoka, Japan.

<sup>‡</sup>Graduate School of Human and Socio-Environment Studies, Kanazawa University, Kanazawa, Japan.

<sup>§</sup>Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN

<sup>¶</sup>Department of Occupational Therapy, Faculty of Rehabilitation, International

University of Health and Welfare, Fukuoka, Japan

E-mail: <sup>1</sup>staka@neurophy.med.kyushu-u.ac.jp, <sup>2</sup>nakajima@desgin.kuushu-u.ac.jp, <sup>3</sup>remijn@staff.kanazawa-u.ac.jp, <sup>4</sup>takeichi@riken.jp, <sup>5</sup>ygoto@iuhw.ac.jp, <sup>6</sup>tobi@neurophy.med.kyushu-u.ac.jp

**Abstract** We examined the mechanisms of human auditory temporal perception using an illusion called 'auditory assimilation' [1]. Despite the robustness of the phenomenon, its underlying neural mechanisms have not been yet clarified. We used a 2-alternative-forced-choice (2AFC) duration discrimination task, where we systematically changed two neighboring time intervals. A correlation between behavioral data and event-related potentials (ERPs) was analyzed. Behavioral results showed a bilateral temporal assimilation; the subjects judged the two neighboring intervals as equal when  $-40 \leq T1-T2 \leq +40$  [ms]. Electrophysiological data showed a P300 component appeared in parietal area, and contingent negative variation (CNV)-like component emerged in the frontal area, at approximately 300 ms after T1. Human auditory temporal perception can be established by complex of several higher-order brain functions that are derived from broader brain areas.

**Keywords** Time perception, Assimilation, P300, CNV

### 1. Introduction

Our aim was to elucidate the mechanisms of human auditory temporal perception by taking up phenomena known as "auditory temporal assimilation" [1], [2]. In this study, we recorded event-related brain potentials (ERPs) while listeners judged the equality of two neighboring empty time intervals in various temporal patterns in order to examine the neural processing responsible for the temporal assimilation. Previous studies suggest that perceiving temporal intervals requires several brain functions such as working memory and attention [3]. ERP components related to those brain functions are often referred to CNV and P300. It enables us to know our cognitive temporal processing by examining how these components appear when the assimilation occurs.

### 2. Experiment

#### 2.1. Participants

Eight healthy volunteers (aged 21-35 years, 2 males and 6 females) participated in the experiment. All of

them were right handed and had no reported hearing deficits. Informed consent was obtained from each participant after an explanation of the purpose and procedures of the experiment.

#### 2.2. Apparatus and Stimuli

The experiment was conducted in an electrically shielded soundproof room (Yamaha Music Cabin, SC-3 and SC-5). The background noise was always kept below 30 dBA. Stimulus patterns were generated by J software (sampling frequency was 44.1 kHz) run on a Dell Dimension 4500C personal computer. Stimuli were presented by an AV Tachistoscope (Iwatsu, IS-703) diotically via headphones (Stax SR-303) to the participant through low-pass filters (NF DV8FL; cut-off frequency was 8 kHz) and a driver unit (Stax SRM-313). All stimulus patterns consisted of two neighboring time intervals marked by three successive 1-kHz tone bursts of 20 ms (5 ms rise/fall times and 10 ms plateau). The sound pressure level was 77 dBA (the level of a continuous tone

of the same amplitude). Levels were measured by a precision sound level meter (Node, type 2075), mounted with an artificial ear (Brüel and Kjær 4153). There were eleven stimuli: six standard stimuli where a first interval (T1) varied from 100 to 280 ms (i.e., 100, 120, 160, 200, 240, and 280 ms), and a second interval (T2) was fixed at 200 ms. There were two dummy stimuli (T1/T2: 240/240 and 200/100 ms) where the combinations of T1 and T2 differed from those of the standard stimuli (Fig. 1).

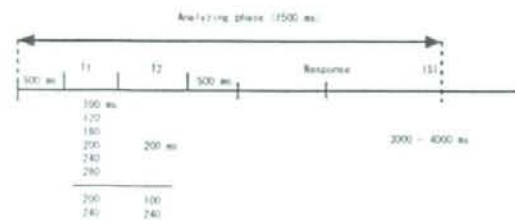


Fig. 1. Experimental procedures. Stimulus epochs began 500 ms prior to the stimulus and continued 1000.2 ms after the stimulus offset. Inter stimulus intervals (ISIs) were randomly varied from 3 to 5 seconds.

### 2.3. Procedure

The experiment consisted of two tasks. The experimental task was a 2-alternative-forced-choice (2AFC) duration discrimination. The participant judged whether the durations of T1 and T2 were the *same* or *different* by pressing a button as quickly as possible. The control task consisted of passive listening to the stimuli and pressing the button without making a judgment. For both the experimental and the control task, the standard and the dummy stimuli were presented about 96 times in pseudo-random order. The tasks were divided into 6 blocks of 60 trials (i.e., 8 stimuli  $\times$  7 or 8 presentations  $\times$  12 repetitions). Inter-stimulus intervals (ISIs) were randomly varied between 3 and 5 s. The allocation of button-pressing to the same or to the different was counterbalanced among participants. The participant first performed the control task and then the experimental task taking four separate days.

### 2.4. ERP recordings

ERPs were recorded by Electroencephalograph (EEG-1100 Neurofax, Nihon Koden) from 11 scalp locations (Fz, F3, F4, Cz, C3, C4, Pz, P3, P4, T3 and T4, international 10-20 system) with reference to an electrode at the nose tip. Horizontal and vertical electro-oculograms (EOGs) were also recorded using four electrodes placed over the outer canthi and in the superior and inferior areas of the orbit. Electrode impedance was kept below 5

k $\Omega$ . The ERP and EOG data were filtered with a bandpass of 0.27-300 Hz and sampled at a rate of 0.68 kHz. For the ERP analysis, stimulus epochs began 500 ms prior to, and continued 1000.2 ms after the stimulus onset (Fig. 1). The participant was instructed to close his/her eyes yet to stay alert. Trials that included artifacts such as  $\alpha$  waves, defined as waves for which the voltage exceeded  $\pm 50 \mu$  V at any electrode, were excluded from the analysis.

### 2.5. ERP analysis

For ERP analysis, we picked out frontal (Fz) central (Cz) and parietal (Pz) areas to check spatiotemporal characteristics of the ERP components related to the temporal judgment. We assumed that the differences in ERP components between the experimental task and the control task reflected the neural activities involved in the temporal judgments. Thus, for each standard stimulus, we compared the amplitudes and the latencies of CNV and P300 at the Fz and Pz electrodes, respectively.

## 3. Results

### 3.1. Behavioral data

#### 3.1.1. Response ratio

Fig. 2 shows the results of the same-different judgments. Participants judged the two neighboring time intervals as equal when  $-40 \leq T1-T2 \leq +40$  [ms]. The data from the experimental task were subjected to a one-way analysis of variance (ANOVA). There was a significant main effect of stimulus pattern [ $F(7, 49) = 42.5, p < .001$ ]. Dunnett's post-hoc *t*-tests revealed that the response ratios differed significantly from that obtained for the stimulus where  $T1=T2$  (200/200), when T1 was 100, 120, 280, or a dummy stimulus of 200/100.

This can be connected to our prediction that T1 should be perceived as equal to T2 when T1 is 120, 160, and 240 ms. The asymmetrical temporal assimilation indeed took place. This result confirmed the robustness of temporal assimilation and also proved that this illusion was not restricted to the experimental paradigms such as the method of adjustment employed in most of the previous psychophysical studies.

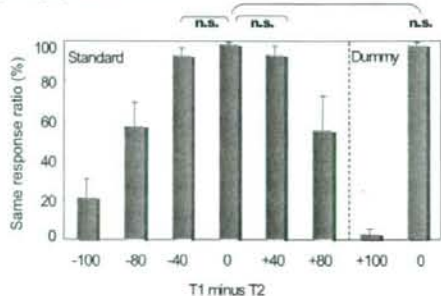


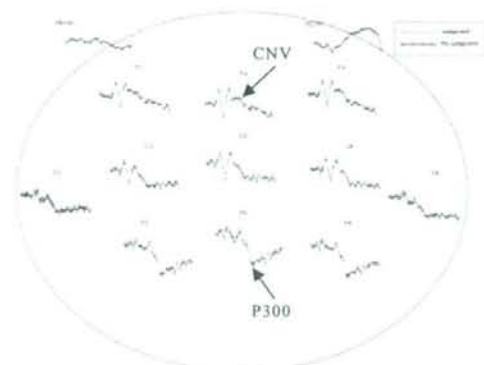


Fig. 2. Results of similarity judgments. Bar graphs show the ratio of similar judgments (i.e., T1 and T2 were perceived as having the same duration). The duration of T1 was perceived as equal when T1-T2 was -40 to +40 ms. The results indicate that bilateral temporal assimilation took place between T1 and T2.

### 3.2. ERP data

Electrophysiological data showed that the component related to judgment task was reflected on the front-central slow negative component and parietal positive component around 300 ms after T1 (Fig. 3). Each component can be defined as CNV and P300. In this experiment, the relative timing of the auditory ERPs was sensitive to the timing at which the second sound markers were presented; they emerged corresponding to the differences of T1.

(a) 100/200 (unequal)



(b) 200/200 (equal)

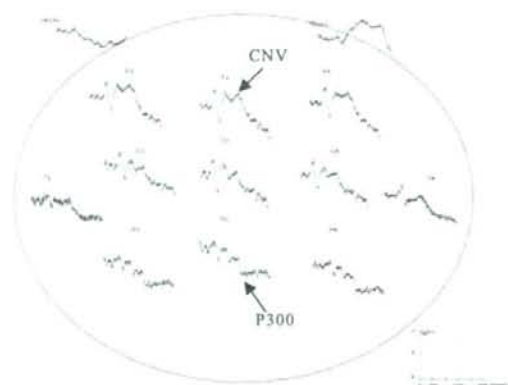


Fig. 3. Grand averaged ERP responses in 8 subjects that were elicited in a condition in which T1 minus T2 were -100 ms and 0 ms. (a) The two neighboring time intervals were perceived as

unequal. (a) The two neighboring time intervals were physically as well as perceptually as unequal. Solid lines represent the ERPs when participants made temporal judgments, while the red lines represent the control task, in which they listened to the stimuli passively. The CNV and P300 emerged in front central and parietal areas.

### 4. Discussion

The present experiment examined the neural processing of auditory temporal assimilation. Behavioral results showed a clearly bilateral temporal assimilation: participants judged the two neighboring time intervals as equal when  $-40 \leq T1-T2 \leq +40$  [ms]. This is in accordance with previous psychophysical findings showing that the temporal assimilation takes place when  $-80 \leq T1-T2 \leq +50$  [ms] [1]. In ERPs, two components such as CNV and P300 appeared in responding to the temporal judgment. Because CNV has been associated with "expectancy" [4] for the next stimulus, it emerged when participants attended to the first interval (T1) to compare it to the second interval (T2). P300 is widely known as the index of "attention". In our experiments, P300 latencies to the perceptually unequal stimuli were earlier than those to the perceptually equal stimuli. Moreover, its amplitude tended to be augmented in the perceptually unequal ones. These results suggested that participants took more cost when they judged the two neighboring intervals as unequal. The results could reflect the effect of attention on the judgment task, rather than the temporal judgment processing itself. Participants could discriminate the differences among stimuli only attending to the first and second marker (i.e., T1); it was unnecessary for the participants to attend to the whole pattern of the stimulus. This strategy may result in the difficulty of occurrence of auditory asymmetrical assimilation [5], expected to take place in this temporal range.

### 5. Acknowledgement

We are grateful to Sachiko Koyama and Hiroyuki Mitsudo for their helpful comments on this research. T.M. was supported by the JSPS Research Fellowships for Young Scientists. This study was supported in part by the Grant-in-aid for the 21st century COE program and Grants-in-aid for Scientists, No 16390253 and No 16200005 from the Ministry of Education, Culture, Sports, Science and Technology in Japan, and the 2008 Support Program for Research Activity from the Yamaha Music Foundation.

### References

- [1] R. Miyauchi, Y. Nakajima, Bilateral assimilation of two neighboring empty time intervals. *Music Perception*, vol.22, pp.411-424, 2005.
- [2] G. ten Hoopen, T. Sasaki, Y. Nakajima, G.B. Remijn, B. Massier, K.S. Rhebergen, W. Holleman, Time-shrinking and categorical temporal ratio perception: evidence for a 1:1 temporal category. *Music Perception*, vol.24, pp.1-22, 2006.
- [3] K. Rubia and A. Smith, The neural correlates of cognitive time management: a review. *Acta Neurobiol Exp*, vol. 64, pp. 329-340, 2004.
- [4] W.G. Walter, R. Cooper, V.J. Aldridge, W.C. McCallum, A.L. Winter, Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature*, vol.203, pp.380-384, 1964.
- [5] Y. Nakajima, G. ten Hoopen, T. Sasaki, K. Yamamoto, M. Kadota, M. Simons, and D. Suetomi Time-shrinking: the process of unilateral temporal assimilation. *Perception*, vol.33, pp.1061-1079, 2004.



## ミスマッチ陰性電位を用いた統合失調症 の視覚情報自動処理過程の検討

*A study on automatic visual information processing in schizophrenia using mismatch negativity*

前川 敏彦   平野 昭吾   大林 長二   平野 羊嗣  
*MAEKAWA Toshihiko   HIRANO Shogo   OBAYASHI Choji   HIRANO Yoji*

鬼塚 俊明   飛松 省三\*\*   神庭 重信\*  
*ONITSUKA Toshiaki   TOBIMATSU Shozo   KANBA Shigenobu*

今月のテーマ 統合失調症の事象関連電位



## ミスマッチ陰性電位を用いた統合失調症の視覚情報自動処理過程の検討

*A study on automatic visual information processing in schizophrenia using mismatch negativity*

前川 敏彦 平野 昭吾 大林 長二 平野 羊嗣  
 MAEKAWA Toshihiko HIRANO Shogo OBAYASHI Choji HIRANO Yoji

鬼塚 俊明 飛松 省三\*\* 神庭 重信\*  
 ONITSUKA Toshiaki TOBIMATSU Shozo KANBA Shigenobu

- ・統合失調症 (SZ) における視覚性ミスマッチ陰性電位 (v-MMN) と P300の特徴を分析した。
- ・v-MMN は早期成分 (MMN1) と後期成分 (MMN2) に分離された。健常対照群と比較して有意に MMN1 と P300 潜時が延長していた。
- ・MMN1/P300 潜時は年齢・服薬量と相関があり、P300 潜時は精神症状と相関を認めた。
- ・SZ の視覚情報処理過程の前注意 (自動) 処理過程と注意処理過程両方の異常が示唆された。

**KEY WORDS** 視覚情報自動処理, ミスマッチ陰性電位, 統合失調症

### はじめに

統合失調症 (Schizophrenia, SZ) は、双極性障害 (Bipolar disorder, BP) とならんで、現在でも原因不明の内因性精神疾患である。SZ に関する長年の研究にもかかわらず、臨床現場では症候学的診断にとどまり、客観指標に乏しい。恐らく SZ は生物学的に多様な集団で構成されていると考えられるので、SZ から生物学的に一様な群 (エンドフェノタイプ) を抽出するために、客観的な診断・治療マーカーを開発することは非常に重要である。事象関連電位 (ERP) は非侵襲かつ簡便なうえ、安価であるため日常臨床検査に適した手法である。かつて、自発脳波検査の進歩によっててんかんが SZ から分離されたように、ERP 異常に

よって特徴づけられる一群が SZ の中から分離されるかもしれない。本稿では特に、視覚刺激を用いた P300, ミスマッチ陰性電位 (MMN) による SZ の ERP 反応の特徴を述べる。

SZ の聴覚性 P300, 聴覚性 MMN (a-MMN), 視覚性 P300 は対象群 (NC) と比較するとどれも潜時は遅れ振幅も減衰している<sup>1)2)</sup>。一方、視覚性 MMN についてはこれまでその存在自体に議論があったため、SZ での記録はない。しかし、これらの電位が感覚情報処理のなかでも比較的後期の処理を反映していることを考えると、聴覚特異的というよりもさらに上位の認知機能メカニズム (トップダウン処理) に関連して誘発されており、他の感覚モダリティー刺激でも誘発できると考えられる<sup>3)4)</sup>。われわれは先行研究において厳密な意味での v-MMN を誘発し、健常人における機能

九州大学大学院医学研究院精神病態医学 \*教授 \*\*同 脳神経病研究施設臨床神経生理学 教授  
 Address/MAEKAWA T: Dept. of Neuropsychiatry and Clinical Neurophysiology, Graduate School of Medical Sciences, Kyushu University, FUKUOKA 812-8582



特性を明らかにした<sup>5)</sup>。その機能特性は、注意非依存性、内因性、感覚モダリティーに一致した発生源、逸脱依存性、感覚記憶関連電位、注意シフト関連電位などを含んでいた。また、a-MMNと同様に前注意刺激情報処理を反映するという特徴とともにa-MMNでは分離しづらい、刺激性状処理関連電位と注意シフト関連電位が二つピークをなして明瞭に分離されていた。これらの健常人の成績を基に同じ実験条件でSZを対象に視覚性P300とv-MMNを計測し、比較した。

## ■ ■ 方 法

### 1. 対 象

九州大学病院精神科神経科にて治療中のSZ患者18人(男性15人、左利き1人、21~54(平均35.8)歳)。診断は構造化面接を行い、精神疾患の分類と診断の手引き(DSM-IV)に基づき診断した。九州大学病院臨床研究等倫理委員会で承認を得た手続きにより、文書および口頭でインフォームドコンセントを行い、本人かつ主治医の同意が得られたものを対象者とした。

### 2. 患者評価

患者評価項目は、年齢、性別、利き手、発症年齢、罹病期間、服薬量(chlorpromazine換算)、精神症状(Positive and negative syndrome scale; PANSSを用いた)であった。検者側のバイアスを最小にするため評価は脳波解析には関与しない2名の精神科医によって行われた。

### 3. 実験条件

被験者には防音された暗室の中の安楽椅子に座り、イヤホンからの物語に注意集中しながら前方の20-inch モニタ画面中心を固視するように指示した。正面のモニタ画面中心には視角5.8度の2種類のウインドミルボタン(標準刺激、逸脱刺激)と同じ大きさの白色円形刺激(標的刺激)を刺激呈示時間200ms、刺激間隔800ms、8:1:1の割合でランダムに呈示し、標的刺激でボタン

を押すように指示した(図1)。行動指標として、標的刺激に対するボタン押しの正答率、反応時間を計測し、実験終了時に物語の内容に関する質問紙(4者択一、10問)により被験者の注意が物語の内容と標的刺激の同定に向いていたことを確認した。実験中は持続して脳波を記録した。

### 4. 記録と解析

国際10-20法に従って、頭皮上10箇所(Fz, Cz, Pz, Oz, P3, P4, O1, O2, T5, T6)に記録電極を配置し、サンプリング周波数833Hz、周波数帯域0.05~200HzでERPを測定した。基準電極は鼻尖部に置いた。得られたデータはオフライン処理し、刺激ごとに加算平均を行い、逸脱刺激に対するERPから標準刺激に対するERPを引算してv-MMNを抽出し、標的刺激に対する反応からP300を同定した。SZ群のv-MMN、P300の大きさ、潜時を以前に行った9人のNC群(男性4人、左利き1人、19~31(平均24)歳)と比較検討し、群間で有意差を認めた場合は、SZ群内で患者評価項目との相関を検定した。

## ■ ■ 結 果

患者評価結果は表1のとおりであった。18例中1例(患者No.18)は気分不良により脳波実験を最後まで遂行できなかった。1例(患者No.9)は $\alpha$ 波の混入が多量で平均波形からN1が同定できなかったため解析から除外した。残りの16人のうちP300は全員誘発されたが、4人(No.1, 3, 5, 10)はv-MMNが誘発されなかった。

### 1. 行動指標

物語質問紙の正答率はNC群と比較してSZ群で有意に低下していた( $0.97 \pm 0.04$  vs.  $0.70 \pm 0.31$ ,  $P=0.002$ )。ボタン押しの正答率もSZ群の方が低下していたが、有意差はなかった( $0.97 \pm 0.05$  vs.  $0.93 \pm 0.10$ )。反応時間はSZ群の方が有意に遅かった( $385.6 \pm 24.7$  vs.  $451.8 \pm 85.9$  ms,  $P=0.016$ )。

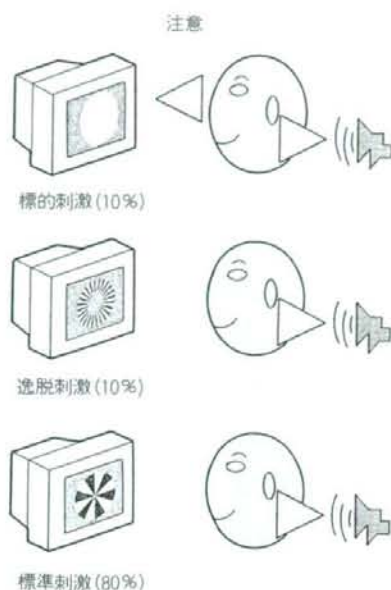
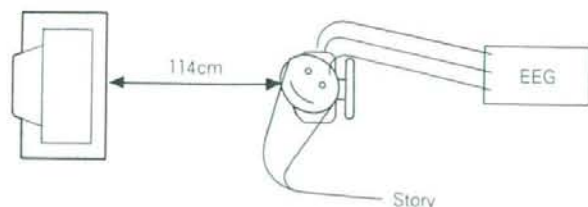


図1 実験条件

被験者には防音された暗室の中の安楽椅子に座りイヤホンからの物語に注意集中しながら前方の20-inch モニタ画面中心を固視するように指示した(上図)。正面のモニタ画面中心には視角5.8度の2種類の白黒円形ウィンドミルボタン(標準刺激(S)、逸脱刺激(D))と同じ大きさの白色円形刺激(標的刺激(T))を刺激呈示時間200ms、刺激間隔800ms、8:1:1の割合でランダムに呈示し、Tでボタンを押すように指示した。行動指標として、Tに対するボタン押しの正答率、反応時間を計測し、実験終了時に物語の内容に関する質問紙(4者択一、10問)により被験者の注意が物語の内容とTの同定に向いていたことを確認した。実験中は持続して脳波を記録した。

表 1

患者No	性別	年齢 (歳)	罹病期間 (年)	服薬量 (CP換算, mg)	PANSS		
					P	N	G
1	M	21	1	80	18	22	47
2	F	22	1	400	29	27	61
3	F	26	10	235	15	32	40
4	M	26	9	555	37	42	83
5	M	27	7	175	18	21	50
6	M	28	9	85.5	30	17	68
7	M	29	1	262.5	10	36	46
8	M	31	6	1211.1	35	21	65
9	F	32	9	80	24	40	75
10	M	33	7	450	15	41	60
11	M	36	11	972.5	18	18	41
12	M	43	14	1200	23	37	9
13	M	44	22	500.7	9	14	25
14	M	44	22	390	24	35	62
15	M	46	20	515.9	41	38	74
16	M	48	12	80	16	31	49
17	M	54	19	1022.4	35	39	79
18	M	54	39	862.5	33	40	77

略語: M: Male, F: Female, CP: chlorpromazine, PANSS: Positive and negative syndrome scale, P: 陽性尺度, N: 陰性尺度, G: 総合精神病理尺度