



## Inaudible High-Frequency Sounds Affect Brain Activity: Hypersonic Effect

Tsutomu Oohashi, Emi Nishina, Manabu Honda, Yoshiharu Yonekura, ... [Show all Authors](#) ▾

Published Online: 01 JUN 2000 // <https://doi.org/10.1152/jn.2000.83.6.3548>

### Abstract

Although it is generally accepted that humans cannot perceive sounds in the frequency range above 20 kHz, the question of whether the existence of such “inaudible” high-frequency components may affect the acoustic perception of audible sounds remains unanswered. In this study, we used noninvasive physiological measurements of brain responses to provide evidence that sounds containing high-frequency components (HFCs) above the audible range significantly affect the brain activity of listeners. We used the gamelan music of Bali, which is extremely rich in HFCs with a nonstationary structure, as a natural sound source, dividing it into two components: an audible low-frequency component (LFC) below 22 kHz and an HFC above 22 kHz. Brain electrical activity and regional cerebral blood flow (rCBF) were measured as markers of neuronal activity while subjects were exposed to sounds with various combinations of LFCs and HFCs. None of the subjects recognized the HFC as sound when it was presented alone. Nevertheless, the power spectra of the alpha frequency range of the spontaneous electroencephalogram (alpha-EEG) recorded from the occipital region increased with statistical significance when the subjects were exposed to sound containing both an HFC and an LFC, compared with an otherwise identical sound from which the HFC was removed (i.e., LFC alone). In contrast, compared



when an HFC and an LFC were presented together, the rCBF in the brain stem and the left thalamus increased significantly compared with a sound lacking the HFC above 22 kHz but that was otherwise identical. Simultaneous EEG measurements showed that the power of occipital alpha-EEGs correlated significantly with the rCBF in the left thalamus.

Psychological evaluation indicated that the subjects felt the sound containing an HFC to be more pleasant than the same sound lacking an HFC. These results suggest the existence of a previously unrecognized response to complex sound containing particular types of high frequencies above the audible range. We term this phenomenon the “hypersonic effect.”

## INTRODUCTION

It is generally accepted that audio frequencies above 20 kHz do not affect human sensory perception since they are beyond the audible range ([Durrant and Lovrinc 1977](#); [Snow 1931](#); [Wegel 1922](#)). Thus for example, most of the conventional commercial digital audio formats [e.g., compact disks (CDs), digital audio tapes (DATs), and digital audio broadcasting] have been standardized to a frequency range that does not allow such high-frequency components (HFCs) of sounds to be included. As a premise for determining these formats, several psychological experiments were performed to evaluate sound quality subjectively by means of questionnaires, according to the recommendation of the Comité Consultatif International Radiophonique ([CCIR 1978](#)) or its modified versions. Studies by [Muraoka et al. \(1978\)](#) and [Plenge et al. \(1979\)](#), as well as other studies, concluded that listeners did not consciously recognize the inclusion of sounds with a frequency range above 15 kHz as making a difference in sound quality. Nevertheless, and interestingly enough, artists and engineers working to produce acoustically perfect music for commercial purposes are convinced that the intentional manipulation of HFC above the audible range can positively affect the perception of sound quality ([Neve 1992](#)). Indeed, the Advanced Audio Conference organized by the [Japan Audio Society \(1999\)](#) proposed two next-generation advanced digital audio formats: super audio compact disk (SACD) and digital versatile disk audio (DVD-audio). These formats have a frequency response of up to 100 kHz and 96kHz, respectively. However, the proposal was not based on scientific data about the biological effects of the HFCs that would become available with these advanced formats. Although recently there have been several attempts to explore the psychological effect of inaudible HFCs on sound perception using a digital audio format with a higher sampling rate of 96 kHz ([Theiss and Hawksford 1997](#); [Yamamoto 1996](#); [Yoshikawa et al. 1995,1997](#)), none of



determining sound quality, which is solely based on a subjective evaluation obtained via questionnaires.

There are two factors that may have some bearing on this issue. First, it has been suggested that infrasonic exposure may possibly have an adverse effect on human health (Danielsson and Landstrom 1985), suggesting that the biological sensitivity of human beings may not be parallel with the “conscious” audibility of air vibration. Second, the natural environment, such as tropical rain forests, usually contains sounds that are extremely rich in HFCs over 100 kHz. From an anthropogenetic point of view, the sensory system of human beings exposed to a natural environment would stand a good chance of developing some physiological sensitivity to HFCs. It is premature to conclude that consciously inaudible high-frequency sounds have no effect on the physiological state of listeners.

In the present study, therefore, we addressed this issue by using quantifiable and reproducible measurements of brain activity. To measure human physiological responses to HFCs, we selected two noninvasive techniques: analysis of electroencephalogram (EEG) and positron emission tomography (PET) measurements of the regional cerebral blood flow (rCBF). These methods have complementary characteristics. EEG has excellent time resolution, is sensitive to the state of human brain functioning, and places fewer physical and mental constraints on subjects than do other techniques such as functional magnetic resonance imaging (fMRI). This is of special importance because some responses might be distorted by a stressful measurement environment itself. On the other hand, PET provides us with detailed spatial information on the neuroanatomical substrates of brain activity. Combining these two techniques with psychological assessments, we provide evidence herein that inaudible high-frequency sounds have a significant effect on humans.

## METHODS Subjects

Twenty-eight Japanese volunteers (15 males and 13 females, 19–43 years old) participated in the EEG experiments; 12 Japanese volunteers (8 males and 4 females, 19–34 years old) participated in the PET experiment; and 26 Japanese volunteers (15 males and 11 females, 18–31 years old) participated in the psychological experiment. None of the subjects had any history of neurological or psychiatric disorders. Written informed consent was obtained from all subjects before the experiments. The PET and EEG experiments were performed in accordance with the approval of the Committee of Medical Ethics, Graduate School of



## Sound materials and presentation systems

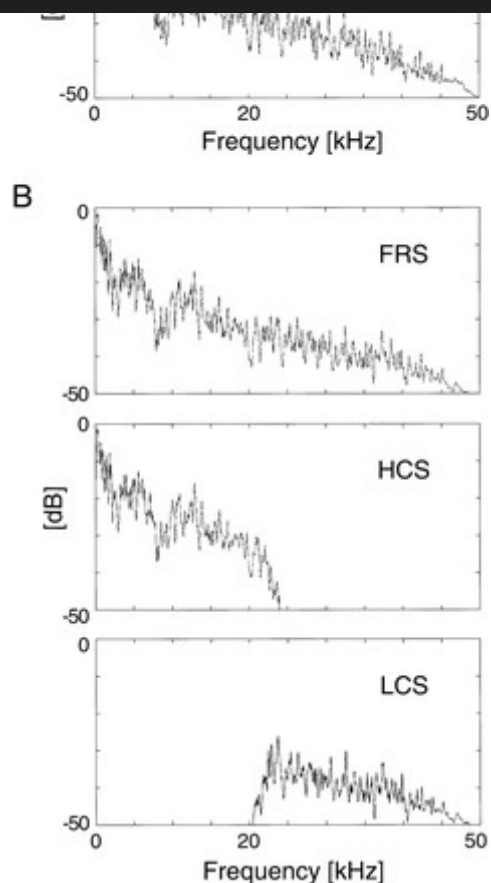
Traditional gamelan music of Bali Island, Indonesia, a natural sound source containing the richest amount of high frequencies with a conspicuously fluctuating structure, was chosen as the sound source for all experiments. A traditional gamelan composition, “Gambang Kuta,” played by “Gunung Jati,” an internationally recognized gamelan ensemble from Bali, was recorded using a B&K 4135 microphone, a B&K 2633 microphone preamplifier, and a B&K 2804 power supplier, all manufactured by Brüel and Kjær (Nærum, Denmark). The signals were digitally coded by Y. Yamasaki's high-speed one-bit coding signal processor (United States Patent No. 5351048) ([Yamasaki 1991](#)) with an A/D sampling frequency of 1.92 MHz and stored in a DRU-8 digital data recorder (Yamaha, Hamamatsu, Japan). This system has a generally flat frequency response of over 100 kHz.

Most of the conventional audio systems that have been used to present sound for determining sound quality were found to be unsuitable for this particular study. In the conventional systems, sounds containing HFCs are presented as unfiltered source signals through an all-pass circuit and sounds without HFCs are produced by passing the source signals through a low-pass filter ([Muraoka et al. 1978](#); [Plenge et al. 1979](#)). Thus the audible low-frequency components (LFCs) are presented through different pathways that may have different transmission characteristics, including frequency response and group delay. In addition, inter-modulation distortion may differentially affect LFCs. Therefore it is difficult to exclude the possibility that any observed differences between the two different sounds, those with and those without HFCs, may result from differences in the audible LFCs rather than from the existence of HFCs. To overcome this problem, we developed a bi-channel sound presentation system that enabled us to present the audible LFCs and the nonaudible HFCs either separately or simultaneously. First, the source signals from the D/A converter of Y. Yamasaki's high-speed, one-bit coding signal processor were divided in two. Then, LFCs and HFCs were produced by passing these signals through programmable low-pass and high-pass filters (FV-661, NF Electronic Instruments, Tokyo, Japan), respectively, with a crossover frequency of 26 or 22 kHz and a cutoff attenuation of 170 or 80 dB/octave, depending on the type of test. Then, LFCs and HFCs were separately amplified with P-800 and P-300L power amplifiers (Accuphase, Yokohama, Japan), respectively, and presented through a speaker system consisting of twin cone-type woofers and a horn-type tweeter for the LFCs and a dome-type super tweeter with a diamond diaphragm for the HFCs. The



response of over 100 kHz. The level of the presented sound pressure was individually adjusted so that each subject felt comfortable; thus the maximum level was approximately 80–90 dB sound pressure level (SPL) at the listening position.

Using the bi-channel sound presentation system, four different sound combinations were prepared as follows: 1) full-range sound (FRS) = HFC + LFC; 2) high-cut sound (HCS) = LFC only; 3) low-cut sound (LCS) = HFC only; and, 4) baseline = no sound except for ambient noise. All experiments were performed in an acoustically shielded room. In the PET experiment, there was a very low-level fan noise from the PET scanner, which did not annoy the subjects. Figure 1A shows the averaged power spectrum of the source signal obtained from the music with a CF-5220 fast Fourier transform (FFT) analyzer (Ono Sokki, Tokyo, Japan) over an analysis period of 200 s. It contained a significant amount of HFCs above the audible range, often exceeding 50 kHz and, at certain times, 100 kHz. Figure 1B shows the averaged power spectra of the actual sounds reproduced with a 22 kHz cutoff frequency for the filter and recorded at the subject's head position. The spectrum of FRS was essentially the same as that of the source and contained both LFCs below and HFCs above 22 kHz. None of the blindfolded subjects could distinguish LCS (i.e., HFC only) from silence when it was presented alone. Therefore we concluded that the HFC employed in the present experimental setting was, at least, a consciously unrecognizable air vibration.



**Fig. 1.**

Power spectra of the sound used in this study. *A*: the averaged power spectrum calculated from the entire 200-s period of the recorded sound source signal using a CF-5220 fast Fourier transform (FFT) analyzer (Ono Sokki, Tokyo, Japan). It contains a significant amount of high-frequency components above the audible range. *B*: the averaged power spectra of the sounds reproduced by the bi-channel sound presentation system (see text) in different conditions. The power was calculated from the signal actually recorded at the subject's head position using a B&K 4135 microphone (Brüel and Kjær, Nærum, Denmark). The *top*, *middle*, and *bottom* panels represent full-range sound (FRS), high-cut sound (HCS), and low-cut sound (LCS), respectively. The power spectrum of FRS is essentially identical to the spectrum of the source and contains both a low-frequency component (LFC) (i.e., the one used in the HCS condition) and a high-frequency component (HFC) (in the LCS condition).

[Download figure](#) | [Download PowerPoint](#)

## EEG recordings and analysis



distance from the speakers to the subjects' ears was approximately 2.5 m. Special attention was paid to the subjects' immediate environment to avoid discomfort. For example, the room was decorated with plants, lacquered masks, and landscape paintings. The equipment for the EEG recordings was hidden from the subjects' view and all cables for the experimental equipment were in a pit below the floor. The subjects were instructed to enjoy the music without any cognitive tasks during the sound presentation. The subjects were able to see outdoors through a wide, double-glass window that acoustically shielded the experimental room from outside sounds. Two different EEG experiments were performed. In the first experiment, to explore the physiological effect of sounds with a nonaudible frequency range, we employed a strictly controlled experimental setting of sound presentation combined with conventional EEG measurements. In the second experiment, the same effect was examined under more ordinary listening conditions.

### ***EXPERIMENT 1.***

To examine the physiological effect of sounds with an inaudible frequency range, 11 subjects were presented with the FRS, HCS, and baseline conditions. In this experiment, a cutoff frequency of 26 kHz with a steeper cutoff attenuation of 170 dB/octave was employed to separate HFCs from LFCs. This relatively high cutoff frequency was chosen because when a cutoff frequency lower than 26 kHz is used the skirts of the power spectrum of the filtered HFCs extend below 20 kHz and generate sounds containing components below 20 kHz. It is widely known that the upper limit of the audible range of humans varies considerably. It usually corresponds to around 15 or 16 kHz in young adults and sometimes below 13 kHz in the elderly, and some people can recognize air vibrations of 20 kHz as sound. When a cutoff frequency of 26 kHz is employed with the steeper cutoff attenuation, the power spectrum of the filtered HFCs under 20 kHz falls below the system noise level. Therefore we selected a cutoff frequency of 26 kHz, which is sufficiently high to completely exclude contamination by audible sound components in all of the subjects. In accordance with conventional recordings of background EEG activity, subjects were asked to keep their eyes naturally closed during the experiment to eliminate any effects of visual input. The presentation of the sounds in both FRS and HCS conditions lasted 200 s, which included the entire piece of music. The baseline condition also lasted 200 s without sound presentation. The inter-session intervals were 10 s. Two recording sessions were repeated for each condition in the following order: baseline–FRS–HCS–FRS–HCS–baseline.

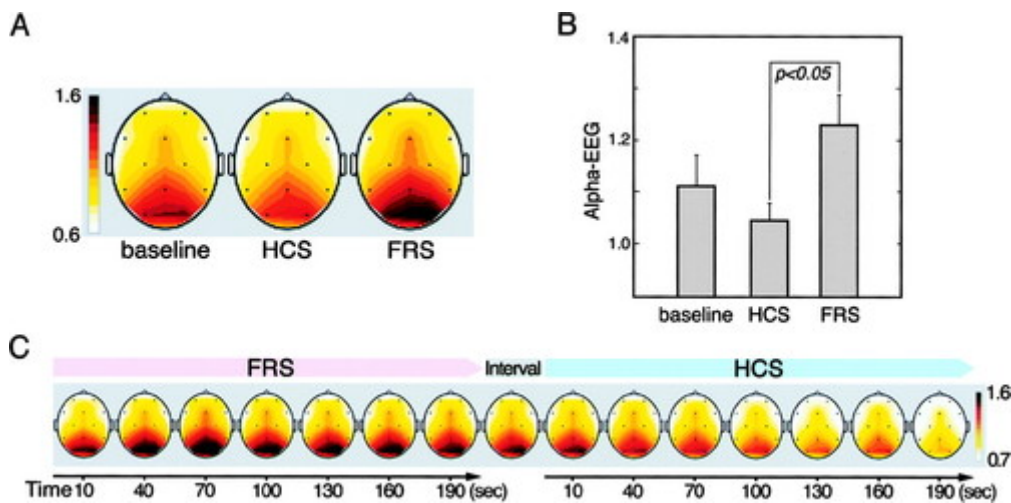


The validity of the digital audio format internationally employed for CDs was evaluated under more ordinary listening conditions. Seventeen subjects were presented with sounds using a cutoff frequency of 22 kHz, which corresponds to the upper range of sounds recorded by a CD. Subjects were then asked to keep their eyes naturally open as they usually do when they listen to music. The open-eye condition was also appropriate to control the subjects' vigilance. Each subject was presented with four types of conditions: FRS, HCS, and baseline, as in Experiment 1, plus LCS to elucidate the effect of an HFC when it is presented alone. As in Experiment 1, each condition lasted 200 s. Before the actual recording sessions, HCS was presented once to familiarize the subjects with the experimental environment. To avoid any influence by the order of presentation, the four different conditions were performed in random order across the subjects. After a 10-min rest, the same four conditions were repeated in reverse order. Neither the subjects nor the experimenters knew which conditions were being performed.

The EEGs, recorded using the WEE-6112 telemetric system (Nihon-Koden, Tokyo, Japan) to minimize constraint on the subjects, were stored on magnetic tape for off-line analysis. The EEGs were recorded continuously, including the intervals between the sessions. Data were recorded from 12 scalp sites (Fp1, Fp2, F7, Fz, F8, C3, C4, T5, Pz, T6, O1, and O2 according to the International 10-20 System) using linked earlobe electrodes as the reference with a filter setting of 1–60 Hz (–3 dB). The impedance of all electrodes was kept below 5 k $\Omega$ . The EEGs obtained were subjected to power spectra analysis. The power spectrum of the EEG at each electrode was calculated by fast Fourier transform (FFT) analysis for every 2-s epoch, with an overlap of 1 s, at a frequency resolution of 0.5 Hz with a sampling frequency of 256 Hz. Then the averaged power spectrum within a 10-s time window was calculated. Each analysis window was designated by the time at its middle point measured from the beginning of the sound presentation. For example, the time window labeled as 100-s contains data from 95 to 105 s from the beginning. Then the square root of the averaged power level in a frequency range of 8.0–13.0 Hz at each electrode position was calculated as the equivalent potential of EEGs in an alpha band (alpha-EEG). To eliminate a possible effect of inter-subject variability, the alpha-EEG at each electrode position was normalized with respect to the mean value across all time epochs, conditions, and electrode positions for each subject. To obtain an overview of the data, to check for contamination by artifacts, and to characterize the spatial distribution of the alpha-EEG, we constructed colored contour line maps using 2,565 scalp grid points with linear interpolation and extrapolation. This type of



by averaging the alpha-EEGs at the electrodes on the posterior one-third of the scalp. The BEAMs and occipital alpha-EEGs were averaged over multiple time epochs and subjected to a statistical evaluation of condition effects. Since the time course of the alpha-EEG change revealed a considerable time lag with respect to the sound presentation (see RESULTS and Fig.2C), we made a statistical evaluation of the data obtained from all time epochs as well as of the data from only the latter half of the session (from the 100-s to 200-s class marks). We used analysis of variance (ANOVA) followed by Fishers' protected least significant difference (PLSD) post hoc test to assess statistical significance for the different conditions.



**Fig. 2.**

Normalized potentials from the alpha frequency range of the spontaneous electroencephalogram (alpha-EEG) under each experimental condition (FRS, HCS, and baseline) and time course in the successive FRS and HCS conditions in EEG Experiment 1. *A*: brain electrical activity maps (BEAMs) averaged across the 11 subjects over the entire time epoch of sound presentation. Darker red indicates higher alpha-EEG potential. Note that the alpha-EEG is enhanced in the parieto-occipital region exclusively in the FRS condition. *B*: mean and standard error of the occipital alpha-EEG for all 11 subjects. FRS significantly enhanced the occipital alpha-EEG relative to HCS. *C*: time course of grand average BEAMs across all 11 subjects. Two sessions for each condition were averaged in this figure. The occipital alpha-EEG shows a gradual increase during the FRS presentation and a gradual decrease while HCS was successively presented.

The sound presentation equipment was installed and calibrated in the PET laboratory of Kyoto University Hospital. Subjects lay supine, with their eyes naturally open, on the PET scanner bed in a quiet, dimly lit room. Their heads were fixed in individually molded helmet-shaped rests that were contoured to leave their ears undisturbed. The distance from the speakers to the subjects' ears was approximately 1.5 m. As in the EEG study, special attention was paid to the immediate environment to minimize the subjects' discomfort. Six of the subjects were studied using FRS, HCS, and baseline conditions, and the other six were studied using FRS, LCS, and baseline conditions. The order of the conditions was randomized across the subjects and a total of six scans was performed on each subject with intervals of 7 min. For each of the FRS, HCS, and LCS presentations, 30 mCi of  $^{15}\text{O}$ -labeled water was injected into the right cubital vein 80 s after the beginning of each session. The same procedure was carried out for the baseline condition after a minimum 1-min rest without any presentation other than the ambient background noise of the PET scanner room. Following the injection, the head was scanned for radioactivity with a multi-slice PET scanner (PCT3600W, Hitachi Medical Co., Tokyo, Japan) for 120 s. The scanner acquired 15 slices with a center-to-center distance of 7 mm and an axial resolution of 6.5 mm full-width at half-maximum (FWHM) at the center ([Endo et al. 1991](#)). The in-plane spatial resolution with stationary mode acquisition used in this protocol was 6.7 mm of FWHM, which was blurred to ~10 mm in the reconstructed PET images. The field of view and pixel size were 256 mm and  $2 \times 2$  mm, respectively. Prior to the emission measurements, transmission data were obtained using a  $^{68}\text{Ge}/^{68}\text{Ga}$  standard plate source for attenuation correction. Reconstructed images were obtained by summing up the activity throughout the 120-s period. No arterial blood sampling was performed; therefore the images collected were of tissue activity. Tissue activity recorded by this method is linearly related to rCBF ([Fox et al. 1984](#); [Fox and Mintun 1989](#)).

The PET data were analyzed with statistical parametric mapping (SPM96 software, Wellcome Department of Cognitive Neurology, London, UK) implemented in MATLAB (Mathworks, Inc., Sherborn, MA). Statistical parametric maps are spatially extended statistical processes that are used to characterize regionally specific effects in imaging data ([Friston et al. 1991, 1994, 1995b](#); [Worsley et al. 1992](#)). The scans from each subject were realigned using the first image as the reference ([Friston et al. 1995a](#)). After realignment, the images were transformed into a standard anatomical space ([Friston et al. 1995a](#); [Talairach and Tournoux 1988](#)). As a result, each scan was resampled into voxels that were  $2 \times 2 \times 4$

variation in normal gyral anatomy and to increase signal-to-noise ratio. The effect of global differences in rCBF between scans was removed by scaling the activity in each pixel proportional to the global activity so as to adjust the mean global activity of each scan to 50 ml/100g/min. To explore regions showing significant differences in rCBF among different conditions, the general linear model with contrasts was employed at each voxel (Friston et al. 1995b). Since the different conditions were run in different subjects, the contrasts of FRS versus HCS and HCS versus baseline were examined for six subjects, and those of FRS versus LCS and LCS versus baseline were examined for the other six subjects. The contrast of FRS versus baseline was examined for all 12 subjects, inclusive. The resulting set of voxel values for each contrast constituted a statistical parametric map of the  $t$  statistic. The  $t$  values were transformed into the unit normal distribution ( $Z$  score), which was independent of the degree of freedom of error, and were thresholded at 3.09. To account for multiple non-independent comparisons, the significance of the activation in each brain region detected was estimated by the use of distributional approximations from the theory of Gaussian fields in terms of spatial extent and/or peak height (Friston et al. 1994). An estimated  $P$  value of 0.05 was used as a final threshold for significance. The resulting set of  $Z$  scores for the significant brain regions was mapped onto a standard spatial grid (Talairach and Tournoux 1988).

In all of the subjects, EEGs were simultaneously recorded throughout the PET measurement, which lasted approximately 60 min, from 12 electrodes as in the EEG experiment. The EEGs obtained during the total 200-s sound presentation were subjected to power spectra analysis and, in particular, those during each 120-s PET scan were used for correlation analysis with the rCBF. The data of one subject were excluded because of an excessive amount of electrical noise in the EEG. We used ANOVA followed by Fisher's PLSD post hoc test to assess the statistical significance of the different conditions. In addition, we used SPM software to calculate a correlation map between rCBF and the occipital alpha-EEG, to examine the relationship between them. An estimated  $P$  value of 0.05 with correction for multiple comparisons was used as the final threshold for significance.

## Psychological evaluation of sound quality

We also evaluated the subjective perception of sound quality. Since the subjective impression of sounds is closely related to the subjects' psychological condition, this evaluation was performed separately from the EEG and PET experiments. We used the

randomized across the subjects. After an intermission of 3 min, another pair of FRS and HCS was presented in reverse order. Therefore the stimuli were presented in an A-B-B-A fashion, in which FRS and HCS were assigned to A and B or B and A, respectively, in a randomly counterbalanced way across the subjects. Neither the subjects nor the experimenter knew what the sound conditions were, although they did know that the presentation was in an A-B-B-A fashion. The subjects filled out a questionnaire to rate the sound quality in terms of 10 elements, each expressed in a pair of contrasting Japanese words (e.g., soft vs. hard). Each element of each condition was graded on a scale of 5 to 1. The scores were statistically evaluated by the paired comparison method described by [Scheffé \(1952\)](#). Note that the method used in the present study differs from that recommended by the [CCIR \(1978\)](#) and its modified version, which were widely used to determine the digital format of CDs around 1980 (e.g., [Muraoka et al. 1978](#); [Plenge et al. 1979](#)). In the previous studies, sound materials were never longer than 20 s and the interval between two successive sound materials was 2–3 s or less. Therefore if neuronal response to sound stimuli is characterized by delay and persistence for longer than 20 s, it is difficult to exclude the possibility that those studies might have introduced a subjective evaluation that might not precisely correspond to each sound condition.

## RESULTSEEG Experiment 1

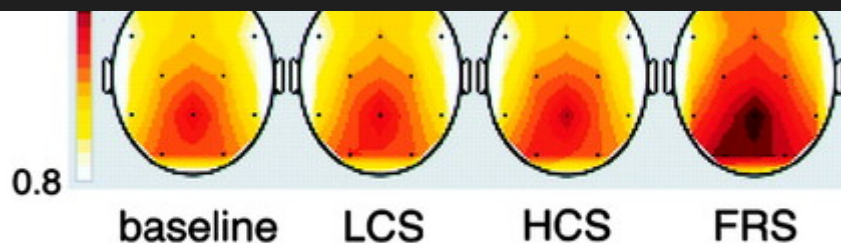
Figure 2, *A* and *B*, shows the grand average BEAMs and occipital alpha-EEGs, respectively, for the 11 subjects, calculated over the entire period of the sound presentation. The alpha-EEGs were enhanced during FRS compared with those during the other conditions. This enhancement was especially predominant in the occipital and parietal regions (Fig. 2*A*). ANOVA on the occipital alpha-EEG revealed a significant main effect of condition [ $F(2,63) = 3.74$ ,  $P < 0.05$ ]. The post hoc tests showed that the occipital alpha-EEG during FRS was significantly greater than that during HCS ( $P < 0.05$ ) (Fig. 2*B*). There was a similar tendency when FRS was compared with the baseline ( $P = 0.10$ ). Figure 2*C* shows the averaged time course of the BEAMs calculated for each 30 s of the FRS and HCS conditions for all subjects, inclusive. The alpha-EEG showed a gradual increase during the first several tens of seconds of FRS; there was a gradual decrease at the beginning of the following HCS. Taking into account the delay and persistence of the enhancement of the alpha-EEG, statistical evaluation was also made of the data from the latter half of the recording session (from the 100-s to 200-s class mark). In this analysis, compared with the data obtained by analyzing



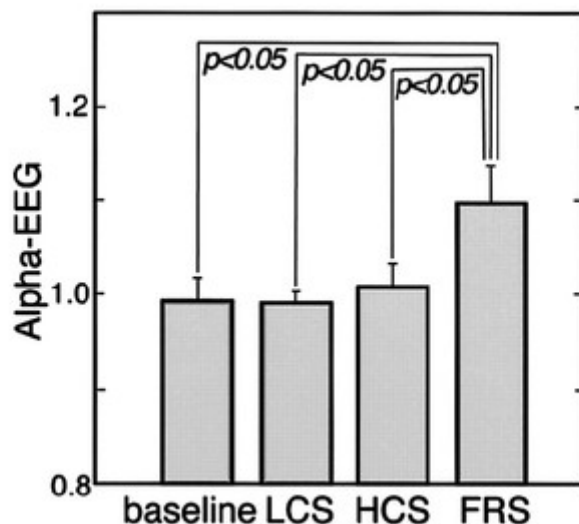
between FRS and HCS ( $P < 0.01$ ).

## EEG Experiment 2

The grand average BEAMs and occipital alpha-EEGs across all 17 subjects over the latter half of the session (from the 100-s to 200-s class mark) are shown in Fig. 3. The amount of eye movement did not differ for different conditions. The alpha-EEG showed significant enhancement in FRS compared with the other conditions (Fig. 3A). This enhancement was predominant in the occipital and parietal regions. ANOVA on the occipital alpha-EEG revealed a significant main effect of condition [ $F(3,131) = 3.74, P < 0.05$ ]. The post hoc tests showed that the occipital alpha-EEG in FRS was significantly greater than that in the other three conditions (Fig. 3B). There was no significant difference among HCS, LCS, and baseline ( $P > 0.8$  for all comparisons). A similar but weaker tendency was recognized when the data from the entire period of the sound presentation were subjected to the analysis (main effect of condition,  $P = 0.26$ ; FRS vs. baseline,  $P = 0.05$ ). This is reasonable because the time course of the grand average occipital alpha-EEG in this experiment showed, as in Experiment 1, a gradual increase over the first several tens of seconds of FRS (data not shown).



**B**



**Fig. 3.**

Normalized alpha-EEG potentials in each experimental condition (FRS, HCS, LCS, and baseline) during the latter half of the sound presentation in EEG Experiment 2. *A*: BEAMs averaged across all 17 subjects over the time period from the 100- to 200-s class marks. *B*: mean and standard error of the occipital alpha-EEG for all 17 subjects. FRS significantly enhanced the occipital alpha-EEG relative to the other conditions.

[Download figure](#) | [Download PowerPoint](#)

## PET experiment

When the conditions with audible sounds (i.e., FRS or HCS) were compared with those without audible sounds (i.e., LCS or baseline), the bilateral temporal cortex, presumably the primary and secondary auditory cortex, always showed significantly increased rCBF as expected (Table 1; see also Fig. 5C). More importantly, when FRS was compared with HCS,

brain stem (Fig.4B) and the lateral part of the left thalamus (Fig.4C). The same areas also showed an increased rCBF when FRS was compared with either the baseline or LCS (Fig.5, A and B). This tendency was also recognizable in the comparison of FRS versus baseline with a lower threshold ( $Z > 1.64$  with correction for multiple comparisons) (Fig. 5C and Table 1). Conversely, when HCS was presented, these areas in fact showed a decreased rCBF compared with the baseline (Fig. 5, A and B). When LCS was compared with the baseline, no significant differential activation was observed anywhere in the brain and neither the left thalamus nor the brain stem showed changes in rCBF.



Table 1. Location and significance of activation in each area explored by SPM software

[Enlarge table](#)

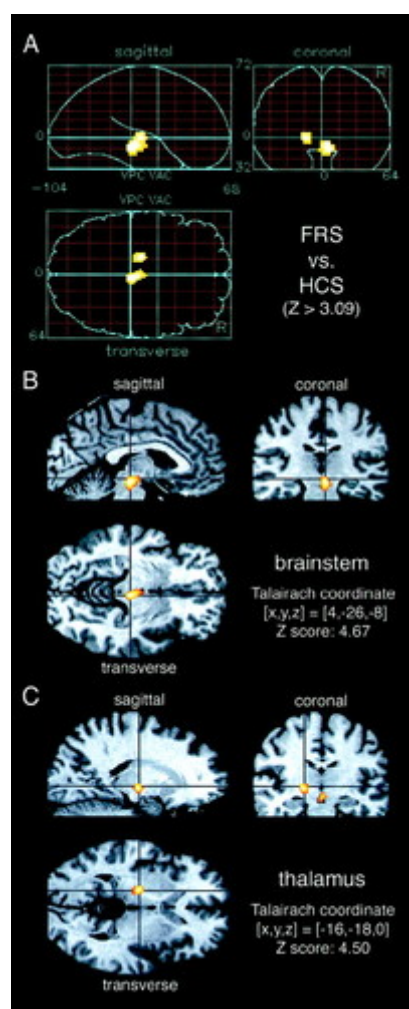
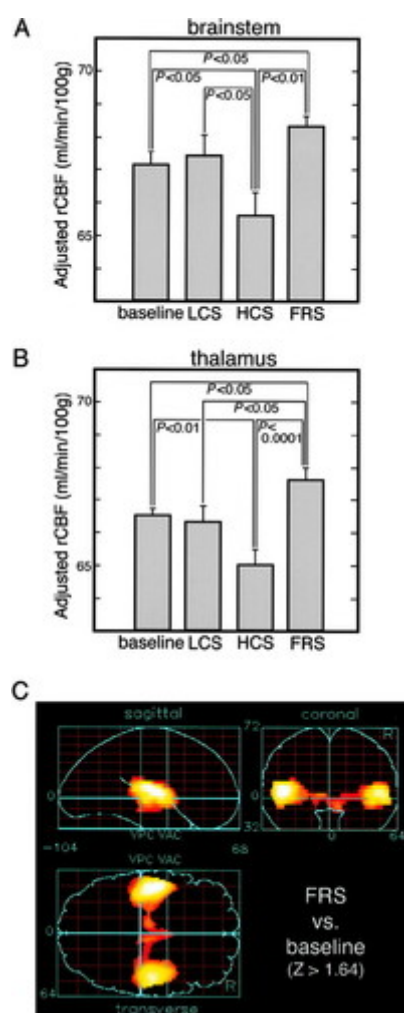


Fig. 4.



Z scores for the regions where activity was significantly increased during FRS as compared with HCS are shown in a standard anatomical space (Talairach and Tournoux 1988) viewed from the back (coronal view), the right side (sagittal view), and the top (transverse view) of the brain. Maps are illustrated by a color scale, with the lower Z score represented in darker red and the higher Z score in brighter yellow. *B* and *C*: activated foci in the brain stem (*B*) and left thalamus (*C*) are superimposed onto the spatially normalized magnetic resonance image (MRI) as shown in neurological convention. Maps are illustrated by a color scale, with the lower Zscore represented in darker red and the higher Z score in brighter yellow.

[Download figure](#) | [Download PowerPoint](#)



**Fig. 5.**

Averaged regional cerebral blood flow (rCBF) at the activated foci in a comparison of FRS and HCS across the 12 subjects under each condition (see Table 1). *A*: brain stem; *B*: left thalamus. Each *P* value indicates the significance calculated by Fisher's protected least significant difference (PLSD) following



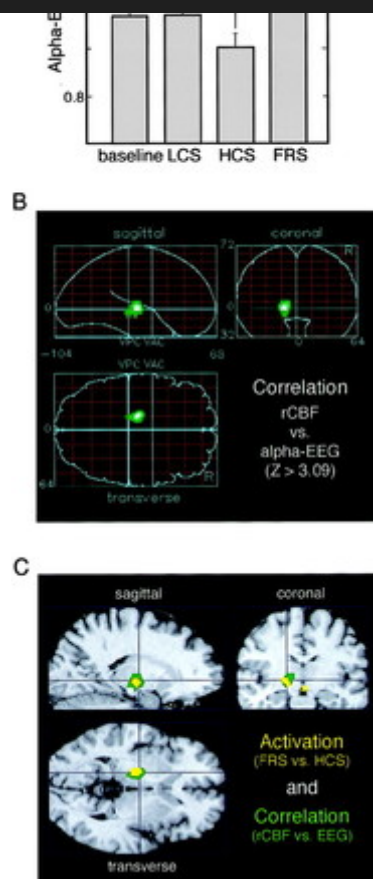


other conditions (baseline, LCS, and HCS). rCBF decreased during HCS compared with the baseline. On the other hand, presentation of LCS did not lead to any change in rCBF compared with the baseline values. *C*: projection images of SPM contrasting FRS with the baseline (in the same format as in Fig. 4A) with a lower threshold ( $Z > 1.64$  and  $P < 0.05$  with correction for multiple comparisons).

[Download figure](#) | [Download PowerPoint](#)

## EEG–rCBF correlation

The EEGs measured simultaneously with PET showed that FRS significantly increased alpha-EEG activity compared with HCS ( $P < 0.05$ ) (Fig. 6A), which is in complete agreement with the findings of the EEG experiments performed independently of the PET experiment. In contrast, when HCS was compared with the baseline, alpha-EEG activity decreased slightly in parallel with the changes we observed in the rCBF. The normalized EEG potentials showed a significant correlation with the rCBF equivalent value in the lateral part of the thalamus ( $r = 0.539$ ,  $P < 0.0001$ ). The maximum correlation in the brain was observed at  $x = -16$ ,  $y = -16$ , and  $z = 0$  ( $Z$  score = 4.30) in the stereotaxic space, which corresponds to the pixel immediately adjacent to the maximally significant point in the left thalamus as determined by the rCBF experiments (Fig. 6, *B* and *C*, and Table 1).



**Fig. 6.**

Normalized alpha-EEG potentials under each experimental condition and their correlation with adjusted rCBF equivalent values as measured by positron emission tomography (PET) scanning. *A*: grand average normalized alpha-EEG potentials with standard error for 11 subjects. The data of one subject were excluded because of excessive electrical noise in the EEG. ANOVA followed by Fisher's PLSD post hoc test showed a significant main effect of condition ( $P < 0.05$ ) and a significant increase during the presentation of FRS compared with HCS ( $P < 0.05$ ). *B*: the brain areas in which the rCBF equivalent values were significantly correlated with the alpha-EEG potentials are shown in a standard format (the same as in Fig. 4*A*). Maps are illustrated in a color scale, with the higher  $Z$  score represented in brighter green. *C*: the same area as shown in *B* (green) and the activated area in the comparisons of FRS and HCS as shown in Fig. 4 (yellow) are superimposed. The same area in the left thalamus that was activated by FRS was most significantly correlated with the alpha-EEG potentials. The maximum correlation in the brain was observed at  $[x, y, z] = [-16, -16, 0]$  ( $Z$  score = 4.30) in the stereotaxic space, which is the pixel immediately adjacent to the maximally significant local point observed in the rCBF experiment (FRS vs. HCS).

Table 2 shows the subjective evaluation of sound quality examined by Scheffé's paired comparison method (Scheffé 1952). A significant difference was evident between FRS and HCS in some elements of sound quality. Subjects felt that FRS was softer, more reverberant, with a better balance of instruments, more comfortable to the ears, and richer in nuance than HCS.



Table 2. Subjective evaluation of sound quality under FRS and HCS conditions

[Enlarge table](#)

## DISCUSSIONPhysiological effects of inaudible high-frequency sounds

Despite the fact that nonstationary HFCs were not perceived as sounds by themselves, we demonstrated that the presentation of sounds that contained a considerable amount of nonstationary HFCs (i.e., FRS) significantly enhanced the power of the spontaneous EEG activity of alpha range when compared with the same sound lacking HFCs (i.e., HCS). In parallel experiments employing exactly the same stimulus and methods, PET rCBF measurement revealed that FRS activated the deep-lying brain structures, including the brain stem and thalamus, compared with HCS. In addition, subjective evaluation by questionnaire revealed that FRS intensified the subjects' pleasure to a significantly greater extent than HCS did. We conclude, therefore, that inaudible high-frequency sounds with a nonstationary structure may cause non-negligible effects on the human brain when coexisting with audible low-frequency sounds. We term this phenomenon the “hypersonic effect” and the sounds introducing this effect the “hypersonic sound.” We do not think that the hypersonic effect is specific to the sound material used in the present study because we previously confirmed, by EEG analysis, that the same effect can be introduced by different sound sources containing a significant amount of nonstationary HFCs (e.g., Oohashi et al. 1994).

In contrast to the fact that the primary auditory cortex in the bilateral temporal lobes was similarly activated by FRS and HCS, it is noteworthy that the brain stem and thalamic foci activated by the presentation of FRS showed a decrease in rCBF when HCS was presented,



and LFCs, not HFCs alone, that specifically induces the enhancement of alpha-EEG and activation in the deep-lying structures. We interpret these findings to mean that the hypersonic effect does not simply result from a neurophysiological response to isolated frequencies above an audible range, but from a more complex interaction to which HFCs and LFCs both contribute.

The alpha rhythm of EEG is considered to occur in relaxed yet alert subjects and to be sensitive to the subjects' emotional as well as arousal state ([Drennen and O'reilly 1986](#); [Iwaki et al. 1997](#)). Although there is considerable inter-subject variability in the amount of alpha rhythm, normal alpha rhythm can be treated as an intra-individually stable trait in terms of its test–retest reliability ([Fernandez et al. 1993](#); [Gasser et al. 1985](#); [Kohrman et al. 1989](#)). Although the mechanisms underlying generation of the alpha rhythm have yet to be fully clarified, an animal model suggests the involvement of at least the thalamocortical and intracortical networks ([Steriade et al. 1990](#)). Our finding of a significant positive correlation between the rCBF in the thalamus and the occipital alpha-EEG suggests that the occipital alpha-EEG may reflect an aspect of activity in deep-lying structures, including the thalamus. This finding does not contradict our earlier report ([Sadato et al. 1998](#)), which did not address the physiological effect of inaudible high-frequency sounds.

### Explanation of the discrepancy between the present and previous studies

The fact that we used an entire piece of natural music lasting 200 s as sound stimuli instead of short fragments of sounds might explain the discrepancy between our findings and those of previous studies carried out around 1980 to determine the format for digital audio CDs (e.g., [Muraoka et al. 1978](#); [Plenge et al. 1979](#)), which concluded that the presence of sounds containing a frequency range above 15 kHz was not recognized as making a difference in sound quality. The [CCIR \(1978\)](#), and the current International Telecommunication Union–Radio communication sector ([ITU-R 1997](#)), have recommended that sound samples used for the comparison of sound quality should not last longer than 15–20 s ([CCIR 1978](#); [ITU-R 1997](#)), and that intervals between sound samples should be about 0.5–1 s ([CCIR 1978](#)) because of short-term human memory limitations. Most of the previous psychological experiments, including the studies by [Muraoka et al. \(1978\)](#) and [Plenge et al. \(1979\)](#), were carried out using, essentially, the sound presentation method recommended by the CCIR. We also examined the psychological evaluation using the same material and sound presentation system as was used for the present study, but followed the presentation method



In our EEG and PET experiments, we focused on physiological brain responses and objectively evaluated the effect of the combination of audible sounds and inaudible HFCs on brain activity, independent of a subjective evaluation of sound quality. According to the EEG measurements, the occipital alpha-EEG gradually increased over several tens of seconds after the exposure to FRS began, and this increase persisted for several tens of seconds after FRS ended. These findings suggest that the phenomenon that we call the hypersonic effect may involve some neuronal mechanisms that can be characterized by delay and persistence for as long as several tens of seconds. It seems, therefore, that an exposure to FRS shorter than 20 s, as recommended by the CCIR and ITU-R, may be insufficient to introduce a physiological effect. By the same token, a short exposure to HCS following FRS with a short interval of 0.5–1 s may not be enough to withdraw physiological effects, if any, induced by the preceding FRS. Based on this physiological consideration, we performed our psychological experiment with sound materials of longer duration. The results showed a significant difference between FRS and HCS in some elements of sound quality. That difference was evident despite the fact that a long presentation time should make it more difficult to detect subtle differences between two materials due to the limitation of short-term auditory memory. Our findings suggest the possibility that the results of the previous psychological studies may not be valid in a situation where humans are continuously exposed to auditory stimuli such as music or environmental sounds.

### **Hypothetical explanation of neuronal mechanisms of the hypersonic effect**

From an authentic view of human auditory physiology, it is not straightforward to explain the neuronal basis of the hypersonic effect characterized by the fact that HFCs showed significant physiological and psychological effects on listeners only when presented with audible sounds. Although how inaudible HFCs produce a physiological effect on brain activity is still unknown, we need to consider at least two possible explanations. The first is that HFCs might change the response characteristics of the tympanic membrane in the ears and produce more realistic acoustic perception, which might increase pleasantness. However, this hypothesis is unlikely to explain the fact that the subjects who showed significant hypersonic effect were not necessarily aware of the difference of sounds in a conscious manner. An alternative explanation is that HFCs might be conveyed through pathways distinct from the usual air-conducting auditory pathway and therefore might affect the CNS, including the deep-lying brain structure. It was reported that the vibratory



as those whose hearing is totally impaired ([Lenhardt et al. 1991](#)). Recently evidence has accumulated that stimuli outside the frequency and amplitude boundaries of an auditory neuron's receptive field can influence responses to stimuli inside the classical receptive field determined with pure tone stimuli (e.g., [Schulze and Langner 1999](#)). This modulatory interaction between inside and outside the classical auditory receptive range is noteworthy. However, we cannot conclude that the neural mechanisms incorporating ultrasound hearing, including the bone-conducting auditory pathway, are the system responsible for the hypersonic effect, which involves the brain stem and thalamus. These regions showed decreased activity compared with the baseline when HCS was presented and thus may not belong to the conventional auditory perception system. Therefore participation of nonauditory sensory systems such as somatosensory perception also needs to be considered in further investigations.

We pay special attention to the fact that FRS is accompanied by an intensification of the pleasure with which the sound is perceived, and envisage the participation of the neuronal pathways in connection with reward-generating systems ([Cooper 1991](#); [Olds and Milner 1954](#); [Wise 1980](#)), which effectively control various aspects of human behavior. The present PET result does not seem to be contradictory to this view. The brain stem contains distinct neuronal groups that are the major source of monoaminergic projections to various parts of the brain ([Nieuwenhuys et al. 1988](#); [Role and Kelly 1991](#)). These monoaminergic systems are thought to be the primary sites for the action of many stimulants and antipsychotic drugs ([Kandel 1991](#)). The rCBF in this area was reported to increase after oral amphetamine challenge ([Devous et al. 1995](#)). These fibers lie in the medial forebrain bundle, which is considered to be intimately connected with registering pleasurable sensations ([Thompson 1988](#)). The monoaminergic neurons or the opioid-peptidergic neurons in the deep-lying brain structures are characterized by long neurotransmitter residence times at synaptic junctions and the participation of an intracellular messenger in the postsynaptic neurons ([Hartzell 1981](#); [Kehoe and Marty 1980](#); [Schwartz and Kandel 1991](#)). These characteristics seem to support the delay and persistence of the hypersonic effect observed in the present EEG experiments. The activation of the thalamus may reflect its function as part of the limbic system, which also plays an important role in the control of emotions ([LeDoux 1993](#); [Vogt and Gabriel 1993](#)). It might also reflect the role of the thalamus in gating sensory input to the cortex ([Andreasen et al. 1994](#)). We speculate that changes of activity in the deep-lying structure may introduce some modulatory effects on the perception of audible sounds and



as a message carrier and frequencies above the audible range, together with those in the audible range, function as a modulator of sound perception through the brain systems, including the reward-generating system. Further investigations are clearly required to examine this hypothetical model.

In conclusion, our findings that showed an increase in alpha-EEG potentials, activation of deep-seated brain structures, a correlation between alpha-EEG and rCBF in the thalamus, and a subjective preference toward FRS, give strong evidence supporting the existence of a previously unrecognized response to high-frequency sound beyond the audible range that might be distinct from more usual auditory phenomena. Additional support for this hypothesis could come from future noninvasive measurements of the biochemical markers in the brain such as monoamines or opioid peptides.

We thank the staff of the Kyoto University PET Center for valuable contributions to this work; Dr. Yoshio Yamasaki, Waseda University, for the use of his recently developed signal processing system; the Yamashiro Institute of Science and Culture for recording the sound sources; Dr. Norihiro Sadato, National Institute for Physiological Sciences, for valuable comments on an early version of the manuscript; and Dr. Masako Morimoto, Japan Society for the Promotion of Science, for valuable technical support.

This work was supported in part by the Japan Ministry of Education, Science and Culture, through the Grants-in-Aid for Scientific Research (A) (09490031) to T. Oohashi, on Priority Areas to H. Shibasaki, and for International Scientific Research Program (10041144) to T. Oohashi, and by the Japan Society for the Promotion of Science through the Research for the Future Program JSPS-RFTF 97L00201 to H. Shibasaki.

## FOOTNOTES

- The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “*advertisement*” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

## AUTHOR NOTES

- Address for reprint requests: T. Oohashi, Dept. of KANSEI Brain Science, ATR Human Information Processing Laboratories, 2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-



JOURNALS ▾

[Previous](#)


[Back to Top](#)

[Next](#)

*Sign up for alerts*

**SIGN UP**



 **the-aps.org** [Privacy policy](#) | [Disclaimer](#)

Copyright © 2018 The American Physiological Society