Mechanism for generating peaks and notches of head-related transfer functions in the median plane

Hironori Takemoto,[a] Parham Mokhtari, Hiroaki Kato, and Ryouichi Nishimura
National Institute of Information and Communications Technology, c/o ATR 2-2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

Kazuhiro Iida
Chiba Institute of Technology, 2-17-1, Tsudanuma, Narashino, Chiba 275-0016, Japan

It has been suggested that the first spectral peak and the first two spectral notches of head-related transfer functions (HRTFs) are cues for sound localization in the median plane. Therefore, to examine the mechanism for generating spectral peaks and notches, HRTFs were calculated from four head shapes using the finite-difference time-domain method. The comparison between HRTFs calculated from the whole head and the pinna-related transfer functions calculated from the segmented pinna indicated that the pinna determines the basic peak–notch pattern of the HRTFs. An analysis of the distribution patterns of pressure nodes and anti-nodes on the pinna computed in the steady state for sinusoidal excitations confirmed that the first three peaks correspond to the first three normal modes of the pinna. The analysis also revealed that at the first spectral notch frequencies, one or two anti-nodes appeared in the cymba and the triangular fossa, and a node developed in the concha. Furthermore, according to changes in the instantaneous pressure distribution patterns on the pinna, three types of mechanisms were hypothesized for inducing the node in the concha depending on the source elevation angle. © 2012 Acoustical Society of America.
[http://dx.doi.org/10.1121/1.4765083]

Pages: 3832–3841

I. INTRODUCTION

Humans can perceive the azimuth and elevation of a sound source in space. The interaural time difference (ITD) and interaural level difference (ILD) provide primary cues for perception of the azimuth (e.g., Blauert, 1997). ITD and ILD, however, are not sufficient for elevation perception: ITD and ILD are approximately equal for any places on a sagittal plane. Instead, spectral features of the head-related transfer function (HRTF) provide cues of the elevation. HRTFs consist of peaks and sharp notches generated mainly by the pinna. Shaw and Teranishi (1968) reported that the center frequency of the lowest sharp spectral notch systematically increased from 6 to 10 kHz as the source elevation angle increased from −45° to 45°. Butler and Belendiuk (1977) also reported a systematic change in sharp notches from −30° to 30°, and Bloom (1977) showed that manipulation of the center frequency of notches could evoke the percept of elevation angles from −45° to 45°. Several other studies implied that sharp spectral notches are perceptual cues for elevation (e.g., Gardner and Gardner, 1973; Hebrank and Wright, 1974; Wright *et al.*, 1974; Moore *et al.*, 1989; Hofman *et al.*, 1998). Presumably, the perception of the elevation angle using those sharp notches as cues would require relatively broad-band sources.

In the median plane, HRTFs have a typical peak–notch pattern. Figure 1 schematically shows the typical pattern for the lower three peaks (P1–P3) and the lower three notches (N1–N3). The horizontal and vertical axes indicate the frequency and the elevation angle, respectively. This schematic, based on Figs. 1 and 8 in Raykar *et al.* (2005), was derived from the CIPIC HRIR database (Algazi *et al.*, 2001b). Note that this figure is rotated 90° from its original orientation. As shown in Fig. 1, the notch frequencies increase as the elevation angle increases, while the peak frequencies are constant with elevation angle. From the results of psychophysical experiments, Iida *et al.* (2007) concluded that N1 and N2 are the spectral cues to sound source elevation and P1 could be the reference information used by the human hearing system to analyze the notches. The mechanisms by which these spectral peaks and notches are generated, however, are not entirely clear. The purpose of the present study is to clarify the mechanisms by simulating sound pressure patterns on the pinna from stimulation from different directions.

Figure 2 shows the pinna. The pinna has the following four small cavities, which are partly separated by the crus of helix (f in Fig. 2) and the antihelix (g in Fig. 2): scaphoid fossa (a in Fig. 2), triangular fossa (b in Fig. 2), cymba conchae (c in Fig. 2), and cavity of concha (d in Fig. 2). Sound waves coming from any direction are reflected and diffracted in complex ways in these cavities before reaching the ear canal (e in Fig. 2). Through this process, some frequency components of the incoming waves are amplified, while other components are attenuated. Thus, the pinna acts as an acoustic filter. Hereafter, we simply refer to these four cavities as the pinna cavities, the cymba conchae as the cymba, and the cavity of concha as the concha. In addition, we

---

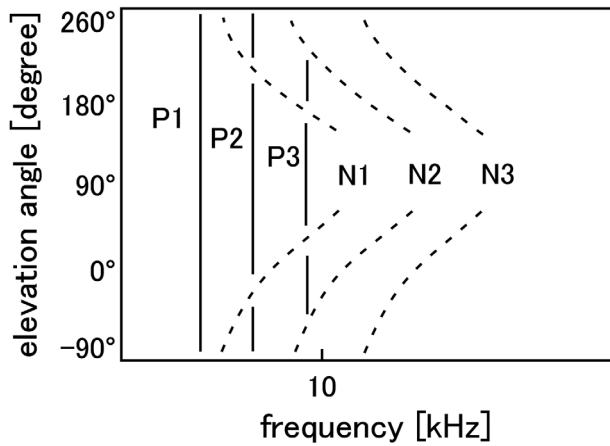[a] Author to whom correspondence should be addressed. Electronic mail: takemoto@nict.go.jp

FIG. 1. A schematized peak–notch pattern of HRTFs in the median plane, based on Figs. 1 and 8 in Raykar *et al.* (2005). Note that this figure is rotated 90° from its original presentation, and thus the horizontal axis indicates frequency and the vertical axis indicates elevation angle.

define three of these cavities—scaphoid fossa, triangular fossa, and cymba—as the upper cavities.

The mechanism for generating peaks is better understood than that of generating notches. Shaw (1997) measured pressure distribution patterns on the pinnae of ten subjects under blocked meatus conditions at the frequencies of the lower six peaks and plotted the average location and the amplitude together with the relative phase of the pressure nodes and the anti-nodes on the pinna. As a result, he showed that as the peak frequency increased, the number of pressure nodes and anti-nodes increased. This indicates that peaks are generated by normal modes of the pinna. Kahana and Nelson (2006) simulated pressure distribution patterns on various pinnae [e.g., two pinnae of KEMAR (Burkhard and Sachs, 1975)] by the boundary element method (BEM). They also
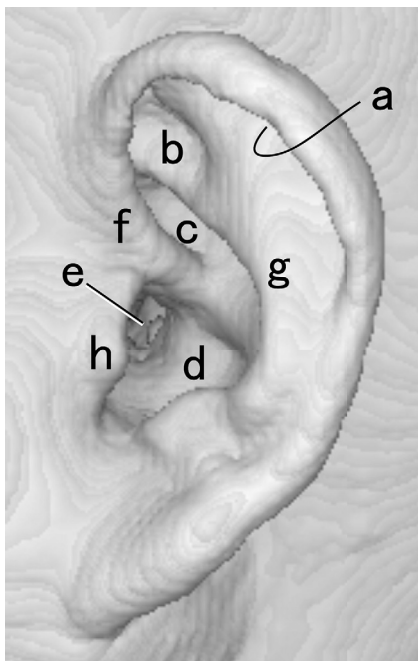

FIG. 2. Anatomical nomenclature of the pinna. a: Scaphoid fossa. b: Triangular fossa. c: Cymba conchae (cymba). d: Cavity of concha (concha). e: Ear canal. f: Crus of helix. g: Antihelix. h: Tragus.

plotted absolute pressure distribution patterns on the pinna at peak frequencies and confirmed that those peaks were caused by the normal modes of the pinna, whereas the number of normal modes varied according to the pinna geometry. In short, it is safe to say that spectral peaks are derived from the normal modes of the pinna.

In contrast, the mechanism for generating notches remains unclear. One hypothesis is that a direct sound wave is canceled by a sound wave reflected on the posterior concha wall (Batteau, 1967; Hebrank and Wright, 1974; Raykar *et al.*, 2005). Based on this hypothesis, Raykar *et al.* (2005) estimated reflection points on the pinna from notch frequencies for elevation angles from −45° to 90° for four subjects and plotted the points on the pinna image of the corresponding subject. As a result, the estimated reflection points appeared roughly on the antihelix and the crus of helix. This indicated that the pinna geometry has relevance to the changing patterns of notch frequencies. Raykar *et al.* (2005), however, stated that the mechanism for spectral notches for elevations behind the ear is not clear. In short, we still do not understand the mechanism by which spectral notches are generated for the complete range of elevation angles.

To understand how spectral peaks and notches are generated, three-dimensional acoustic simulation methods are beneficial because the pressure distribution patterns on the pinna can be calculated and visualized, as Kahana and Nelson (2006) reported. In the present study, therefore, we adopted a three-dimensional acoustic simulation method based on the finite-difference time-domain (FDTD) method (Mokhtari *et al.*, 2007; Takemoto *et al.*, 2010). We confirmed that the FDTD method can calculate HRTFs comparable with the conventional acoustic measurement method (Mokhtari *et al.*, 2010). The FDTD method calculates a frequency response much faster than does the BEM, although the FDTD method is less accurate if the spatial resolution is the same. In addition, because the FDTD method uses an orthogonal grid system, three-dimensionally reconstructed images obtained by magnetic resonance imaging (MRI) can be directly used as analysis data. Thus, in the present study, we measured four human heads by MRI for acoustic analyses. In our analyses, we first compared the HRTFs in the median plane calculated from the whole head and those from the pinna segmented from the head to show that the pinna alone can generate the basic peak–notch pattern of the HRTFs. Note that we omitted the effects of body parts other than the head (such as the torso) because of MRI scanning limitations, even though these parts also affect the HRTFs, especially in the median plane (e.g., Kreuzer *et al.*, 2009; Algazi *et al.*, 2001a). Then, we calculated the distribution patterns of pressure nodes and anti-nodes within the pinna cavities at peak and notch frequencies and displayed them for visualization in a way comparable to those used by Shaw (1997) and Kahana and Nelson (2006).

II. MATERIALS AND METHODS

A. MRI data

The heads of two adult Japanese males (M1 and M2) and two adult Japanese females (F1 and F2) were measured by

Takemoto *et al.*: Generation of peaks and notches by the pinna   3833
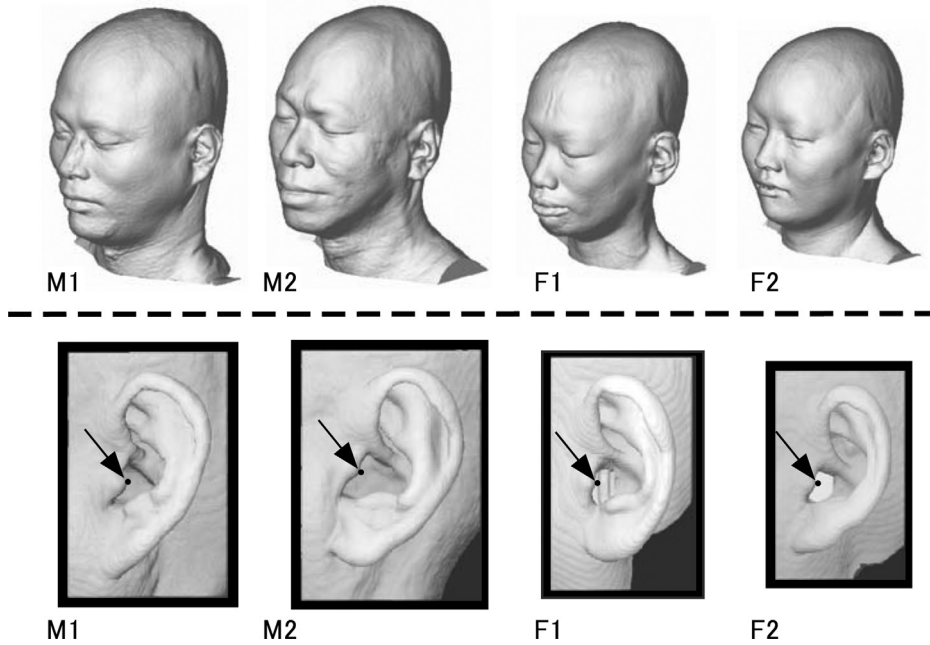
FIG. 3. Whole heads of four subjects (upper panel) and the left pinnae segmented from the heads (lower panel). The black circles pointed out with black arrows in the lower panel indicate the source point for calculation of PRTFs and HRTFs.

MRI [Shimadzu-Marconi MAGNEX ECLIPSE 1.5T Power Drive 250 (Shimadzu Corporation, Kyoto, Japan)] installed in the Brain Activity Imaging Center at ATR-Promotions Inc. The spatial resolution of the MRI data (i.e., the size of each voxel) was 1.2 mm for M1 and M2, 1.1 mm for F1, and 1.0 mm for F2, while 4.47 ms echo time and 12 ms repetition time were common among the four subjects. From these images, three-dimensional head data were reconstructed.

The surface of the head for each subject was segmented from the air by binarization and a region growing process that eliminated isolated points (i.e., noise) in the air after the bilateral ear canals were blocked. Then, the left pinnae were segmented from the whole heads. Figure 3 shows the whole heads (upper panel) and the segmented left pinnae (lower panel). Note that the ear canals are not observed in the lower panel, because they were blocked in the modeling process. The whole heads in the upper panel all have the same scale, and the segmented pinnae in the lower panel have a common scale.

B. Calculation of HRTFs and pinna-related transfer functions (PRTFs)

The FDTD method described in Takemoto *et al.* (2010) was used for calculating HRTFs and PRTFs for the left ear of each subject. The accuracy of the finite-difference method was second order in time and fourth order in space: The temporal derivative was approximated by a second-order centered-difference, and the spatial derivative was approximated by a fourth-order centered-difference. The analysis field was surrounded by a perfectly matched layer (Berenger, 1994) with a depth of 12 cells, which absorbed the outgoing sound waves. To simulate wall reflection, the surface impedance method proposed by Yokota *et al.* (2002) was implemented.

Figures 4(a) and 4(b) illustrate the setting for calculating the HRTFs and the PRTFs. In calculating the HRTFs, according to the acoustic reciprocity theorem (e.g., Pierce, 1989; Zotkin *et al.*, 2006), a source point was placed one

voxel (1.2 mm for M1 and M2, 1.1 mm for F1, and 1.0 mm for F2) away from the blocked meatus (the small white circles in the lower panel in Fig. 3), and observation points were placed at 10° intervals in the median plane on the
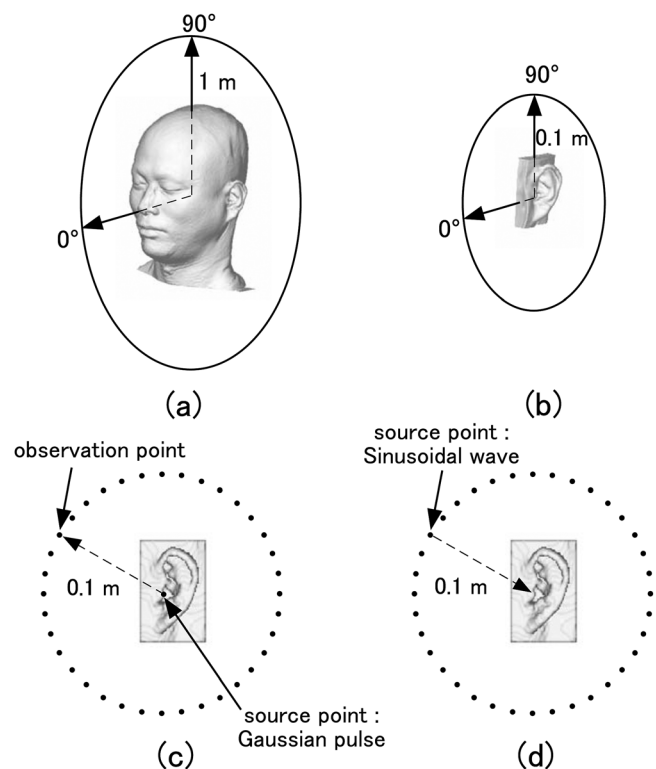

FIG. 4. (a) HRTFs were calculated at 1.0 m in the median plane. (b) PRTFs were calculated at 0.1 m in the sagittal plane passing through the source point denoted by the white circles in the lower panel of Fig. 3. (c) When PRTFs were calculated, the source point was placed just outside of the blocked meatus and a total of 36 observation points were placed on the circumference of a circle with 0.1 m radius. A Gaussian pulse was input to the source point and pressure changes for all the observation points were obtained in a single computational run. (d) In contrast, when the pressure distribution pattern on the pinna was calculated, the source point was placed at the observation point used in the calculation of PRTFs.

circumference of a circle with a 1.0 m radius. In calculating the PRTFs, a source point was also set one voxel away from the blocked meatus, but the observation points were placed on the circumference of a circle with 0.1 m radius at 10° intervals, as Fig. 4(d) shows. In addition, the center of the circle for calculating the PRTFs was equal to the source point; that is, the observation points were not on the median plane of the head but on the sagittal plane passing through the source point. The simulation frequency was 1.0 MHz, the pressure changes at all observation points were calculated during 4 ms, and the HRTFs and the PRTFs were calculated up to 24 kHz. Note that the elevation angle for the direction directly in front was set at 0° and that for the direction directly above was set at 90°.

In calculating both the HRTFs and the PRTFs, the volume velocity given by the following Gaussian pulse gp(t) was input to the source point:

$$\mathrm{gp}(t) = e^{-\{\Delta tn - T/0.29T\}^2} [\mathrm{m}^3/\mathrm{s}], \tag{1}$$

where $T = 0.646/f_0$ and $f_0 = 10$ kHz. Figure 5(a) shows the input volume velocity with 0.132 ms duration, Fig. 5(b) shows the pressure change in the open air at 0.1 m from the source point, and Fig. 5(c) indicates the power spectrum of Fig. 5(b). As shown in Fig. 5(c), this pulse has a nearly flat frequency characteristic across a wide frequency range. The frequency characteristic of the Gaussian pulse was used for normalizing the pressure changes at each observation point to obtain the HRTFs and the PRTFs.

C. Calculation of distribution patterns of pressure nodes and anti-nodes on the pinna

To understand the mechanism for generating spectral peaks and notches, the distribution patterns of pressure nodes and anti-nodes on the pinna at peaks and notches were calculated and displayed for visualization. From the numerous peaks and notches on the PRTFs, particular peaks and notches were selected and each frequency and elevation
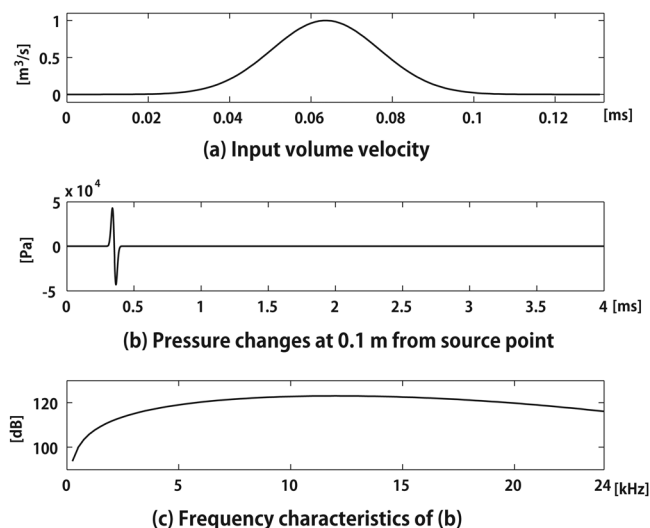
FIG. 5. (a) Input volume velocity (Gaussian pulse) given by Eq. (1). (b) Calculated pressure changes at 0.1 m from the source point in free space. (c) Frequency characteristics of pressure changes of (b).

angle was extracted. For the elevation angle, a source point was selected from among the 36 points placed on the circumference of a circle of 0.1 m radius, as shown in Fig. 4(d). Then, a sinusoidal wave with the extracted frequency was input to the selected source point as a volume velocity source to excite the pinna cavities. The amplitude of the sinusoidal wave was 1.0 m³/s. In this calculation, although the pressure distribution pattern at each simulation step in the whole analysis field can be output as a computational file for a desired duration, the total size of output files became enormous. Thus, after reaching a steady state, the instantaneous pressure distribution patterns were output as a series of files at five simulation step intervals (i.e., 200 kHz sampling rate) for 0.5 ms.

In the time course of the instantaneous pressure distribution pattern on the pinna, the change in pressure was locally minimal at a node, and locally maximal at an anti-node. Thus, time-averaged absolute pressure distribution patterns were calculated over a cycle. In the present study, voxels having a time-averaged absolute pressure value below 9.8×10^3 Pa were visualized as nodes, whereas voxels with a value above 1.8×10^4 Pa were visualized as anti-nodes. These values were determined in order to compare our results with those of Shaw (1997) and Kahana and Nelson (2006). Although the pressure nodes can be visualized in this process, multiple anti-nodes lost their relative phase information. Consequently, if two pressure anti-nodes had opposite phases, these two anti-nodes were visualized in the same color. To avoid this, phase information was extracted and retained from the instantaneous pressure distribution pattern. Using a volume rendering method, pressure nodes were colored green, anti-nodes with a negative phase were colored blue, and anti-nodes with a positive phase were colored red in the present study.

III. RESULTS AND DISCUSSION

A. Comparison between HRTFs and PRTFs

Figure 6 shows the HRTFs and the PRTFs for the left ear of the four subjects. As Fig. 6 shows, the peak–notch pattern was generally common among the HRTFs and the PRTFs for all four subjects, although the fine spectral structures were different. This indicates that the pinna generates the basic peak–notch pattern of the HRTFs, and thus acoustic analyses of the segmented pinna can reveal the mechanisms of generating spectral peaks and notches of the HRTFs in the median plane.

B. Peak–notch patterns on PRTFs

By setting the grayscale to show only the positive log-power spectral components (i.e., greater than or equal to 0 dB), Figs. 7(a), 7(c), 7(e), and 7(g) represent only the peak patterns of the PRTFs of the four subjects. In Figs. 7(a), 7(c), 7(e), and 7(g), most peaks are observed as vertical bands. This indicates that normal modes generating peaks could be induced by sources at a certain range of elevation angles. In other words, peak frequencies are not sensitive to changes in the source elevation angle. According to Kahana and Nelson
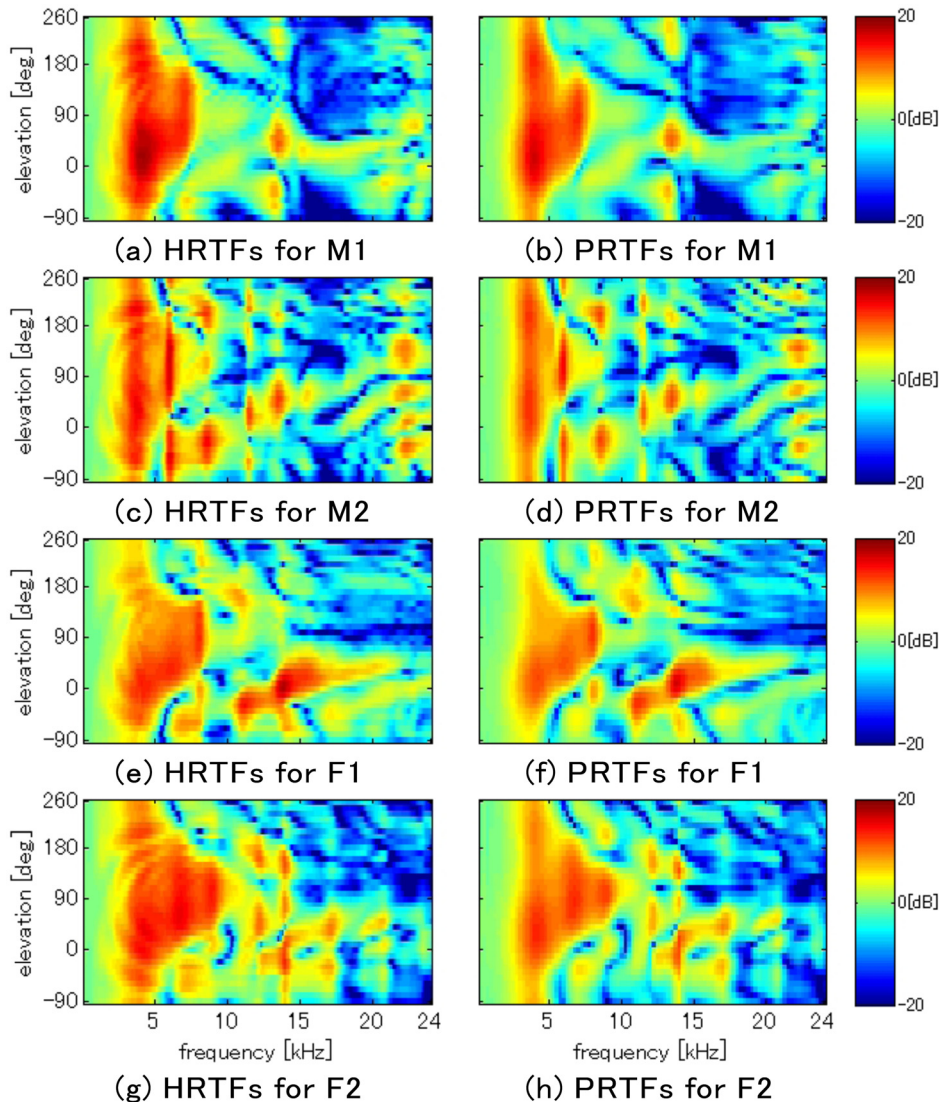
FIG. 6. (Color online) HRTFs and PRTFs of four subjects. The scale bars indicate the log-power in dB.

(2006), five or six normal modes are observed below 20 kHz. The present study confirmed their results: M1 had five peaks and the other subjects had six peaks. As shown in Fig. 3, the pinna of M1 is larger than that of F2, while the pinna of M1 has fewer peaks than that of F2. This fact indicates that the number of peaks below a certain frequency is not determined simply by the overall size of the pinna. Instead, the configuration of the pinna cavities could be a factor.

In contrast, only the negative components of the PRTFs are plotted in Figs. 7(b), 7(d), 7(f), and 7(h). Here, the notches appear as curved lines. This indicates that, unlike peaks, notches were generated at specific frequencies depending on the elevation angle. In other words, notches are sensitive to changes in the source elevation angle. Thus, it is reasonable to assume that notches can be considered as perceptual cues for elevation. In most cases, notch trajectories increased in frequency as the elevation angle approached approximately 90° (directly above), while they decreased as the elevation angle deviated from 90°.

N1 and N2, which are considered as perceptual cues for elevation (Iida *et al.*, 2007), were the focus in this study. N1 was observed between P1 and P2 at an elevation angle of −90° (directly below). As the elevation angle increased, N1

increased in frequency and interrupted P2 and P3 to reach a maximum frequency at approximately 90°–120°. Above P3 frequency, the N1 trajectories became ambiguous: N1 became shallow and broad for M1 and F1, other notches appeared for M2, and N1 disappeared for F2. The N1 trajectories were again visible from approximately 120° to 150° and then decreased in frequency and interrupted P3 and P2 again to approach P1 asymptotically near 260°. When the frequency of N1 approached and crossed the frequencies of P2 and P3, in some cases N1 appeared at the intersection while in other cases P2 and P3 appeared. In short, N1 and P2 or P3 interacted with and interrupted each other, and the log-power spectral amplitude at the intersection therefore varied from case to case depending on the relative strength of the peak–notch interactions.

The N2 trajectories appeared at much more limited elevation angles than did the N1 trajectories. They were observed at the following elevation angles: 150°–260° for M1, 0°–50° for F1, and −90°–10° for F2. That is, in most cases, the N2 trajectory was observed only as curve segments and not as a continuous curve. Thus, it was difficult to examine the mechanism by which N2 was generated and how its frequency shifted with the elevation angle. In the following sections,
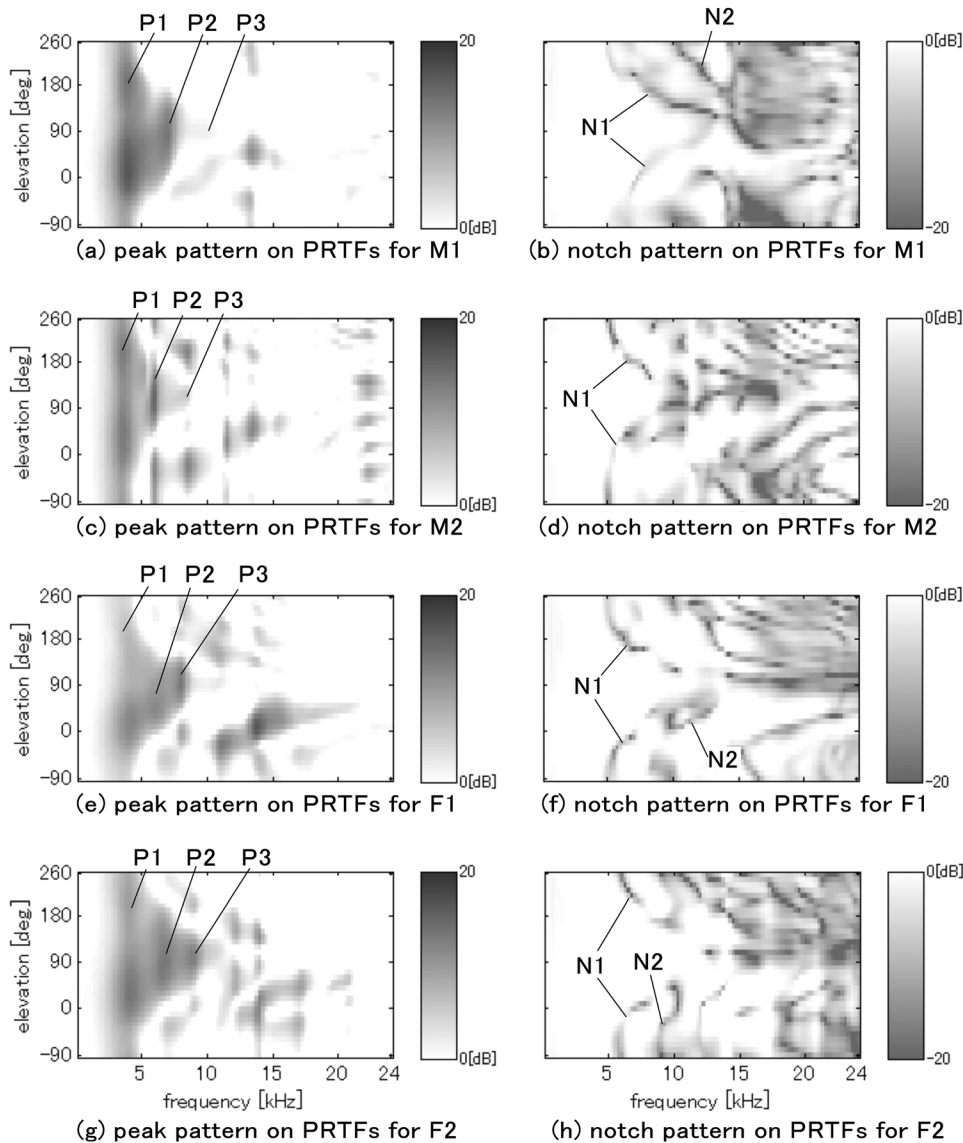
FIG. 7. Peak patterns on PRTFs (left) and notch patterns on PRTFs (right). The scale bars indicate the log-power in dB. Note that the dark regions in panels (a), (c), (e), and (g) indicate higher log-power spectral amplitudes of PRTFs, while the dark regions in panels (b), (d), (f), and (h) indicate lower log-power spectral amplitude of PRTFs.

therefore, we confirm the first, second, and third normal modes of the pinna (i.e., P1–P3), whose frequency range was almost the same as the N1 frequency range. Then, we examine the mechanism for generating N1 and discuss how the N1 frequency changes with the elevation angle.

C. Distribution patterns of pressure nodes and anti-nodes at P1–P3

Figure 8 shows the distribution patterns of pressure nodes and anti-nodes at P1–P3 for M2, that is, the first, second, and third normal modes of the pinna of M2. Note that these pressure nodes and anti-nodes were computed in the steady state for sinusoidal excitation. For P1–P3, the excitation source directions were 0°, 60°, and 120° and the excitation frequencies were 3.5, 6.0, and 8.25 kHz, respectively. As Shaw (1997) described, the first mode was "a simple quarter-wavelength depth resonance." Indeed at P1, a single pressure anti-node developed in the concha, the cymba, and the triangular fossa, as shown in Fig. 8(a). The average absolute pressure values were high on the base of the concha, and they decreased with increasing distance from the base. This reso-

nance was excited across all source elevation angles for all subjects, and thus it was confirmed to be omnidirectional.

The second and third modes were "vertical modes," as Shaw (1997) and Kahana and Nelson (2006) reported. In the second mode, a nodal surface developed along the crus helix, and two anti-nodes with opposite phases developed in the concha and the combined region from the cymba to the triangular fossa, as shown in Fig. 8(b). For M2, this mode was excited at all elevation angles other than those at which N1 intersected P2. For the other subjects, this mode was excited at a more limited range of elevation angles around approximately 90°. In the third mode, three pressure anti-nodes and two nodes appeared. The first pressure anti-node developed in the anterior part of the concha, the second one spread through the posterior part of the concha and the cymba, and the third one appeared in the triangular fossa, as shown in Fig. 8(c). The first and third anti-nodes had the same phase, while the second one had the opposite phase. Therefore, a pressure node was generated between two adjacent anti-nodes. The pressure node appeared vertically between the first and second anti-nodes, and another node emerged between the cymba and triangular fossa. Compared with the

(a) P1: mode 1
0°, 3.5 kHz

(b) P2: mode 2
90°, 6.0 kHz
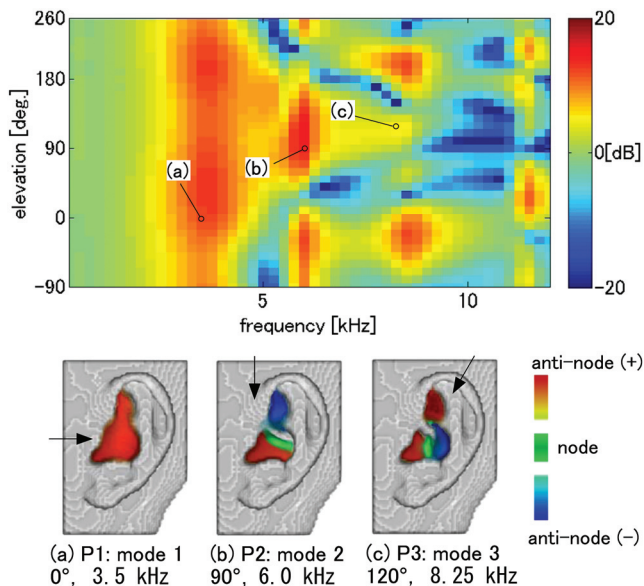
(c) P3: mode 3
120°, 8.25 kHz

FIG. 8. The upper panel shows PRTFs for M2 below 12 kHz to indicate the frequency and each elevation angle chosen for the lower panel. The scale bar indicates the log-power in dB. The lower panel indicates the distribution patterns of pressure nodes and anti-nodes at the first, second, and third normal modes of the pinna. Arrows represent the source direction.

second resonance mode, the third resonance mode was excited at more limited elevation angles. This resonance was relatively weak for M1.

The distribution patterns of the pressure nodes and the anti-nodes at the second and third modes shown by Shaw [(1997), Fig. 7] agreed well with those of M2 in the present study. However, the distribution pattern of the third mode shown by Kahana and Nelson [(2006), Fig. 7(c)] differed from that of M2: Kahana and Nelson (2006) showed that the three anti-nodes appeared in the concha, the cymba, and the scaphoid fossa. This difference might be explained by the different pinna geometry, because the distribution pattern of F1 in the present study was the same as that of Kahana and Nelson (2006).

D. Distribution patterns of pressure nodes and anti-nodes at N1

We examined the distribution patterns of the pressure nodes and the anti-nodes at the N1 frequency at various elevation angles for the four subjects and categorized the patterns into three types. Figure 9 shows the typical node and anti-node patterns of these three types for M2, and Fig. 10 shows where each type appeared on the N1 trajectory of each subject.

Figures 9(a) and 9(b) present two patterns of type 1 for M2 at elevation angles of −30° and 0°, where N1 appeared at 5.25 and 5.5 kHz, respectively. In type 1, a single anti-node developed in the triangular fossa and the cymba, and a node appeared in the concha. This type was the major pattern, observed when the sound source was placed approximately below the horizontal plane and approximately in front of the subject, as shown in Fig. 10 by the elevation angles for each subject: below 20° and above 190° for M1, below 20° and

above 210° for M2, below 0° and above 160° for F1, and below 20° and above 190° for F2. For M2, as the source elevation angle increased from −30° to 0°, the N1 frequency increased from 5.25 to 5.5 kHz. Along with these changes, the anti-node in the triangular fossa decreased in volume, while the anti-node in the cymba increased in volume. That is, as the source elevation changed, the location and volume of the single anti-node also changed to maintain a node in the concha and, thus, the N1 frequency change was smooth.

Figures 9(c) and 9(d) show two patterns of type 2 for M2 at elevation angles of 30° and 50°, where N1 appeared at 6.25 and 7.0 kHz, respectively. In type 2, two anti-nodes with opposite phases appeared in the triangular fossa and the cymba, and partly in the posterior part of the concha. Because the node in the concha involved another node generated between the two anti-nodes, a large node developed across the concha and the cymba. This type was observed when the sound source was placed in antero-superior directions at the

(a) type 1
−30°, 5.25 kHz

(b) type 1
0°, 5.5 kHz

(c) type 2
30°, 6.25 kHz

(d) type 2
50°, 7.0 kHz

(e) type 3
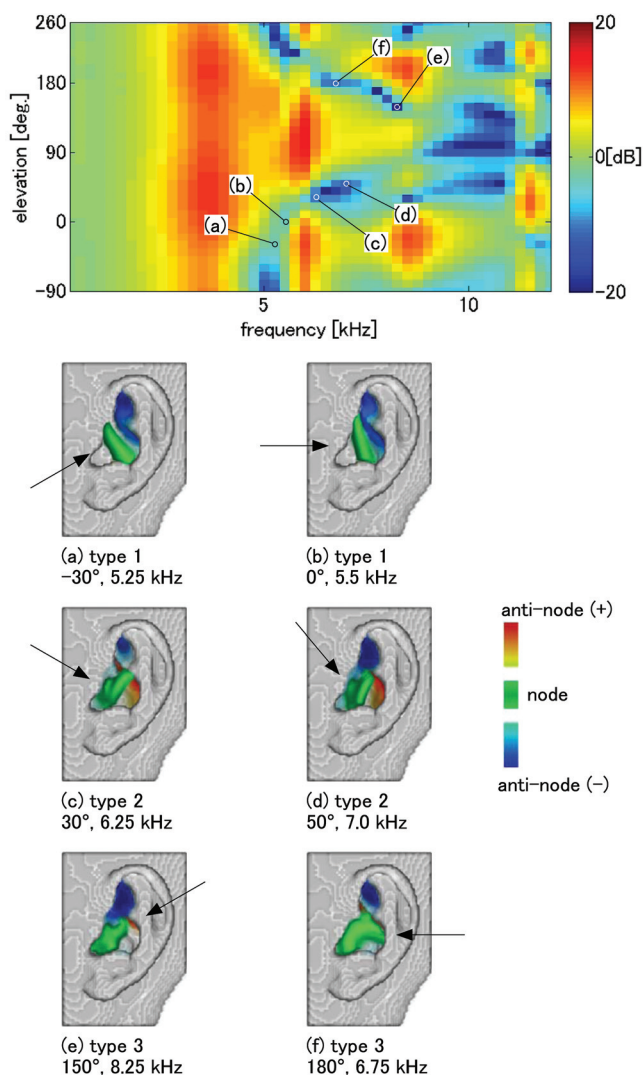150°, 8.25 kHz

(f) type 3
180°, 6.75 kHz

FIG. 9. The upper panel shows PRTFs for M2 below 12 kHz to indicate the frequency and each elevation angle chosen for the lower panels. The scale bar indicates the log-power in dB. The lower panels indicate the three types of distribution patterns of pressure nodes and anti-nodes on the pinna (see the text). Arrows represent the source direction.

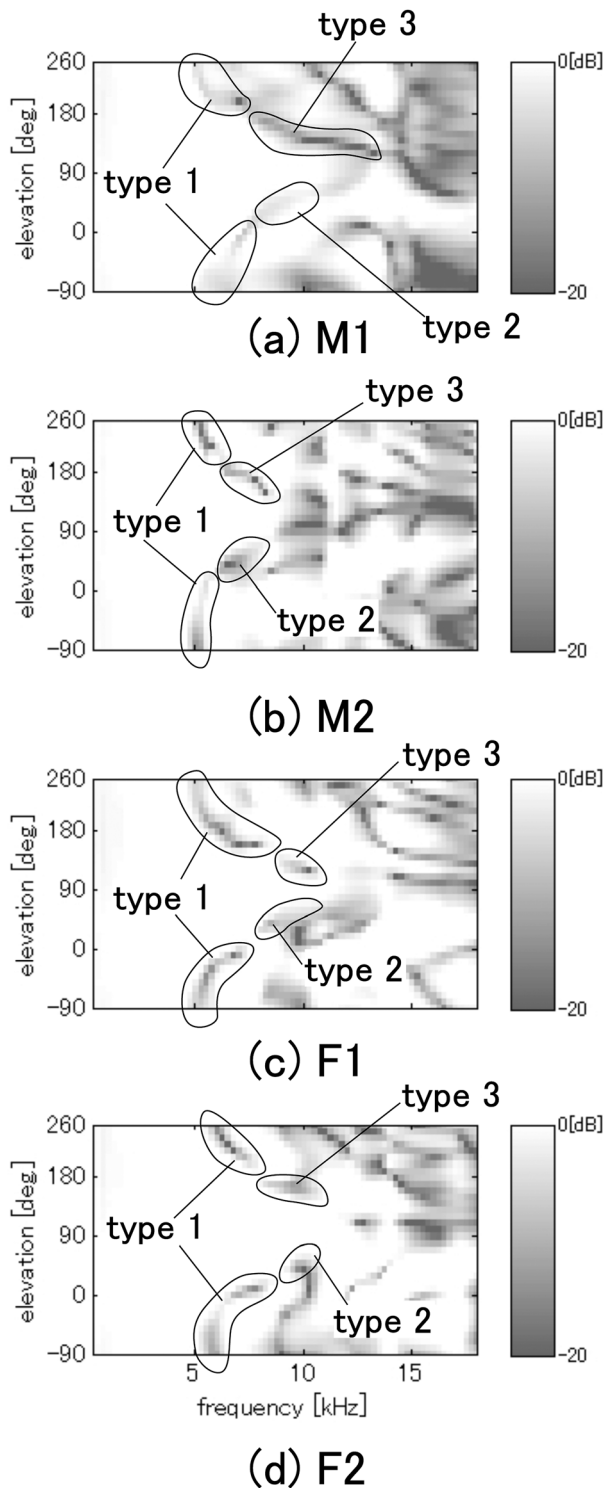Takemoto *et al.*: Generation of peaks and notches by the pinna

FIG. 10. Notch patterns of PRTFs of all four subjects, indicating the three types of distribution patterns of pressure nodes and anti-nodes on the pinna (see the text) for each N1 trajectory. The scale bars indicate the log-power in dB.

following elevation angles as shown in Fig. 10: 30°–60° for M1, 30°–70° for M2, 30°–60° for F1, and 30°–60° for F2. For M2, as the source elevation angle increased from 30° to 50°, the N1 frequency increased from 6.25 to 7.0 kHz. Along with these changes, the anti-node in the triangular fossa (colored blue) increased in volume; that is, the pressure changes in the triangular fossa became larger. Note that when the source ele-

vation further increased, the pressure changes in the triangular fossa became much larger. Although the results of the increase of pressure changes in the triangular fossa were not clear, one possible explanation is that the increase might be necessary to maintain a node in the concha as the source elevation increased.

Figures 9(e) and 9(f) show two patterns of type 3 for M2 at elevation angles of 150° and 180°, where N1 appeared at 8.25 and 6.75 kHz, respectively. In type 3, as in type 2, two anti-nodes with opposite phases appeared in the triangular fossa and the cymba, but not in the posterior part of the concha. Note that the amplitude of the anti-node within the cymba was large, even though the anti-node is not clearly visible in Figs. 9(e) and 9(f). Because the triangular fossa extends laterally to the cymba (compared with the direction of the incoming acoustic waves), these two anti-nodes partially overlapped in the medio-lateral direction. Between the two anti-nodes, a nodal surface was generated and combined with a node in the concha. Although the distribution patterns of nodes and anti-nodes were very similar between types 2 and 3, the mechanism for inducing the node was different. In Sec. III E, we discuss the difference. As Fig. 10 shows, the type 3 pattern was found when the sound source was placed in the postero-superior direction at the following elevation angles for each subject: 120°–180° for M1, 150°–200° for M2, 120°–150° for F1, and 150°–180° for F2. For M2, as the source elevation angle increased from 150° to 180°, the N1 frequency decreased from 8.25 to 6.75 kHz. Along with these changes, in contrast to type 2, the anti-node in the triangular fossa decreased in volume, and the pressure changes in the triangular fossa became smaller.

In short, type 1 was observed for inferior directions of the sound source, type 2 was observed for antero-superior directions, and type 3 was observed for postero-superior directions. Based on the number of anti-nodes, type 1 can be considered as the first mode of N1, whereas types 2 and 3 can be considered as the second mode. Because the second mode was higher in frequency than the first mode, the major factor of the N1 frequency change with the source elevation angle was the mode change: The first mode occurred for sound waves coming from below the subject, and the second mode occurred for sound waves coming from above the subject. A minor (or secondary) factor was a change in the arrangement pattern of the anti-nodes within the same mode. In all three types, when the source elevation slightly changed, the volume and location of the anti-nodes slightly changed to maintain a node in the concha, and consequently, the N1 frequency slightly changed with elevation angle.

E. Hypothesis of mechanism for inducing a node in the concha

The pressure nodes and anti-nodes mentioned previously were those of standing waves in the pinna cavities induced by propagating sinusoidal waves. Normally, a node is placed midway between two adjacent anti-nodes with opposite phase, e.g., the nodes in Figs. 8(b) and 8(c). However, as shown in Fig. 9, the node in the concha was not always placed between the two anti-nodes of standing waves. One

(a) type 1
M2, −30°, 5.25 kHz

(b) type 2
M2, 40°, 6.5 kHz
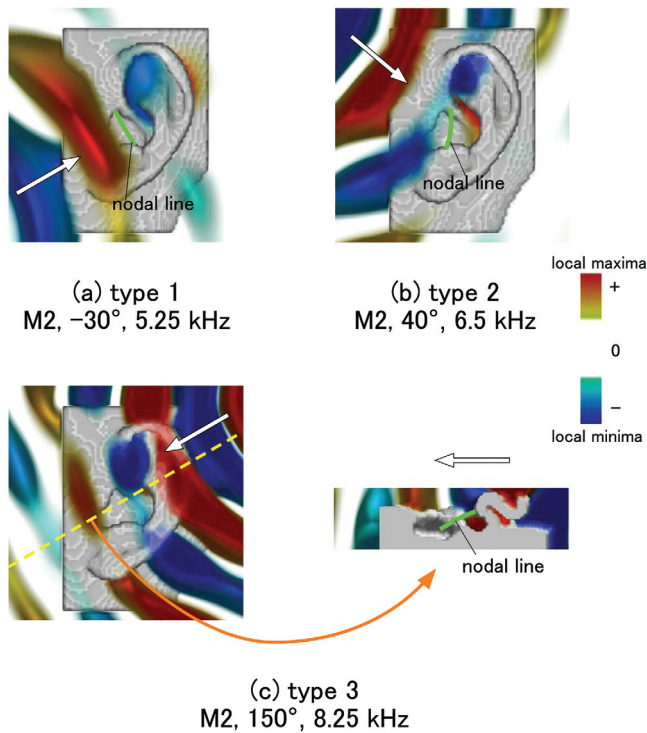
(c) type 3
M2, 150°, 8.25 kHz

FIG. 11. The mechanisms by which the propagating waves induce a node in the concha for the three types of distribution patterns of pressure nodes and anti-nodes on the pinna of M2, as shown in Fig. 9.

possible explanation is that this node is induced by interactions between the local maxima or minima of the propagating waves and the anti-node with opposite phase in the upper cavities. In this section, therefore, we describe a hypothesis on how the node in the concha was induced.

Figures 11(a)–11(c) show instantaneous pressure distribution patterns around the pinna of M2 at elevation angles of −30°, 40°, and 150°, where N1 appeared at 5.25, 6.5, and 8.25 kHz, respectively. In Figs. 11(a)–11(c), voxels with pressure values less than -1.8×10^4 Pa are colored blue, voxels with pressure values more than 1.8×10^4 Pa are colored red, and voxels with pressure values close to 0 Pa are transparent. Thus, in addition to the anti-nodes in the pinna cavities, the local pressure minima and maxima of propagating waves are also colored blue and red, respectively. Note that the nodes are omitted in Figs. 11(a)–11(c).

Figure 11(a) shows the instantaneous pressure distribution patterns when a positive wave front of propagating waves coming from the antero-inferior direction (−30°, denoted by the white arrow) reached the antero-inferior edge of the pinna. At this instant, a negative anti-node developed in the triangular fossa and the cymba. It is unlikely that the previous negative wave front caused this negative anti-node because most parts of the previous negative wave front were canceled by the positive anti-node previously developed in the triangular fossa and the cymba. Thus, a nodal line was generated in the concha, midway between the positive wave front and the negative anti-node. Because the anti-node always had a phase opposite to the incoming wave front at all times throughout each cycle, pressure changes in the concha were very small and a nodal line always appeared there.

In this way, a node developed in the concha in the type 1 pattern, as shown in Figs. 9(a) and 9(b).

Figure 11(b) shows the instantaneous pressure distribution patterns when a negative wave front of propagating waves coming from the antero-superior direction (40°, denoted by the white arrow) reached the antero-superior edge of the pinna. At this instant, a positive anti-node developed in the cymba, while a negative anti-node developed in the triangular fossa. Thus, a nodal line was generated in the concha, midway between the negative wave front and the positive anti-node. Because the anti-node induced in the cymba always maintained a phase opposite to the incoming wave front at all times throughout each cycle, the nodal line always appeared there. In this way, a node developed in the concha in the type 2 pattern, as shown in Figs. 9(c) and 9(d).

The left-hand side of Fig. 11(c) shows the instantaneous pressure distribution patterns when a negative wave front of propagating waves coming from the postero-superior direction (150°, denoted by the white arrow) passed above the cymba. The right-hand side of Fig. 11(c) represents a cross section of the left-hand side along the yellow dashed line. Both the right- and left-hand sides of Fig. 11(c) show a positive anti-node within the cymba, while the left-hand side shows that a negative anti-node simultaneously developed in the triangular fossa. In addition to these two anti-nodes with opposite phase, the local pressure maxima or minima of the propagating waves seemed to contribute to generating the node in the concha, because the anti-node in the triangular fossa overlapped only the antero-superior part of the anti-node in the cymba. Between these two anti-nodes and simultaneously between the negative wave front and the positive anti-node in the cymba, a nodal surface was generated [its cross section is denoted by the green "nodal line" on the right-hand side of Fig. 11(c)]. In this case, the nodal surface was large enough to cover the concha. Because the anti-node in the cymba always had an opposite phase relative to the anti-node in the triangular fossa and the incoming wave front at all times throughout each cycle, the nodal surface was generated and maintained at every cycle. In this way, a node developed in the concha in the type 3 pattern, as shown in Figs. 9(e) and 9(f). As just described, type 3 and type 2 were distinct in their mechanisms for inducing a node in the concha, even though their distribution patterns of nodes and anti-nodes were similar.

IV. CONCLUSION

The present study examined the distribution patterns of pressure nodes and anti-nodes of the pinna of four subjects. We confirmed normal modes of the pinna at transfer function peaks and clarified the mechanism for generating N1 and shifting the N1 frequency. First, the heads of two males and two females were measured by MRI (Fig. 3) and the HRTFs in the median plane were calculated from the heads by the FDTD method. The PRTFs were then calculated from the left pinnae segmented from the heads (Fig. 3) and compared with the HRTFs (Fig. 6). The results showed that the basic pattern of spectral peaks and notches was common between the HRTFs and the PRTFs for all subjects, but the fine

structures were different. This result indicated that the basic peak–notch pattern of the HRTFs originated from the pinna.

We confirmed that P1–P3, respectively, were the first, second, and third normal modes of the pinna, as Shaw (1997) and Kahana and Nelson (2006) reported. We obtained essentially the same distribution patterns of pressure nodes and anti-nodes at each mode as Shaw (1997) measured and Kahana and Nelson (2006) computed. The first mode was excited across all elevation angles, and thus it was omnidirectional. The second and third modes were vertical modes. At the second and third modes, the distribution patterns of nodes and anti-nodes within a subject were stable at all elevation angles where these modes were excited. Among subjects, however, the frequency and the location of nodes and anti-nodes at each mode were different.

We also examined the distribution patterns of pressure nodes and anti-nodes at various N1 frequencies at various elevation angles for the four subjects and categorized the patterns into three types as shown in Figs. 9 and 10. Type 1 appeared when the sound source was placed approximately below the horizontal plane. In type 1, one pressure anti-node appeared in the upper cavities and one node appeared in the concha. The location and the volume of the anti-node changed with the elevation angle, and thus the N1 frequency also changed. Type 2 appeared when the sound source was placed in anterosuperior directions. In contrast, type 3 appeared when the sound source was placed in postero-superior directions. In both types 2 and 3, two anti-nodes developed in the triangular fossa and the cymba, and one combined node developed in the concha. When the source elevation angle slightly changed, the amplitude of the anti-node in the triangular fossa slightly changed, and consequently the N1 frequency changed, even if the mechanism was unknown.

Based on the number of anti-nodes, type 1 could be considered as the first mode of N1, and types 2 and 3 could be considered as the second mode. Because the second mode was higher in frequency than the first, a modal change from the first to the second mode was the major factor determining the N1 frequency change according to the elevation angle: The N1 frequency increased as the height of the sound source increased. A minor factor was the more gradual changes in the location and the volume of the anti-nodes. These two factors could account for systematic changes of the N1 frequency, except for the directions above the subject (at approximately an elevation angle of $90°$). Although the N1 trajectories were not consistently observed at directions approximately above the subjects in the present study, it is possible that some other subjects might have clear N1 trajectories for those directions. In that case, even higher modes that have three or more anti-nodes may possibly be observed.

In the present study, we revealed the mechanisms for generating P1–P3 and N1. However, we could not clarify the relationship between the individual pinna geometry and the frequencies and amplitudes of spectral peaks and notches, because the pinna shape is so complex that it is difficult to extract the physiological parameters. To examine this relationship, it would be necessary to develop a simplified pinna model that can generate the typical peak–notch pattern.

Algazi, V. R., Avendano, C., and Duda, R. O. (**2001a**). "Elevation localization and head-related transfer function analysis at low frequencies," J. Acoust. Soc. Am. **109**, 1110–1122.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001b**). "The CIPIC HRTF Database," in *Proceedings of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (Mohonk Mountain House, New Paltz, NY), pp. 99–102.

Batteau, D. W. (**1967**). "The role of the pinna in human localization," Proc. R. Soc. London, Ser. B **168**, 158–180.

Berenger, J. P. (**1994**). "A perfectly matched layer for the absorption of electromagnetic waves," J. Comput. Phys. **114**, 185–200.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization Revised Edition* (MIT, Cambridge, MA), Chap. 2.4, pp. 137–177.

Bloom, P. J. (**1977**). "Creating source elevation illusions by spectral manipulation," J. Audio Eng. Soc. **25**, 560–565.

Burkhard, M. D., and Sachs, R. M. (**1975**). "An Anthropometric manikin for acoustic research," J. Acoust. Soc. Am. **58**, 214–222.

Butler, R. A., and Belendiuk, K. (**1977**). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. **61**, 1264–1269.

Gardner, M. B., and Gardner, R. S. (**1973**). "Problem of localization in the median plane: Effect of pinnae cavity occlusion," J. Acoust. Soc. Am. **53**, 400–408.

Hebrank, J., and Wright, J. (**1974**). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. **56**, 1829–1834.

Hofman, P., Van Riswick, J., and Van Opstal, A. (**1998**). "Relearning sound localization with new ears," Nat. Neurosci. **1**, 417–421.

Iida, K., Itoh, M., Itagaki, A., and Morimoto, M. (**2007**). "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," Appl. Acoust. **68**, 835–850.

Kahana, Y., and Nelson, P. A. (**2006**). "Numerical modeling of the spatial acoustic response of the human pinna," J. Sound Vib. **292**, 148–178.

Kreuzer, W., Majdak, P., and Chen, Z. (**2009**). "Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range," J. Acoust. Soc. Am. **126**, 1280–1290.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2007**). "Comparison of simulated and measured HRTFs: FDTD simulation using MRI head data," *123rd Audio Engineering Society (AES) Convention*, New York, Preprint No. 7240, pp. 1–12.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2010**). "Computer simulation of KEMAR's head–related transfer functions: Verification with measurements and acoustic effects of modifying head shape and pinna concavity," in *Principles and Applications of Spatial Hearing* (World Scientific, Singapore), pp. 179–194.

Moore, B. C. J., Oldfield, S. R., and Dooley, G. (**1989**). "Detection and discrimination of spectral peaks and notches at 1 and 8 kHz," J. Acoust. Soc. Am. **85**, 820–836.

Pierce, A. D. (**1989**). *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, Melville, NY), pp. 180–183.

Raykar, V. C., Duraiswami, R., and Yegnanarayana, B. (**2005**). "Extracting the frequencies of the pinna spectral notches in measured head related impulse responses," J. Acoust. Soc. Am. **118**, 364–374.

Shaw, E. A. G. (**1997**). "Acoustical features of the human ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Erbaum, Mahwah, NJ), pp. 25–47.

Shaw, E. A. G., and Teranishi, R. (**1968**). "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," J. Acoust. Soc. Am. **44**, 240–249.

Takemoto, H., Mokhtari, P., and Kitamura, T. (**2010**). "Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method," J. Acoust. Soc. Am. **128**, 3724–3738.

Wright, D., Hebrank, J. H., and Wilson, B. (**1974**). "Pinna reflections as cues for localization," J. Acoust. Soc. Am. **56**, 957–962.

Yokota, T., Sakamoto, S., and Tachibana, H. (**2002**). "Visualization of sound propagation and scattering in rooms," Acoust. Sci. Tech. **23**, 40–46.

Zotkin, D. N., Duraiswami, R., Grassi, E., and Gumerov, N. A. (**2006**). "Fast head-related transfer function measurement via reciprocity," J. Acoust. Soc. Am. **120**, 2202–2215.