

PAPER

Upper median plane localization when head-related transfer functions of different target vertical angles are presented to the left and right ears

Kazuhiro Iida^{1,*}, Motokuni Itoh^{2,†} and Masayuki Morimoto^{2,‡}

¹*Department of Advanced Media, Faculty of Advanced Engineering, Chiba Institute of Technology,*

2-17-1 Tsudanuma, Narashino, Chiba, 275-0016 Japan

²*Environmental Acoustics Laboratory, Faculty of Engineering, Kobe University, Rokko, Nada, Kobe, 657-8501 Japan*

(Received 26 June 2017, Accepted for publication 5 January 2018)

Abstract: Previous studies showed that the distortion of the spectral information by occluding the pinna cavities of one of the ears decreases the accuracy of vertical localization. However, it is unclear whether vertical localization can be accomplished using the spectral information at only a single ear or requires the spectral information at both ears. In the present study, median plane localization tests in which the head-related transfer functions (HRTFs) for different target vertical angles were presented to each ear of the subjects were carried out. The results showed that the subjects either localized a single sound image to the target vertical angle presented to either the left or right ear or localized two sound images to both target vertical angles, with approximately the same accuracy as when identical target vertical angles of HRTFs were presented to both ears. These results imply that the spectral information at only a single ear provides the spectral cue for vertical localization.

Keywords: Head-related transfer function, Median plane localization, Spectral cue, Human auditory system, Single ear

PACS number: 43.66.Pn, 43.66.Qp [doi:10.1250/ast.39.275]

1. INTRODUCTION

It is generally known that the spectral notches and peaks in the human head-related transfer functions (HRTFs) in the frequency range approximately above 4 kHz act as cues for median plane localization [1–4].

The frequencies of the prominent high-frequency notches systematically increase with elevation [2,5] and are related to the physical dimensions and shape of the pinna [6–8]. Furthermore, the importance of the outline of spectral notches and peaks, rather than the fine structures, in the HRTF has been reported [9–14].

The authors proposed a parametric HRTF model for vertical sound localization [15]. The parametric HRTF is recomposed of all or some of the spectral notches and peaks extracted from a listener's measured HRTF, regarding the peak around 4 kHz, which is independent of the vertical angle of the sound source [5], as the lower-frequency limit. The notches and peaks are labeled in order

of frequency (e.g., P1, N1, P2, N2, and so on), and are expressed parametrically in terms of center frequency, level, and sharpness. The results of sound localization tests using the parametric HRTFs suggest that N1 and N2 play an important role in the localization in the upper median plane.

On the other hand, a number of studies have examined whether the spectral information of the input signals at both ears is necessary for the perception of vertical angle. Gardner [16] investigated the effects of pinna cavities on localization ability in the anterior sector of the median plane (-18° to 18°) by occluding the pinna cavities of one, both, or neither of the listener's ears. The sound was audible, but the pinna-induced cues were disrupted. The results showed that the localization accuracy was worse when one of the ears was occluded than when neither of the ears was occluded. The implication of his result was that the influence of the cavities was monaural in nature, at least insofar as localization in this sector of the median plane is concerned.

Morimoto [17] carried out localization tests in the upper hemisphere by occluding pinna cavities in the same manner as Gardner. The results showed that there were

*e-mail: kazuhiro.iida@it-chiba.ac.jp

[†]Presently, Panasonic Hearing Instruments Co., Ltd.

[‡]Presently, Emeritus Professor of Kobe University

significant differences in mean localization error among the cases in which one, both, or neither of the ears was occluded. He inferred that both ears contribute to the perception of vertical angle even in the median plane.

Some studies have reported that the spectral difference between the two ear-input signals might be used as a cue for median plane localization [18,19]. However, other studies have reported that the binaural disparity cues alone are not sufficient for median plane localization [20,21].

Thus, previous studies clarified that the distortion of the spectral information at one of the ears decreases the accuracy of vertical localization. However, as mentioned above, Gardner inferred that the cue for vertical localization is the spectral information at a single ear, while Morimoto inferred that the spectral information at both ears is necessary. Which hypothesis holds is not yet known and cannot be determined by experiments in which one of the listener's ears is occluded because the spectrum at the occluded ear is flattened.

Consider the direction at which a listener localizes a sound image under the following condition: (1) a sound source exists in the median plane and (2) the left ear of the listener is occluded while the right ear is not occluded. If the spectral information at a single ear contributes as the spectral cue, the listener localizes the sound image at the sound source with the spectral information of the right ear. However, the flattened spectrum at the left ear may decrease the accuracy of localization. Therefore, the following results are possible: (1) the listener localizes a sound image at the sound source, and so the accuracy of localization does not decrease, or (2) the accuracy of localization decreases.

On the other hand, if the spectral information at both ears contributes as the spectral cue, the accuracy of localization is thought to decrease because there exists no vertical angle at which the spectra at both ears coincide.

As mentioned above, the accuracy of localization decreased in both the experiments by Gardner and those by Morimoto. The results of these experiments cannot determine whether vertical localization can be accomplished using the spectral information at only a single ear or requires the spectral information at both ears. Therefore, another experiment, which provides different results depending on the process of extracting spectral cues in the human auditory system, is required.

In the present study, we carried out upper median plane localization tests presenting HRTFs of different target vertical angles to the left and right ears. If the location-related spectral cue is extracted in each ear and the location representations for the two ears are individually preserved until the final judgment, the possible results are as follows: (1) the listener localizes a sound image at either of two HRTFs, or (2) the listener localizes two sound images at

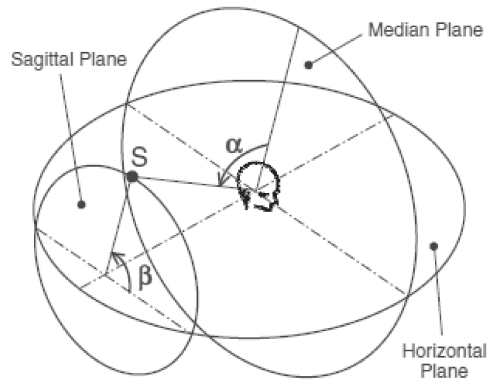


Fig. 1 Definition of the interaural-polar-axis coordinate system. Here, α is the lateral angle, and β is the vertical angle of sound image S.

both of the HRTFs. On the other hand, if the spectral information at both ears contributes as the spectral cue, a sound image is thought to be localized at neither of the HRTFs because the spectral information for both ears does not coincide with any vertical angle.

The interaural-polar-axis coordinate system [22], shown in Fig. 1, was used in the present study. This coordinate system expresses directions in three-dimensional space using the lateral angle, α , and the vertical angle, β , instead of the azimuth and elevation.

2. MEASUREMENT OF HRTFS

The HRTFs of four male subjects (IT, NS, SG, and YG), who were 23 to 30 years of age, were measured for seven vertical angles in the upper median plane (0° to 180° in 30° steps) in an anechoic chamber. The test signal was a swept sine wave, the sampling frequency of which was 48 kHz. The test signal was presented in 30° steps by a loudspeaker having a diameter of 80 mm (FOSTEX FE83E) located in the upper median plane. The distance from the loudspeakers to the center of the subject's head was 1.2 m. Earplug-type microphones [8] were used to sense the test signals at the entrances of the ear canals of the subject.

The earplug-type microphones were fabricated using the subject's ear molds. The ear molds were constructed by the following procedure: (1) an inverse mold was formed by occluding the pinna with silicon, (2) the inverse mold was encased in plaster, and (3) the silicon mold was removed. A miniature electret condenser microphone element of 5 mm in diameter (Panasonic, WM64AT102) was embedded in the silicon resin at the entrance of the ear canal of each ear mold. The microphone and silicon resin were then removed from the ear mold in order to be used as an earplug-type microphone.



Fig. 2 Placement of the earplug-type microphone in the entrance of the ear canal of a subject in the HRTF measurement.

The earplug-type microphones were placed into the ear canals of the subjects (Fig. 2). The diaphragms of the microphones were located at the entrances of the ear canals. This condition is referred to as the blocked-entrances condition [5]. The HRTF was obtained as

$$HRTF_{l,r}(\omega) = G_{l,r}(\omega)/F(\omega) \quad (1)$$

where $F(\omega)$ is the Fourier transform of the impulse response, $f(t)$, measured at the point corresponding to the center of the subject's head in the anechoic chamber without a subject, and $G_{l,r}(\omega)$ is that measured at the entrance of the ear canal of the subject with the earplug-type microphones. Subscripts l and r denote the left and right ears, respectively.

3. LOCALIZATION TESTS

3.1. Method

The subjects were four males (IT, NS, SG, and YG), whose HRTFs had been measured in the manner described in Sect. 2. All of the subjects self-reported normal hearing sensitivity.

The localization tests were conducted in an anechoic chamber. A PC, a digital signal processor (DSP) board for real-time processing, open-air headphones (AKG, K1000), and earplug-type microphones were used for the localization tests.

The sound pressure at the eardrum for the open-ear-canal condition, P , can be obtained by processing the sound pressure at the entrance of the blocked ear canal with a compensation filter, G , through headphones [23], as follows:

$$P = S \times HRTF \times G, \quad (2)$$

$$G = \frac{1}{M \times PTF} \times \frac{Z_{\text{ear canal}} + Z_{\text{headphone}}}{Z_{\text{ear canal}} + Z_{\text{radiation}}} \quad (3)$$



Fig. 3 Headphones used in the experiments.

$$\triangleq \frac{1}{M \times PTF} \times PDR, \quad (4)$$

where S denotes the sound source, M is the transfer function of the earplug-type microphones, $M \times PTF$ is the electroacoustic transfer function of the headphones measured at the entrance of the blocked ear canal, $Z_{\text{ear canal}}$ and $Z_{\text{headphone}}$ denote the impedances of the ear canal and headphones, respectively, and $Z_{\text{radiation}}$ is the free-air radiation impedance as observed from the ear canal. The second term on the right-hand side of Eq. (3) is referred to as the pressure division ratio (PDR). Møller *et al.* defined free-air equivalent coupling to the ear (FEC) headphones as headphones for which the PDR reduces to unity [23]. K1000 headphones (AKG), which were regarded as FEC headphones, were used in the localization tests.

The following procedure was used to measure $M \times PTF$. The subjects sat at the center of the soundproof room. The earplug-type microphones were placed into the ear canals of the subjects. The diaphragms of the microphones were located at the entrances of the ear canals, in the same manner as in the HRTF measurements described in Sect. 2. The subjects wore the headphones, and maximum-length sequence signals (48 kHz sampling) were emitted through the headphones. The signals were received by the earplug-type microphones, and $M \times PTF$ was obtained. The earplug-type microphones were then removed. Since the pinnae of the subject were not enclosed by the headphones (Fig. 3), displacement of the headphones did not occur.

The HRTFs for different vertical angles of 30° steps between 0° and 180° were presented to each ear of the subjects. Thus, a total of 49 (seven vertical angles \times seven vertical angles) combinations of HRTFs were prepared. The source signal was a wide-band white noise from 280 Hz to 11.2 kHz. The white noise was generated in the PC for each stimulus. Namely, each stimulus used a different white noise. Each stimulus was presented 10 times in random order at 60 ± 0.4 dB A-weighted SPL at the entrance of the ear canals of the subjects. The duration

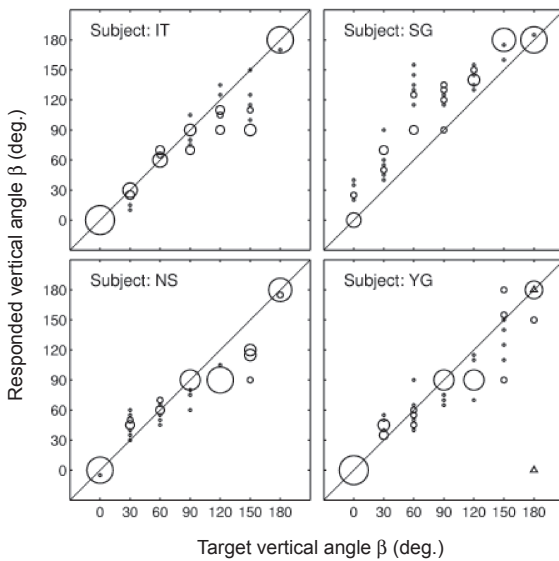


Fig. 4 Responded vertical angle, β , to stimuli having identical target vertical angles ($\beta_l = \beta_r$).

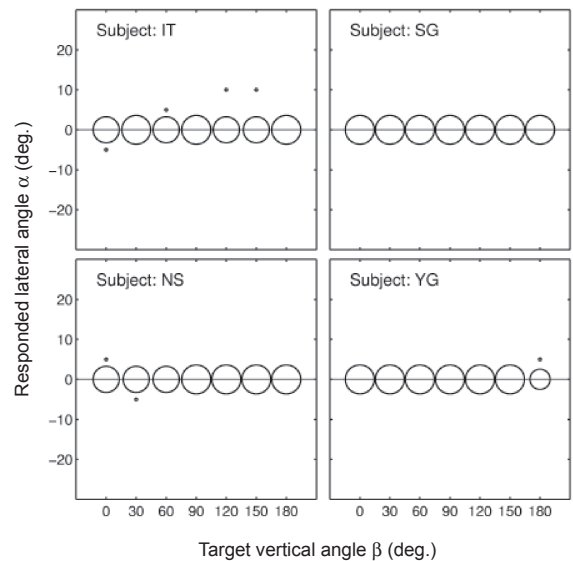


Fig. 5 Responded lateral angle, α , to stimuli having identical target vertical angles ($\beta_l = \beta_r$).

of the stimuli was 1 s, and the interval between two stimuli was 9 s. The task of the subjects was to mark the perceived azimuth and elevation of the center of each sound image on a graphic response form. The response form displayed two circles intersected by perpendicular lines printed upon a sheet of paper. One circle was used to indicate the perceived azimuth angle, and the other was used to indicate the perceived elevation angle, in reference to a spherical coordinate system containing a single vertical pole. The angles marked by subjects were read with a protractor to an accuracy of one degree and were transformed into the lateral angle, α , and the vertical angle, β , after the experiment.

3.2. Results

3.2.1. Case in which identical target vertical angles of HRTFs are presented to both ears

Figure 4 shows the responses for the vertical angle when the HRTFs for identical vertical angles were presented to both ears, as usual median plane localization tests. The ordinate represents the responded vertical angle, and the abscissa represents the target vertical angle. The diameter of each circle is proportional to the number of responses with a resolution of five degrees. In general, the responses are distributed along a diagonal line, indicating that the subjects perceived the vertical angle of a sound image accurately. However, subject IT localized a sound image at 90° to 120° for the target vertical angles of 120° and 150°. Subject NS localized a sound image at 90° for the target vertical angles of 120°, and at 90° to 120° for the

target vertical angles of 150°. Some of subject SG's responses tend to be localized around 120° for the target elevation of 60°. Subject YG localized a sound image at 90° to 120° for the target vertical angles of 120°, and at 90° to 180° for the target vertical angles of 150°. The subject occasionally perceived two sound images at 180° and 0° for the target vertical angle of 180°, indicated by the triangle symbol in Fig. 4.

Figure 5 shows the responses for the lateral angle, α . The ordinate represents the responded lateral angle, and the abscissa represents the target vertical angle. Most of the responses are distributed around 0°.

3.2.2. Case in which different target vertical angles of HRTFs are presented to each ear

Although the subjects mostly reported a single sound image when different target angles were presented to the left and right ears, subjects IT, SF, and YG sometimes reported two sound images. They reported either a single sound image or two sound images outside their heads. The proportion of trials in which the subject reported two sound images was 2.3% (IT), 0% (NS), 6.4% (SG), and 9.8% (YG).

Figures 6 through 9 show the responded vertical angle, β , when a single sound image was perceived. The ordinate represents the responded vertical angle, and the abscissa represents the target vertical angle presented to the right ear, β_r . The broken line indicates the target vertical angle presented to the left ear, β_l . Therefore, a response distribution along the broken line indicates that the subject localized a sound image to the vertical angle presented to

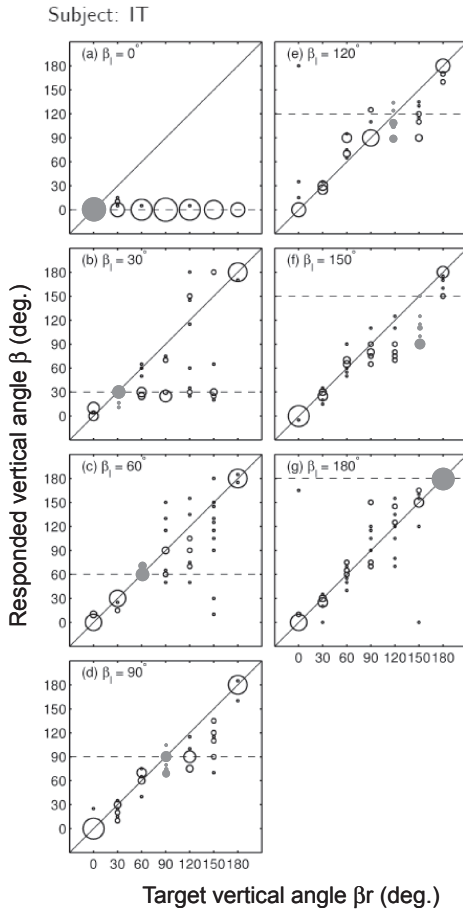


Fig. 6 Responses, which localize one sound image, to stimuli having different target vertical angles to the left and right ears for subject IT. The abscissa represents the target vertical angle presented to the right ear, β_r . The broken line denotes the target vertical angle presented to the left ear, β_l . The ordinate represents the responded vertical angle, β . For comparison, the responded vertical angle when the HRTFs of identical vertical angles were presented to both ears, shown in Fig. 4, is shown by closed gray circles.

the left ear. On the other hand, a distribution along the diagonal line indicates that the subject localized a sound image to the vertical angle presented to the right ear. For comparison, the responded vertical angle when the HRTFs of identical vertical angles were presented to both ears, shown in Fig. 4, is shown by closed gray circles.

Figure 6 shows the responses of subject IT. For $\beta_l = 0^\circ$ (panel (a)), the subject localized a sound image to the vertical angle presented to the left ear. For $\beta_l = 30^\circ$ and 60° (panels (b) and (c), respectively), the subject localized a sound image to the vertical angle presented to the right ear, except for $\beta_r = 60^\circ$ to 150° , for which the subject

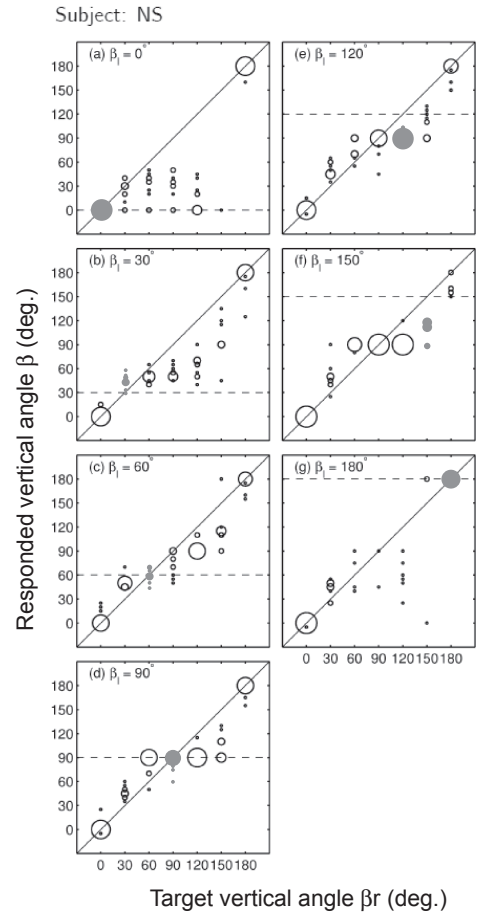


Fig. 7 Same as Fig. 6, but for subject NS.

sometimes localized a sound image to the target vertical angle presented to the left ear, and sometimes to that presented to the right ear. For $\beta_l = 90^\circ$ and 120° (panels (d) and (e), respectively), the subject localized a sound image to the vertical angle presented to the right ear. However, the subject localized a sound image at 90° to 120° for $\beta_r = 120^\circ$ and 150° . This tendency agrees with the responses obtained when identical target vertical angles were presented to both ears (closed gray circles for $\beta_l = \beta_r = 120^\circ$ and 150°). For $\beta_l = 150^\circ$ and 180° (panels (f) and (g), respectively), most of the responses were localized to the vertical angle presented to the right ear.

Figure 7 shows the responses of subject NS. For $\beta_l = 0^\circ$ (panel (a)), the subject localized a sound image to the vertical angle presented to the right ear for $\beta_r = 180^\circ$ and to an angle intermediate to those presented to the left and right ears for $\beta_r = 30^\circ$ to 120° . For $\beta_l = 30^\circ$ (panel (b)), the subject localized most of the sound images to the vertical angle presented to the right ear, with the exception of some responses for $\beta_r = 90^\circ$ to 150° . For $\beta_l = 60^\circ$ to

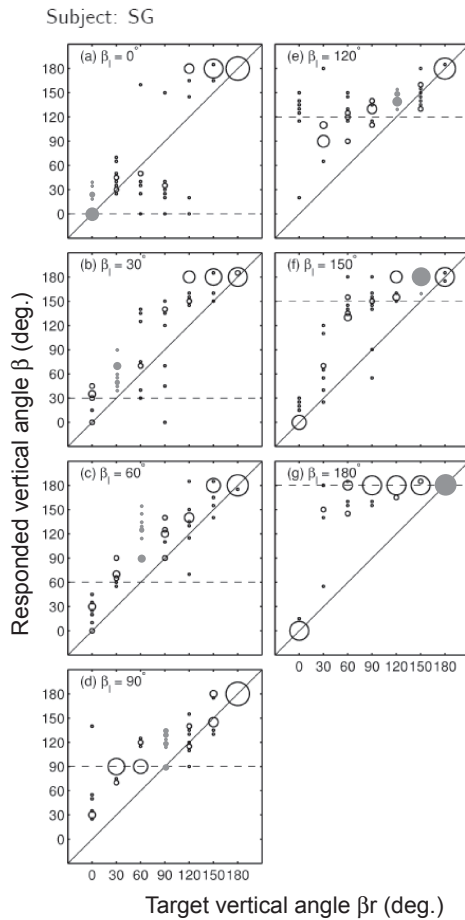


Fig. 8 Same as Fig. 6, but for subject SG.

150° (panels (c) through (f), respectively), the subject localized most of the sound images to the vertical angle presented to the right ear. However, the subject localized a sound image to 90° to 120° for $\beta_r = 120^\circ$ and 150° . This tendency agrees with the responses obtained when identical target vertical angles were presented to both ears (closed gray circles for $\beta_l = \beta_r = 120^\circ$ and 150°). For $\beta_l = 180^\circ$ (panel (g)), most of the responses were localized to the vertical angle presented to the right ear.

Figure 8 shows the responses of subject SG. For the case in which $\beta_l = 0^\circ$ (panel (a)), the subject localized a sound image to the vertical angle presented to the right ear for $\beta_r = 30^\circ, 120^\circ, 150^\circ$, and 180° . Making a detailed observation, the subject localized a sound image around 180° for $\beta_r = 120^\circ$ and 150° . This tendency agrees with the responses obtained when identical target vertical angles were presented to both ears (closed gray circles for $\beta_l = \beta_r = 120^\circ$ and 150°). The subject localized a sound image to an angle intermediate to those presented to the left and right ears for $\beta_r = 60^\circ$ and 90° . In the case of $\beta_l = 30^\circ$

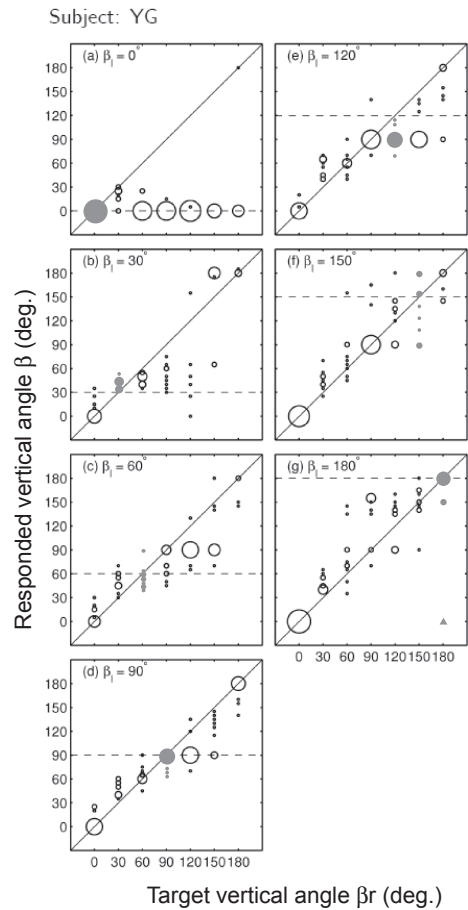


Fig. 9 Same as Fig. 6, but for subject YG.

(panel (b)), the subject localized a sound image to approximately the vertical angle presented to the right ear and occasionally to the vertical angle presented to the left ear for $\beta_r = 60^\circ$ and 90° . For $\beta_l = 60^\circ$ and 90° (panels (c) and (d), respectively), the subject localized most of the sound images to the vertical angle presented to the right ear and the remainder to that presented to the left ear. For $\beta_l = 120^\circ$ (panel (e)), the subject localized sound images to the vertical angle presented to the left ear for $\beta_r = 0^\circ$ to 90° and the remainder to that presented to the right ear. In the case of $\beta_l = 150^\circ$ (panel (f)), the subject localized a sound image to the vertical angle presented to the right ear for $\beta_r = 0^\circ$ and 180° and to that presented to the left ear for $\beta_r = 60^\circ$ to 120° . For $\beta_l = 180^\circ$ (panel (g)), most of the responses were localized to the vertical angle presented to the left ear, except for $\beta_r = 0^\circ$, for which the subject localized a sound image to the target vertical angle presented to the right ear.

Figure 9 shows the responses of subject YG. In the case of $\beta_l = 0^\circ$ (panel (a)), the subject localized a sound image

approximately to the vertical angle presented to the left ear and to an angle intermediate to those presented to the left and right ears for $\beta_r = 30^\circ$. In the case of $\beta_l = 30^\circ$ (panel (b)), the subject localized most of the sound images to the right ear for $\beta_r = 0^\circ$ and 180° and to an angle intermediate to those presented to the left and right ears for $\beta_r = 60^\circ$ to 120° . In the case of $\beta_l = 60^\circ$ and 90° (panels (c) and (d), respectively), the subject localized most of the sound images to the vertical angle presented to the right ear. However, the subject localized a sound image to approximately 90° for $\beta_r = 120^\circ$. This tendency agrees with the responses when identical target vertical angles are presented to both ears (closed gray circles for $\beta_l = \beta_r = 120^\circ$). In the case of $\beta_l = 120^\circ$ to 180° (panels (e) through (g), respectively), most of the responses are localized to the vertical angle presented to the right ear, while a few responses are localized to an angle intermediate to those presented to the left and right ears.

These results indicate that the subjects tend to localize a sound image to the target vertical angle presented to either the left or right ear, for cases in which they perceived a single sound image.

Figure 10 shows the responded lateral angle, α . The ordinate represents the responded lateral angle, and the abscissa represents the target vertical angle presented to the right ear, β_r . Most of the responses are distributed around 0° . A few of IT's and NS's responses are localized up to 20° from the median plane. This shift is thought to have been caused by a small interaural difference that arose from presenting different HRTFs to the left and right ears.

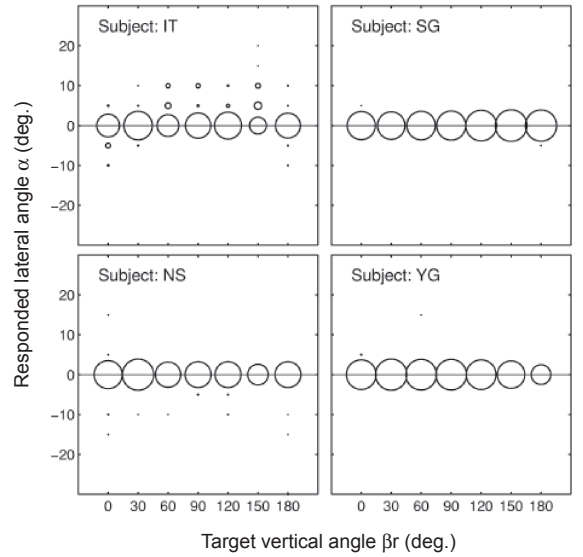


Fig. 10 Responded lateral angle, α , to stimuli for which the target vertical angles β_l and β_r differ.

Subjects IT, SF, and YG sometimes perceived two sound images. All responses to the stimuli, for which two sound images were perceived, are given in the order that the stimuli were presented (Figs. 11 through 13). The two values in parentheses in these figures denote the target vertical angles presented to the left and right ears (β_l and β_r , respectively). For comparison, the responded vertical angle when the HRTFs of identical vertical angles were presented to both ears

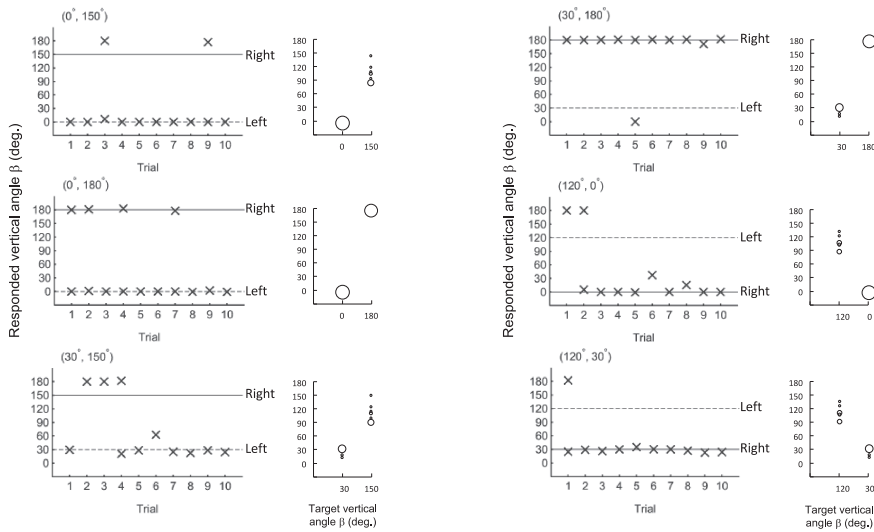


Fig. 11 Responses for subject IT localizing two sound images to stimuli for which the target angles presented to the left and right ears, β_l and β_r , differ. The values in parentheses denote β_l and β_r . For comparison, the responded vertical angle when the HRTFs of identical vertical angles were presented to both ears, shown in Fig. 4, is shown.

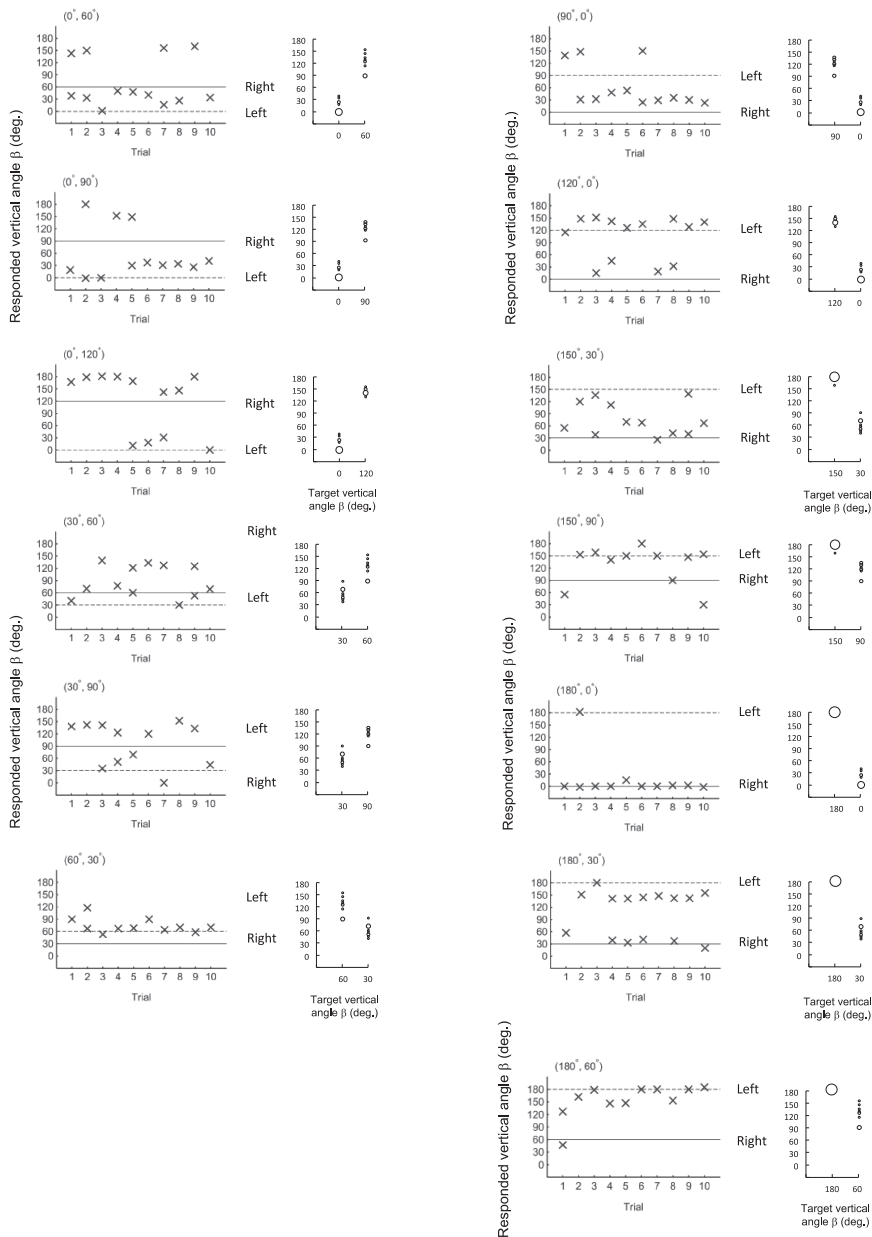


Fig. 12 Same as Fig. 11, but for subject SG.

presented to both ears, shown in Fig. 4, is shown beside each panel.

Figure 11 shows the responses of subject IT. Most of the responses localized two sound images to the target vertical angles presented to the left and right ears. A few responses are localized to 180° when the target vertical angle presented to the left or right ear was 120° or 150° .

Figure 12 shows the responses of subject SG. The localized vertical angles of two sound images are similar to

those when identical target vertical angles are presented to both ears. For instance, the responses tend to localize between 90° to 150° when the target vertical angle presented to either ear is 60° . This tendency agrees with the responses when a target vertical angle of 60° is presented to both ears.

Figure 13 shows the responses of subject YG. Most of the localized vertical angles of two sound images are similar to those when identical target vertical angles are

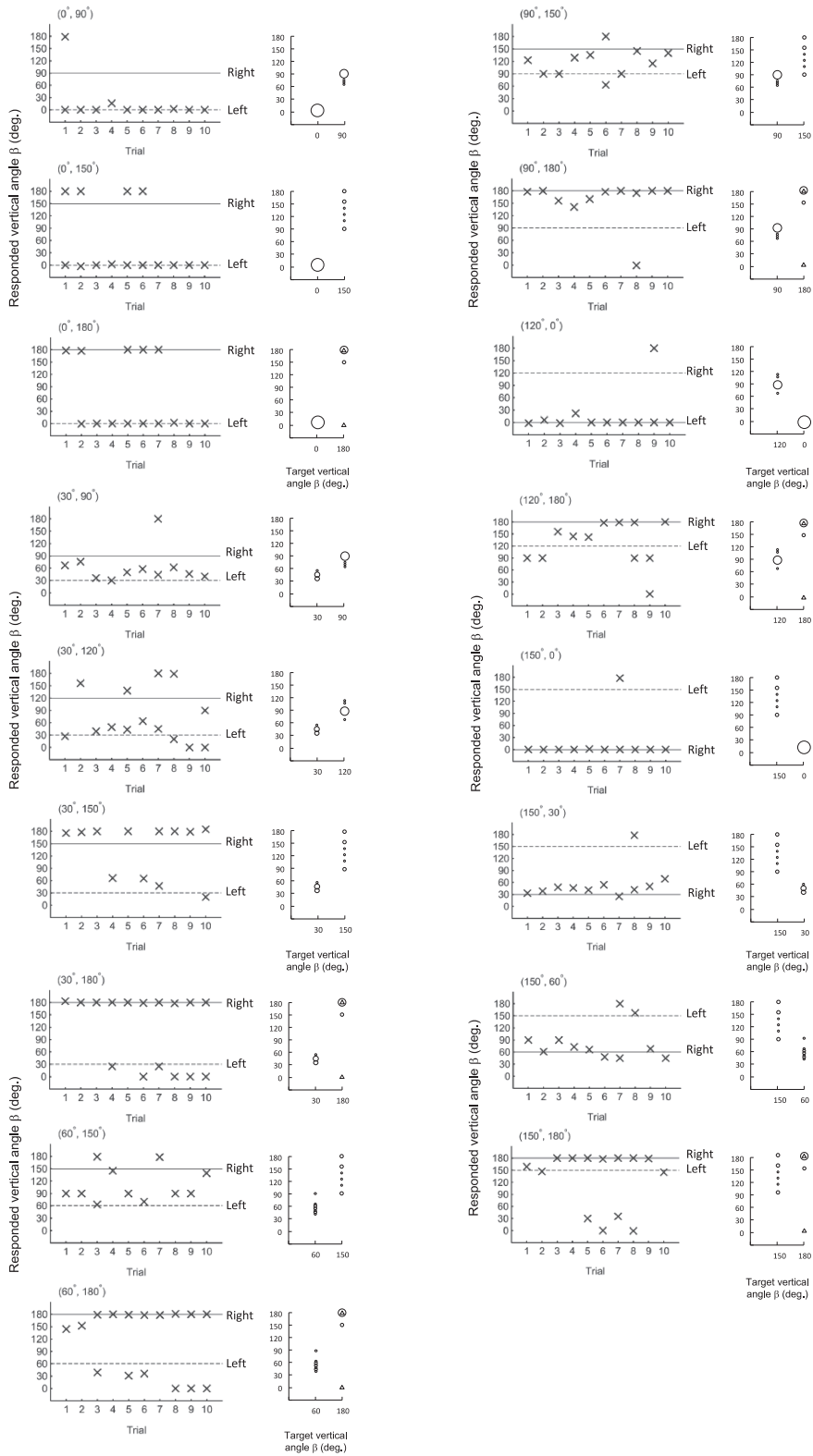


Fig. 13 Same as Fig. 11, but for subject YG.

Table 1 Mean localization error in the vertical angles for stimuli in which identical target vertical angles were presented to both ears ($\beta_l = \beta_r$) and for stimuli in which different target vertical angles were presented to each ear ($\beta_l \neq \beta_r$).

Subject	Mean localization error (deg.)		<i>t</i> -test
	$\beta_l = \beta_r$	$\beta_l \neq \beta_r$	
IT	11.6	9.3	—
NS	13.8	13.9	—
SG	26.3	17.5	*
YG	17.7	14.0	—

* $p < 0.05$

presented to both ears. For instance, the responses tend to localize to 0° when a target vertical angle of 180° is presented to the right ear. This tendency agrees with the responses obtained when a target vertical angle of 180° is presented to both ears. The triangle symbol indicates the subject occasionally perceived two sound images at 180° and 0° when the HRTFs of the vertical angle of 180° were presented to both ears, as shown in Fig. 4.

These results indicate that the subjects tend to localize two sound images to both target vertical angles, for cases in which they perceived two sound images.

3.2.3. Mean localization error

The mean vertical localization error was calculated. The mean vertical localization error is defined as the absolute difference between the responded and target vertical angles averaged over all of the responses.

Table 1 shows the mean localization error of the vertical angles for stimuli in which identical target vertical angles were presented to both ears ($\beta_l = \beta_r$) and for stimuli in which different target vertical angles were presented to each ear ($\beta_l \neq \beta_r$). The smaller error for the two target vertical angles was adopted for the stimuli $\beta_l \neq \beta_r$. Both errors were obtained when the subjects localized two sound images. The chance levels for $\beta_l = \beta_r$ and $\beta_l \neq \beta_r$, in which the smaller error for the two target vertical angles was adopted, were 65.0 (deg) and 36.7 (deg), respectively. The mean localization errors for both $\beta_l = \beta_r$ and $\beta_l \neq \beta_r$ were each smaller than half the chance level for all four subjects. The mean localization error for $\beta_l \neq \beta_r$ was less than that for $\beta_l = \beta_r$, except for subject NS, whose errors for $\beta_l \neq \beta_r$ and $\beta_l = \beta_r$ were approximately the same. This is probably due to adopting the smaller error in the case of stimuli $\beta_l \neq \beta_r$.

Then, a *t*-test was performed in order to determine whether the difference in the localization error between $\beta_l = \beta_r$ and $\beta_l \neq \beta_r$ is statistically significant. With the exception of subject SG, there was no statistically significant difference between the localization errors for $\beta_l \neq \beta_r$ and $\beta_l = \beta_r$. Even for subject SG, the error for $\beta_l \neq \beta_r$ was less than that for $\beta_l = \beta_r$.

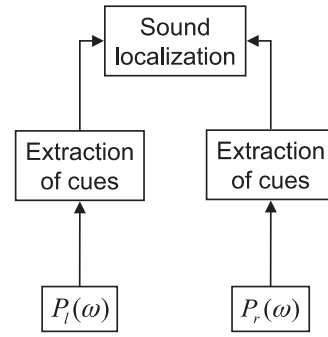


Fig. 14 A possible extraction process of spectral cues in the human auditory system. The spectral cues are independently extracted from the spectrum of the input signal to each ear, $P_l(\omega)$ and $P_r(\omega)$.

These results showed that the subjects localized a single sound image to the target vertical angle presented to either the left or right ear or localized two sound images to both target vertical angles, when different target vertical angles were presented to the left and right ears. These results imply that a listener perceives the vertical angle of a sound image with the spectral information at only a single ear.

4. DISCUSSION

Based on the results shown in Sect. 3, a possible extraction process for spectral cues in the human auditory system would be that spectral cues are extracted from the spectrum of the input signal to each ear independently (Fig. 14).

Middlebrooks [11] proposed a model in which a sound image is localized to the vertical angle at which the cross-correlation coefficient between the spectrum of each ear-input signal and the spectrum of the HRTF is maximized. A similar model was proposed by Chung *et al.* [24]. These models support the validity of the proposed extraction process.

Furthermore, it is known that the dorsal cochlear nucleus (DCN) extracts the spectral notch in the HRTF, and the DCN type IV neurons encode rising spectral edges of the notch [25]. These findings support the hypothesis that the human auditory system extracts the spectral cues from the spectrum of the input signal to each ear independently.

However, it is not clear which of the ears is dominant in the extraction process. Moreover, what determines the dominance remains unclear. Figure 15 shows the amplitude spectrum of the HRTF of subject IT for the left ear and the right ear. The figure indicates that the amplitude spectrum of the HRTF were similar between the left and right ears. Therefore, the dominance of one ear cannot be

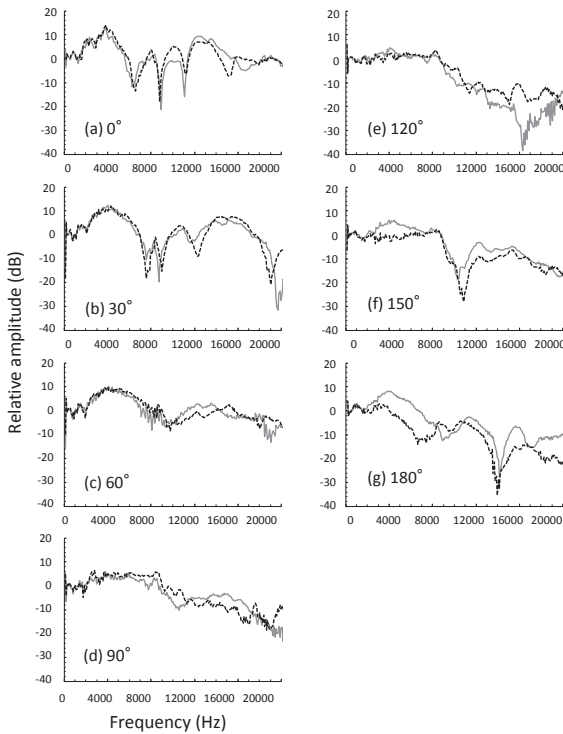


Fig. 15 Amplitude spectrum of the HRTF of subject IT in the upper median plane for the left ear and the right ear. The solid lines and the dotted lines indicate the HRTFs of the left and right ears, respectively.

explained by more salient spectral characteristics than the other ear. This is a problem to be solved in the near future.

5. CONCLUSIONS

In order to examine whether the vertical localization accomplished with the spectral information at only a single ear or the vertical localization requires the spectral information at both ears, we carried out upper median plane localization tests presenting HRTFs of different target vertical angles to the left and right ears.

The results showed that the subjects either localized a single sound image to the target vertical angle presented to either the left or right ear or localized two sound images to both target vertical angles, with approximately the same accuracy as when identical target vertical angle of HRTFs were presented to both ears. These results imply that the spectral information at only a single ear provides the spectral cue for vertical localization.

ACKNOWLEDGEMENTS

The authors would like to thank E. Rin for his assistance in conducting the sound localization tests.

REFERENCES

- [1] J. Hebrank and D. Wright, "Spectral cues used in the localization of sound sources on the median plane," *J. Acoust. Soc. Am.*, **56**, 1829–1834 (1974).
- [2] R. A. Butler and K. Belendiuk, "Spectral cues utilized in the localization of sound in the median sagittal plane," *J. Acoust. Soc. Am.*, **61**, 1264–1269 (1977).
- [3] S. Mehrgardt and V. Mellert, "Transformation characteristics of the external human ear," *J. Acoust. Soc. Am.*, **61**, 1567–1576 (1977).
- [4] A. Musicant and R. Butler, "The influence of pinnae-based spectral cues on sound localization," *J. Acoust. Soc. Am.*, **75**, 1195–1200 (1984).
- [5] E. A. G. Shaw and R. Teranishi, "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," *J. Acoust. Soc. Am.*, **44**, 240–249 (1968).
- [6] V. C. Raykar, R. Duraiswami and B. Yegnanarayana, "Extracting the frequencies of the pinna spectral notches in measured head-related impulse responses," *J. Acoust. Soc. Am.*, **118**, 364–374 (2005).
- [7] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura and K. Iida, "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," *J. Acoust. Soc. Am.*, **132**, 3832–3841 (2012).
- [8] K. Iida, S. Nishioka and Y. Ishii, "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," *J. Acoust. Soc. Am.*, **136**, 317–333 (2014).
- [9] F. Asano, Y. Suzuki and T. Sone, "Role of spectral cues in median plane localization," *J. Acoust. Soc. Am.*, **88**, 159–168 (1990).
- [10] J. C. Middlebrooks, "Narrow-band sound localization related to external ear acoustics," *J. Acoust. Soc. Am.*, **92**, 2607–2624 (1992).
- [11] J. C. Middlebrooks, "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," *J. Acoust. Soc. Am.*, **106**, 1493–1510 (1999).
- [12] A. Kulkarni and H. S. Colburn, "Role of spectral detail in sound-source localization," *Nature*, **396**, 747–749 (1998).
- [13] E. H. A. Langendijk and A. W. Bronkhorst, "Contribution of spectral cues to human sound localization," *J. Acoust. Soc. Am.*, **112**, 1583–1596 (2002).
- [14] E. A. Macpherson and A. T. Sabin, "Vertical-plane sound localization with distorted spectral cues," *Hear Res.*, **306**, 76–92 (2013).
- [15] K. Iida, M. Itoh, A. Itagaki and M. Morimoto, "Median plane localization using parametric model of the head-related transfer function based on spectral cues," *Appl. Acoust.*, **68**, 835–850 (2007).
- [16] M. B. Gardner, "Some monaural and binaural facets of median plane localization," *J. Acoust. Soc. Am.*, **54**, 1489–1495 (1973).
- [17] M. Morimoto, "The contribution of two ears to the perception of vertical angle in sagittal planes," *J. Acoust. Soc. Am.*, **109**, 1596–1603 (2001).
- [18] R. A. Butler, "Monaural and binaural localization of noise bursts vertically in the median sagittal plane," *J. Aud. Res.*, **3**, 230–235 (1969).
- [19] C. L. Searle, L. D. Braida, D. R. Cuddy and M. F. Davis, "Binaural pinna disparity: Another auditory localization cue," *J. Acoust. Soc. Am.*, **57**, 448–455 (1975).
- [20] J. Hebrank and D. Wright, "Are two ears necessary for localization of sound sources on the median plane?," *J. Acoust. Soc. Am.*, **56**, 935–938 (1974).

- [21] M. Morimoto and K. Nomachi, "Binaural disparity cues in median-plane localization," *J. Acoust. Soc. Jpn. (E)*, **3**, 99–103 (1982).
- [22] M. Morimoto and H. Aokata, "Localization cues of sound sources in the upper hemisphere," *J. Acoust. Soc. Jpn. (E)*, **5**, 165–173 (1984).
- [23] H. Møller, D. Hammershøi, C. J. Jensen and M. F. Sørensen, "Transfer characteristics of headphones measured on human ears," *J. Audio Eng. Soc.*, **43**, 203–217 (1995).
- [24] W. Chung, S. Carlile and P. Leong, "A performance adequate computational model for auditory localization," *J. Acoust. Soc. Am.*, **107**, 432–445 (2000).
- [25] L. A. J. Reiss and E. D. Young, "Spectral edge sensitivity in neural circuits of the dorsal cochlear nucleus," *J. Neurosci.*, **25**, 3680–3691 (2005).