Kim, Taeho; Pöntynen, Henri; Pulkki, Ville

Directional-band-dominant spectral region for the sound localisation in the median plane

Published in:
Proceedings of the 24th International Congress on Acoustics: A11: Psychoacoustics

Published: 24/10/2022

Document Version
Publisher's PDF, also known as Version of record

PROCEEDINGS of the
24th International Congress on Acoustics

October 24 to 28, 2022 in Gyeongju, Korea

Directional-band-dominant spectral region for the sound localisation in the median plane

Taeho Kim[1]; Henri Pöntynen[2]; Ville Pulkki[3 1]

[1] Acoustics Lab, Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

ABSTRACT

Previous studies have reported that the direction of a band-limited sound stimulus in the median plane is localised based on its centre frequency instead of its actual location. The frequency band that determines this localisation is referred to as the directional-band. However, since most relevant studies employed a coarse localisation response scale or a limited range of frequency bands, the precise localisation sensitivity of each frequency band over the whole audible range is not yet clearly established. Therefore, this paper employed a comprehensive approach to the localisation of band-limited stimuli, utilising a continuous response scale within the whole median plane circle and consecutive 1/3-octave frequency bands within 100 Hz-16 kHz in the listening test. The results show that the 'directional-bands' fall exclusively within the spectral region between 2.5kHz and 8kHz. Moreover, a 'pitch-height effect' is observed within this spectral region, correlating to the elevation range of around 20° to 90°. Although previous studies find directional-bands dispersed over a wide frequency range, not limited to a particular region, the responses in the directional-band-dominant spectral region from this study were statistically distinct against the other spectral regions for all subjects and source directions.

Keywords: Psychoacoustics, Directional-band, Localisation, Median plane

1. INTRODUCTION

Directional hearing of sound sources has been studied extensively until the present. While sound localisation in the lateral direction relies on interaural disparities (ITD, ILD), the factors that enable localisation in the vertical direction are not fully established. However, in the vertical auditory mid-line between the ears (the median plane), it is generally known that direction-dependent spectral cues mainly contribute to the localisation (1, 2). The studies investigating the effect of spectral cues on vertical localisation can be generally categorised into two groups. One group of studies focuses on analysing the head-related transfer function (HRTF), which accentuates the spectral notch(es) at high frequencies (e.g. (1, 3, 4, 5, 6, 7, 8, 9)). These spectral notches, which appear due to the reflection of sound waves at the outer ear (pinna), are elevation-dependent since their centre frequencies shift systematically within 6-9 kHz as the sound source's elevation varies (9, 10, 11). However, despite the apparent correlation between spectral notches and sound source elevation, localisation in the median plane cannot be explained exclusively by pinna-related spectral cues. Many studies have noticed the contribution of low-frequency features below the spectral notch region (e.g. below 4 kHz) in median plane localisation (e.g. (12, 13, 14)). Moreover, localisation seems to be apparent even with severely degraded pinna cues (15, 16). Likewise, Zonooz et al.(17) reported that mislocalisation in the median plane occurred consistently when the spectral notch region was highly emphasised over the rest of the frequencies, while the best localisation performance was reported for the entire frequency range without any emphasised spectral region. In addition, Head movement is also reported to enhance elevation perception (18, 19).

Another group of studies focuses on the contribution of a spectral region or spectral band to median plane localisation. This topic has been researched with numerous methods. One method is to modify the spectral curve of a specific frequency band to investigate the influence of the selected band on elevation perception. For instance, Asano et al. (20) applied several smoothing filters to a spectral region above

[1] taeho.kim@aalto.fi

a boundary frequency that varied between 500 Hz and 5 kHz. As a result, the cues for the elevation perception were reported to be the macroscopic curve features above 5 kHz in the transfer function. Langendijk and Bronkhorst (7) employed a similar method to their localisation experiment by flattening the curves within frequency regions with different bandwidths (1/2, 1 and 2-oct.) and spectral positions (low, mid and high) above 4 kHz. The result showed that the elevation cues are located in the 6-12 kHz frequency region, which roughly agrees with (20). However, these studies only applied a limited number of high-frequency bands.

Another line of research have applied the research method examining the effect of band-limited stimuli on elevation perception. This method was introduced in Blauert's experiment (21), where he investigated the elevation perception using 1/3-octave bandwidth noises between 125 Hz and 16 kHz. He reported that elevation perception was solely determined by certain frequency bands associated with specific sound directions (e.g. 500 Hz and 4 kHz for 'front', 1 kHz for 'rear' and 8 kHz for 'above' location), not the actual sound location. Those frequency bands were designated as *directional-bands*. Blauert's directional-band theory has been approved, despite some disparities in the research results, in many subsequent studies (3, 22, 23, 24). However, the localisation results reported from those directional-band studies were approximate because the results may have been affected by the fact that they used a relatively coarse response scale and a limited number of frequency bands. For instance, the response scales employed by those studies were three directional regions in the upper hemisphere in (21), nine regions from -30° to 210° in (22) and eight regions in the whole circle in (24). Although a continuous response scale was employed by Middlebrooks (3), he utilised only four discontinuous 1/6-octave frequency bands (6, 8, 10 and 12 kHz) in the vertical localisation experiment. On the other hand, Itoh et al.'s experiment (23) showed a continuous response scale with test stimuli of 1/3 and 1/6-Oct. successive frequency bands within a wide frequency range (800 Hz to 12.5 kHz). This fine frequency resolution allowed to find that the center frequencies of the directional-bands may vary across individuals. However, despite sufficient response data, they analysed the directional-bands applying Blauert's method (21) which utilised only three response regions ('front', 'above' and 'rear').

Therefore, this paper applies a comprehensive approach to the median plane localisation of band-limited stimuli, utilising sufficient frequency resolution to find individual differences, and a continuous response scale within the whole circle on the median plane, which may reveal smaller dependencies on elevation perception and center frequency. Consequently, this study examines the contribution aspects of each frequency band to the elevation perception, thus identifying the directional-bands and examining the localisation precision that might vary with each frequency region and source direction.

2. Methods

Localisation tests were conducted to examine the elevation perception of 1/3-octave white noise in the median plane.

2.1 Subjects

Nine subjects with normal hearing, eight male and one female, participated in the listening tests. Participants were staff members or students of Aalto University aged between 21 and 45.

2.2 Apparatus

Listening tests were carried out in an anechoic chamber in Aalto Acoustics Lab. Sound stimuli were presented from one of three loudspeakers positioned at 0°, 90° and 180° elevation from the listening position located in the centre of the chamber. The distance from each speaker to the listening position was 2,044 mm. A headrest was attached to the chair located in the listening position to prevent the subjects from moving their heads. The chamber was darkened entirely during the test to avoid biases from the actual speaker location. However, a luminous tape strip was attached to the 'front' speaker (0° elevation) for the listeners to identify the front centre location as the reference direction in the darkness.

2.3 Stimuli

Twenty-one 1/3-octave band white noises and a broadband white noise were utilised for the sound stimuli. The centre frequencies of narrow-band noises ranged from 125 Hz to 12.5 kHz. The 1/3-octave bandwidth was selected in this test based on the previous studies, reporting that 1/3-octave bandwidth is sufficient to examine the variance of centre frequencies compared to narrower bandwidths in band-limited stimuli localisation (23, 24). Each stimulus consisted of two noise bursts of 20 ms length and an interval of 80 ms between the bursts.

2.4 Procedure

The listening test was carried out with two identical sessions to avoid fatigue from the subjects. The stimuli were presented in one of the loudspeakers located in three different elevation angles: 0° for the front, 90° for the above and 180° for the rear loudspeaker. Each of 22 sound samples from each stimulus direction was repeated four times for each session; thus, the overall stimuli presentation was 528 times per subject. All stimuli were presented in random order, and the sound pressure level was also randomised within 69-71 dB SPL, A weighted. A remote GUI device was provided to the subjects during the tests, which enabled to record the perceived elevation angle by moving a circular slider shown in the device. Subjects were instructed to fasten their head to the headrest behind, facing the reference location (0° elevation in the centre).

3. Results and discussion

Examination on the contribution of each frequency band to the elevation perception requires the distribution aspects of response data for each frequency band. Figure 1 shows the box plots for each stimulus from each source direction.

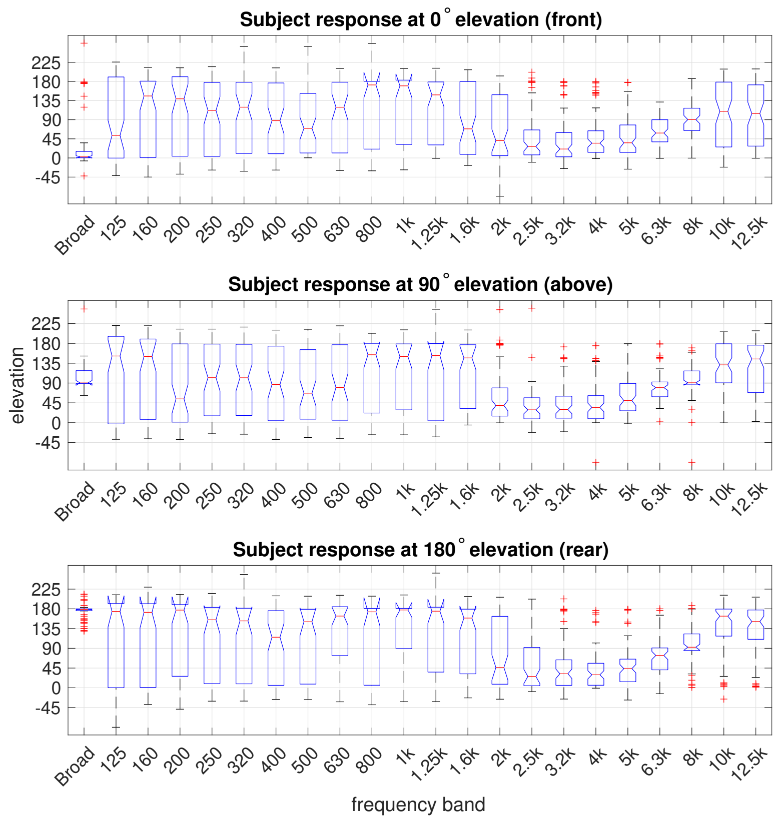

Figure 1: Box-plots presenting the distribution of response data for each 1/3-octave frequency band and broadband white noise from three incidence directions.

Figure 1 shows that, band-limited noises are rarely localised to the source direction. The *error rates*, derived with a coarse response scale of 45°, were above 50% from nearly all frequency bands and source directions, with only two exceptions: 41.7% for the 8kHz band from the 'above' and 34.7% for the 1kHz band from the 'rear' direction, which is in line with Blauert's directional-bands (21). On the other hand, the error rates of the broadband white noise were 22.2%, 37.5% and 22.2% for 'front', 'above' and 'rear', respectively.

However, despite poor localisation responses from band-limited noises, a significant drop in the interquartile range is observed in a frequency region ranging from 2.5 kHz to 8 kHz from all three source directions (see Figures 1 and 2). This result is relevant to the corresponding median values from three source directions in the 2-8 kHz region, shown in Figure 2. These consistent patterns in the median values and interquartile ranges indicate that the responses roughly converge to specific perceived elevations in the 2.5-8 kHz frequency region regardless of the source directions, implying that those frequency bands within such region might function as *directional-bands*.

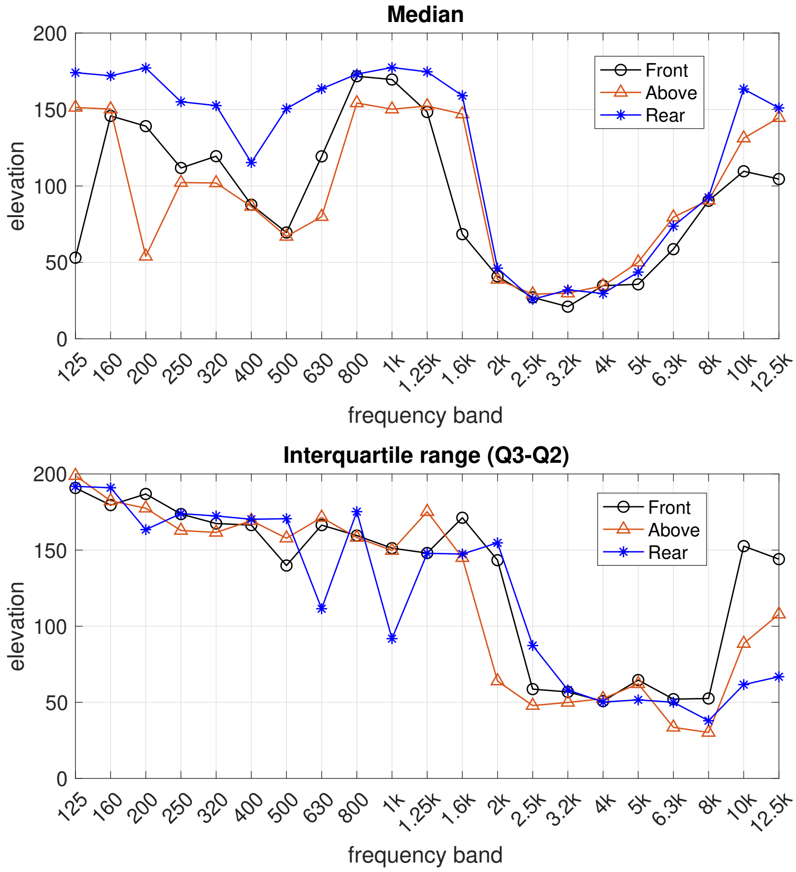

Figure 2: The median values and interquartile ranges of the response data.

The consistent pattern in the 2.5-8 kHz frequency region appears more prominently in the individual response distribution plots. Figure 3 shows the box plots from four test participants, which generally appear similar to the box plots from overall response data (Figure 1). Even the exceptional case from Subject 3, perceiving every frequency band from 125 Hz to 1.6 kHz to the 'rear' direction, shows low interquartile ranges and an increasing trend of median value in the 2.5-8 kHz region. One noticeable finding in the individual box plots is some subjects' localisation to the 'rear' direction around 1 kHz. The box plots of Subjects 3,5, and 6 show median values of about 180° and low interquartile ranges around 1 kHz. Furthermore, Figure 2 shows that the median values around the 1 kHz frequency band (800-1.25 kHz) appear within 150°-180° elevation, although the variance around 1 kHz is much higher than the 2.5-8 kHz frequency region for the overall responses. This pattern agrees with Blauert's directional-band study (21).

To evaluate the distinctiveness of the directional-band region (2.5-8 kHz), significant differences between the directional-band region and the frequency bands outside that region were analysed with 1-way ANOVA and multiple comparisons between each frequency band. First, 1-way ANOVA was conducted to each frequency band's responses from the front, above, and rear directions and overall responses from all directions, resulting in $p < 1.0e-19$ for all conditions. Next, pairwise comparisons were made between each frequency band's responses using a multiple comparison method, the *honestly significant difference* test (HSD) by Tukey (25). Specifically, responses of each frequency band within

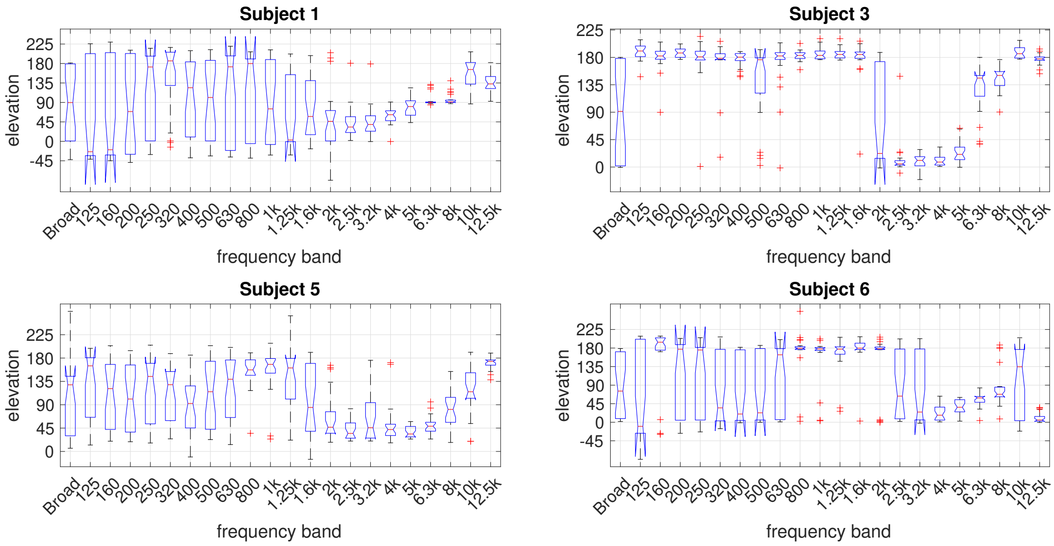

Figure 3: Box-plots presenting the distribution of response data for each 1/3-octave frequency band and broadband white noise from four subjects. Three source directions are pooled into a single box plot for each subject.

the directional-band region (2.5-8 kHz) were compared with every frequency band's responses outside that region (125 Hz-2 kHz, 10 kHz-12,5 kHz). The results show that the means of 14 frequency bands among 15 bands (93.3%) outside the directional-band region are significantly different from the means of 2.5 kHz, 3.2 kHz, 4 kHz and 5 kHz frequency bands in the overall responses from all directions.

The results of the multiple comparison test from each conditions (front, above, rear and all directions) are shown in Table 1. It presents the percentage of frequency bands outside the directional-band region whose means are significantly different ($p < 0.05$) from those of each directional-band (a frequency band within the directional-band region). These results indicate that the responses from the directional-band region are prominently distinguishable from the outside regions. Furthermore, the distinction of the directional-band region was relatively higher when the source directions were 'above' and 'rear' compared to the 'front' direction. The exception in the 8-kHz frequency band can be explained by the assumption that the highly dispersed response data in the non-directional-band region causes the 'mean' values to be located around the middle, which is close to the mean values of 8 kHz frequency band responses: 90° - 100° elevation. Nevertheless, the multiple comparison analysis within the directional-band region (see Table 2) shows that most responses of the 8-kHz frequency band are significantly different ($p < 0.05$) from the others, which implies that the 8-kHz frequency band is the most distinct in the elevation perception.

Table 1: the percentage of non-directional frequency bands whose responses are significantly different from each directional-band.

Direction	2.5 kHz	3.2 kHz	4 kHz	5 kHz	6.3 kHz	8 kHz
Front	66.7%	80%	66.7%	66.7%	20%	0%
Above	93.3%	80%	80%	46.7%	6.7%	0%
Rear	86.7%	93.3%	93.3%	93.3%	40%	0%
Overall	93.3%	93.3%	93.3%	93.3%	73.3%	20%

On the other hand, besides the distinctiveness of the directional-band region, the increasing perceived elevation in the median values implies a "pitch-height effect" reported by Roffler and Butler (26). Although they examined the vertical localisation using tonal sounds, the narrow-band noise in the current study also shows an increasing pattern from 20° to around 90° elevation at 3.2-8 kHz, which is contained within the directional-band region (see Figure 2). Statistical analysis was conducted using *Pearson's*

Table 2: the percentage of significantly different frequency bands within the directional-band region.

Direction	2.5 kHz	3.2 kHz	4 kHz	5 kHz	6.3 kHz	8 kHz
Front	0%	0%	0%	0%	0%	20%
Above	40%	20%	20%	0%	20%	60%
Rear	20%	20%	20%	20%	0%	80%.
Overall	40%	40%	40%	20%	60%	80%.

correlation method, resulting in a moderate correlation between the centre frequency and perceived elevation. The correlation coefficients were 0.314 (p = 1.14e-9), 0.442 (p = 1.36e-18) and 0.375 (p = 1.72e-13) from the front, above and rear source directions, respectively.

4. General discussion

The median and variance values of the present study generally agree with the directional-bands suggested in previous studies (e.g. (21, 23, 24)). For instance, the median values in the 2.5-4 kHz frequency region are observed around 20°-30° elevation, which roughly agrees with Blauert's directional-band around 3 kHz, presenting the 'front' direction. Likewise, the median values at the 8 kHz frequency band show the coinciding result of around 90°, which also agrees with the directional-band of 8 kHz for the 'above' direction from the previous directional-band studies and Hebrank's study (22). Furthermore, Blauert's directional-band of 1 kHz presenting the 'rear' direction accords with the median values around 1 kHz (800-1.25 kHz) in the present study, despite high variance values. The response results from each individual (Figure 3) indicate that the directional-band for the 'rear' direction shows higher individual disparity than other directional-bands.

However, whereas previous studies suggested the directional-bands with discrete localisation directions and specific centre frequencies (e.g. 500 Hz and 4 kHz for 'front', 8 kHz for 'above' and 1 kHz for 'rear'), the present study finds directional-band dominant region (2.5-8 kHz) appeared as a continuously increasing curve ranging from 20° to 90°. Therefore, it is feasible to regard all the frequencies within this frequency range, not only the specific directional-band centre frequencies, as spectral cues correlated to the elevation perception of 20°-90° ('front' to 'above'). Furthermore, the moderate correlation of the continuous spectral curve to the perceived elevation might enable precise control of median plane localisation. Considering Blauert's hypothesis that the perceived elevation is determined by the relative distribution of sound energy in directional-bands (1), weighting the magnitude of a frequency within the directional-band region would provide accurate localisation to the correlated elevation. Since simulating the median plane localisation by frequency-weighting has been already evaluated with decent results by recent studies (27, 28), it will also be valuable to evaluate the localisation by frequency-weighting within the directional-band region in future studies.

Nevertheless, this directional-band dominant spectral region and its correlation to the perceived elevation do not exclusively function as spectral cues for the median plane localisation. Despite individual differences, the directional-band around 1 kHz for the 'rear' direction is validated in the present study. Furthermore, low-frequency elements due to the reflections from the head and shoulders also operate as spectral cues (12). Most of all, the spectral notches from pinna reflections play a crucial role in the median plane localisation. However, when simulating the median plane localisation utilising generic HRTFs, the individual differences in the *pinna cues* can lead to inaccurate elevation perception (28, 29). Since the directional-band region in the present study presents low variance values implying fewer individual differences, the frequency weighting within this spectral region can be an effective simulation method enabling more precise localisation in the median plane.

5. Conclusions

This study examined the median plane localisation of band-limited stimuli, applying a comprehensive approach utilising consecutive 1/3-octave frequency bands, three source directions at 'front', 'above' and 'rear', and a continuous response scale within the whole circle on the median plane. The results are summarised below.

1. The localisation responses from the frequency region at 2.5-8 kHz showed low variance and coinciding median values for all source directions and most subjects, which can be regarded as a directional-band-dominant spectral region.
2. The responses around 1 kHz showed median values to the 'rear' for all source directions, but variance was high due to individual differences.
3. The responses from the directional-band-dominant region were statistically distinct from other spectral regions.
4. The directional-band-dominant region appeared as a *continuously increasing curve*, which is different from previous directional-bands dispersed over discrete center frequency positions.
5. The increasing curve within the directional-band region showed a moderate correlation to the elevation from 20° to 90° for all source directions.

REFERENCES

1. Blauert J. Spatial Hearing: The Psychophysics of Human Sound Localization. MIT Press; 1997.
2. Moore BCJ. An Introduction to the Psychology of Hearing. vol. 6. Boston Academic Press; 2012.
3. Middlebrooks JC. Narrow-band sound localization related to external ear acoustics. The Journal of the Acoustical Society of America. 1992;92(5):2607-24.
4. Shaw EAG, Teranishi R. Sound Pressure Generated in an External-Ear Replica and Real Human Ears by a Nearby Point Source. The Journal of the Acoustical Society of America. 1968 jul;44(1):240-9.
5. Butler RA, Belendiuk K. Spectral cues utilized in the localization of sound in the median sagittal plane. The Journal of the Acoustical Society of America. 1977 may;61(5):1264-9.
6. Bloom PJ. Creating Source Elevation Illusions by Spectral Manipulation. J Audio Eng Soc. 1977 sep;25(9):560-5.
7. Langendijk EH, Bronkhorst AW. The contribution of spectral cues to human sound localization. The Journal of the Acoustical Society of America. 1999;105(2):1036-6.
8. Raykar VC, Duraiswami R, Yegnanarayana B. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. The Journal of the Acoustical Society of America. 2005 jul;118(1):364-74.
9. Iida K, Itoh M, Itagaki A, Morimoto M. Median plane localization using a parametric model of the head-related transfer function based on spectral cues. Applied Acoustics. 2007 aug;68(8):835-50.
10. Takemoto H, Mokhtari P, Kato H, Nishimura R, Iida K. Mechanism for generating peaks and notches of head-related transfer functions in the median plane. The Journal of the Acoustical Society of America. 2012;132(6):3832-41.
11. Iida K, Ishii Y, Nishioka S. Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae. The Journal of the Acoustical Society of America. 2014 jul;136(1):317-33.
12. Algazi VR, Avendano C, Duda RO. Elevation localization and head-related transfer function analysis at low frequencies. The Journal of the Acoustical Society of America. 2001 mar;109(3):1110-22.
13. Otani M, Iwaya Y, Magariyachi T, Suzuki Y. Spatial distribution of low-frequency head-related transfer function spectral notch and its effect on sound localization. Principles and Applications of Spatial Hearing. 2011:195-204.
14. Morimoto M, Yairi M, Iida K, Itoh M. The role of low frequency components in median plane localization. Acoustical Science and Technology. 2003;24(2):76-82.
15. Pulkki V, Pöntynen H, Santala O. Spatial perception of sound source distribution in the median plane. AES: Journal of the Audio Engineering Society. 2019;67(11):855-70.
16. Macpherson EA, Middlebrooks JC. Vertical-plane sound localization probed with ripple-spectrum noise. The Journal of the Acoustical Society of America. 2003;114(1):430-45.
17. Zonooz B, Arani E, Körding KP, Aalbers PATR, Celikel T, Van Opstal AJ. Spectral Weighting Underlies Perceived Sound Elevation. Scientific Reports. 2019;9(1):1-12.
18. Perrett S, Noble W. The contribution of head motion cues to localization of low-pass noise. Percept Psychophys. 1997;59:1018-26.
19. Pöntynen H, Salminen NH. Resolving front-back ambiguity with head rotation: The role of level dynamics. Hearing research. 2019;377:196-207.

20. Asano F, Suzuki Y, Sone T. Role of spectral cues in median plane localization. Journal of the Acoustical Society of America. 1990;88(1):159-68.
21. Blauert J. Sound localization in the median plane. Acta Acustica united with Acustica. 1969;22(4):205-13.
22. Hebrank J, Wright D. Spectral cues used in the localization of sound sources on the median plane. The Journal of the Acoustical Society of America. 1974 dec;56(6):1829-34.
23. Itoh M, Iida K, Morimoto M. Individual differences in directional bands in median plane localization. Applied Acoustics. 2007;68(8):909-15.
24. Wallis R, Lee H. Directional bands revisited. 138th Audio Engineering Society Convention 2015. 2015;1:495-501.
25. Abdi H, Williams LJ. Tukey's honestly significant difference (HSD) test. Encyclopedia of research design. 2010;3(1):1-5.
26. Roffler SK, Butler RA. Localization of Tonal Stimuli in the Vertical Plane. The Journal of the Acoustical Society of America. 1968;43(6):1260-6.
27. Kim T, Pöntynen H, Pulkki V. Vertical Direction Control Using Difference-Spectrum Filters in Stereophonic Loudspeaker Reproduction. Journal of the Audio Engineering Society. 2022 3;70:128-39.
28. Rajendran VG, Gamper H. Spectral manipulation improves elevation perception with non-individualized head-related transfer functions. The Journal of the Acoustical Society of America. 2019;145(3):EL222-8.
29. Mendonça C, Campos G, Dias P, Vieira J, Ferreira JP, Santos JA. On the Improvement of Localization Accuracy with Non-Individualized HRTF-Based Sounds. J Audio Eng Soc. 2012;60(10):821-30.