Audio Engineering Society

Convention Paper

Presented at the AES 161st Convention

2026 October 30-November 1, Nashville, TN, USA

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib) all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Perception of Elevation on the Medial Plane Using Individualized HRTFs

Cayleb Martinez[1], Wesley Bulla[1], and Song Hui Chon[1]

[1] *Belmont University, 1900 Belmont Boulevard, Nashville, TN 27212.*

Correspondence should be addressed to Cayleb Martinez (CaylebJMartinez@gmail.com)

ABSTRACT

It is widely accepted that auditory localization is superior with individualized head related transfer functions (HRTFs) compared to that with generic HRTFs. However, research also suggests that listeners can adapt to non-individual HRTFs due to perceptual "plasticity." This study investigated if subjects' perception of elevation on the medial plane was better when using their own individualized HRTFs or someone else's individualized HRTFs. Individual HRTFs from three vertical angles on the medial plane were captured using binaural microphones on 10 subjects. These 30 recordings were played back to the same subjects. For each stimulus, subjects were asked to determine from which location the sound originated: "above," "ear-level," or "below." Results showed that only one subject achieved significantly higher localization accuracy with their own HRTFs.

1 Introduction

Sound localization is based on three acoustic factors: interaural time differences (ITD), interaural level differences (ILD), and spectral cues [1, 2]. These three co-existent variables, which are unique for each individual, are referred to collectively as the Head Related Transfer Function (HRTF) [3, 4]. Studies have shown that it is easier for listeners to localize sounds using their own individualized HRTFs compared to so-called generic HRTFs [5, 6, 7]. Localization using generic HRTFs can be improved with training because of "the high degree of cross-modal plasticity" of the human auditory sensory system [8, 9, 10]. However, there may be a possible ceiling effect on training improvement using generic HRTFs [11], and some generic HRTFs may be more effective than others [12].

Based on previous research, a hypothesis was formed that subjects would localize better using their own individualized HRTFs on the medial plane than with another person's HRTFs.

2 Methods

2.1 Stimuli

HRTF data were collected from 10 subjects in a semi-anechoic chamber using DPA4560 binaural headset microphones and a playback system compliant with ITU BS.1116-3 specifications [13]. Subjects were seated two meters from a loudspeaker positioned in front of them at the ear level. Each participant tilted their head in two directions, up and down 30˚, to record three HRTFs as an approximation of ear-level, higher, and lower elevation cues respectively. Three pink noise bursts were played at 60 dBA in succession for each position and recorded as 48 kHz 24 bit .wav files.

2.2 Procedure

A blind listening test was conducted with the same 10 subjects. Samples were matched to -11 dB FS and presented binaurally using open-back Sennheiser HD598 headphones. Subjects were asked to indicate their perceived elevation position of each sample. Response options presented were "above," "ear-level," and "below." Ten sets of 30 individualized HRTFs were played back in a randomized order [14], resulting in a total of 300 trials per person. Two-minute breaks were implemented between sets to minimize ear fatigue.

2.3 Subjects

Ten audio engineering graduate students participated. All and had completed a graduate course in critical listening. In a general sense, these subjects would be considered more experienced than "novice" but not yet "expert" listeners.

3 Results

This experiment utilized a Three-Alternative Forced Choice (3AFC) paradigm, yielding trinomial data with probability reflected as 0.33 [15]. Only Subject 3 demonstrated reliable localization accuracy with their own HRTFs overall (n = 30, p = .019 at 19/30 correct), hence the hypothesis was rejected. Three other subjects yielded significantly more accurate localization (p < .05) with HRTFs that belonged to another subject than with their own. In terms of individual elevation levels, three subjects showed significant hits (n = 30, p = .031 at 9/10 correct) with their own HRTFs from the "above" position, and one subject with their HRTFs at "ear-level." All other localization performances failed to reach significance, indicating that subjects were mostly guessing.

Tables 1–3 show the correct identification rate of individual HRTFs in each elevation level, from the easiest to the most difficult. For example, the first row of Table 1 shows that across 100 total trials (10 sets per 10 subjects) using HRTF No. 2 (Subject 2's HRTF) on the ear-level, 61 were answered correctly. The hit counts for "ear-level" (Table 1) and "above" stimuli (Table 2) appeared to be higher than those for "below" (Table 3). A one-way analysis of variance (ANOVA) showed that these differences were strong but not significant (F(2, 27) = 2.99, p = .067).

Table 4 lists each subject's correct identification of their own HRTFs. The last three columns show each subject's number of correct identifications for each elevation level. The numerator in the "overall" column is the sum of the numerators in the last three columns. Shaded cells denote statistical significance, where the subject could reliably identify the correct elevation beyond guessing.

Table 1. Hit counts for 10 individual HRTFs at ear-level, ranked most to least

HRTF No.	Hit Count (n = 100)
2	61
8	59
6	55
4	54
3	45
7	44
1	43
5	43
10	43
9	39

Table 2. Hit counts for 10 individual HRTFs from above, ranked most to least

HRTF No.	Hit Count (n = 100)
1	91
4	73
7	72
8	71
2	67
3	60
9	50
5	48
10	26
6	18

Table 3. Hit counts for 10 individual HRTFs from below, ranked most to least.

HRTF No.	Hit Count (n = 100)
9	73
3	73
6	63
10	54
5	41
2	13
8	10
4	5
1	2
7	1

As mentioned previously, Subject 3 was the only person whose "overall" results reached significance, combined at 23/30 (p < .05) as well as for "above." The other subjects showed at most one significant result for their own HRTFs. Looking at data from each elevation level, three subjects showed

significance for "above," one for "ear-level," and none for "below."

Table 4. Hit count for each subject's own HRTF

Subject	Overall	Above	Below	Ear-level
1	18/30	10/10	0/10	8/10
2	17/30	8/10	1/10	8/10
3	23/30	9/10	8/10	6/10
4	15/30	9/10	0/10	6/10
5	16/30	6/10	4/10	6/10
6	12/30	0/10	6/10	6/10
7	13/30	4/10	0/10	9/10
8	12/30	6/10	2/10	4/10
9	17/30	4/10	5/10	8/10
10	13/30	0/10	7/10	6/10

Table 5 tallies the total number of significant results for each subject across all HRTFs. Therefore, the numbers on this table include the subjects' own HRTF results from Table 4. For example, Subject 2 showed significant performance ($p < .05$) on one HRTF overall ("Overall"), on six HRTFs for above, on one HRTF from ear-level, and no HRTFs from below, across all 10 sets of HRTFs. Summing each column yields seven overall, 22 above, 13 below, and four ear-level in total that reached significance.

Table 5. Each subject's hit counts across all 10 sets of HRTFs

Subject	Overall	Above	Below	Ear-level
1	2	3	0	0
2	1	6	0	1
3	3	2	6	0
4	0	3	0	0
5	0	2	3	1
6	0	1	0	0
7	0	0	1	1
8	0	2	2	0
9	0	1	0	1
10	1	2	1	0

4 Discussion

The goal of this research was to investigate if subjects' perceived elevation on the median plane would be more accurate using their own individualized HRTFs as opposed to using someone else's HRTFs.

Subjects here appeared to correctly localize "above" more often than "ear-level" or "lower" stimuli. This observation may have been the result of a headphone presentation artifact whereby a signal source does not appear on the horizontal plane but rather is perceptually elevated in the headphones in comparison to free-field listening [16]. Additionally, the quality of the stimuli may have affected the results because of the possible variability in subjects' head tilts when capturing the HRTF stimuli. Had this study captured HRTFs from loudspeakers placed at actual elevations rather than having subjects tilt their heads, results may have demonstrated a different outcome.

The hypothesis for this experiment was that subjects would demonstrate greater accuracy using their own individualized HRTFs on the medial plane than with other subject's HRTFs. This study was designed as a simple 3AFC test with 60˚ of total vertical separation and 30˚ in each direction from the horizontal reference of ear-level. Given that vertical localization blur for changes in elevation under binaural headphone conditions has been demonstrated to be from about 4˚ to 17˚ depending upon elevation from the horizontal [17], the differences captured here should have been clearly sufficient to show an effect of the individual HRTFs. Results here did not produce the predicted outcome and therefore, the hypothesis was rejected because results indicated that only one out of 10 subjects were statistically better with their own HRTFs.

References

[1] H. Møller, M.F. Sørensen, D. Hammershøi, and C.B. Jensen, "Head-related transfer functions of human subjects," *J. Audio Eng. Soc.*, vol. 43, pp. 300-321 (1995 May). [Online]. Available: https://scispace.com/pdf/head-related-transfer-functions-of-human-subjects-49tkbmpwce.pdf (accessed April 19, 2026).

[2] P. Majdak, R. Baumgartner, and B. Laback, "Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization," *Front. Psychol.*, vol. 5 (2014 Apr.). https://dx.doi.org/10.3389/fpsyg.2014.00319.

[3] D.Y.N. Zotkin, J. Hwang, R. Duraiswaini, and L.S. Davis, "HRTF personalization using anthropometric measurements," in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (IEEE Cat. No.03TH8684), pp. 157-160, New Paltz, NY (2003 Oct.). https://dx.doi.org/10.1109/ASPAA.2003.1285855.

[4] D. Poirier-Quinot, P. Gutierrez-Parera, and B.F.G. Katz, "One listener, sixteen HRTFs: a perceptual benchmark of individualization methods," in AES Intl. Conf. Headphone

Technology, Espoo, Finland (2025 Aug.). [Online]. Available: https://hal.science/hal-05377331/ (accessed April 19, 2026).

[5] V. Planinec, J. Reijniers, M. Horvat, H. Peremans, and K. Jambrošić, "The Accuracy of Dynamic Sound Source Localization and Recognition Ability of Individual Head-Related Transfer Functions in Binaural Audio Systems with Head Tracking," *Appl. Sci.*, vol. 13, no. 9, (2023 Apr.). https://dx.doi.org/10.3390/app13095254.

[6] A. Roginska, G.H. Wakefield, and T.S. Santoro, "User Selected HRTFs: Reduced Complexity and Improved Perception," in *Undersea Human System Integration Symp.*, Providence, RI (2010 Jul.). [Online]. Available: https://www.researchgate.net/profile/Agnieszka-Roginska/publication/267240183_User_Selected_HRTFs_Reduced_Complexity_and_Improved_Perception/links/5457c9f90cf2bccc491116aa/User-Selected-HRTFs-Reduced-Complexity-and-Improved-Perception.pdf (accessed April 19, 2026).

[7] J.S. Anderson, R. Miccini, S. Serafin, and S. Spagnol, "Evaluation of Individualized HRTFs in a 3D Shooter Game," in *IEEE Immersive and 3D Audio: from Architecture to Automotive* (I3DA), pp. 1-10, Bologna, Italy (2021 Sep.). https://dx.doi.org/10.1109/I3DA48870.2021.9610934.

[8] C.C. Berger, M. Gonzalez-Franco, A. Tajadura-Jiménez, D. Florencio, and Z. Zhang, "Generic HRTFs May be Good Enough in Virtual Reality. Improving Source Localization through Cross-Modal Plasticity," *Front. Neurosci.*. vol. 12, no. 21 (2018 Feb.). https://dx.doi.org/10.3389/fnins.2018.00021.

[9] P. Stitt, L. Picinali, and B.F.G. Katz, "Auditory Accommodation to Poorly Matched Non-Individualized Spectral Localization Cues Through Active Learning," *Sci. Rep.*, vol. 9, no. 1 (2019 Jan.). https://dx.doi.org/10.1038/s41598-018-37873-0.

[10] C. Mendonça, G. Campos, P. Dias, J. Vieira, J.P. Ferreira, and J.A. Santos, "On the improvement of Localization Accuracy with Non-Individualized HRTF-Based Sounds," in *J. Audio Eng. Soc.*, vol. 60, no. 10, pp. 821-830 (2012 Oct.). https://aes.org/publications/elibrary-page/?id=16555. (accessed April 19, 2026).

[11] S.H. Chon and S. Kim, "Auditory localization training using generalized head-related transfer functions in augmented reality," *Acoust, Sci. & Tech.*, vol. 39, no. 4, pp. 312-315 (2018 Feb.). https://dx.doi.org/10.1250/ast.39.312.

[12] S. Kim, S.H. Chon, H. Okumura, and S. Sakamoto, "Comparing training effects associated with two sets of HRTF data on auditory localization performance," presented at the *148th Conv. Audio Eng. Soc,*. Paper 10379, Vienna, Austria (virtual) (2020 May). https://aes.org/publications/elibrary-page/?id=20796 (accessed July 04, 2026).

[13] International Telecommunications Union, "Methods for the subjective assessment of small impairments in audio systems." Rec. ITU-R BS.1116-3 (02/2015) BS Series," 4 February 2015. [Online]. Available: https://www.itu.int/rec/R-REC-BS.1116-3-201502-I/en (accessed April 17, 2026).

[14] RANDOM.ORG, "True Random Number Service," N.D. [Online]. Available: https://www.random.org (accessed April 11, 2026).

[15] "Sign and Binomial Test," GraphPad.com, 2026. [Online]. Available: https://www.graphpad.com/quickcalcs/binomial1/ (accessed April 11, 2026).

[16] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," in *J. Acoust. Soc. Am.,* vol. 94, no. 1, pp. 111-123 (1993 Jul.). https://dx.doi.org/10.1121/1.407089.

[17] J. Blauert, "Spatial Hearing with One Sound Source, Localization and Localization Blur," Spatial hearing: the psychophysics of human sound localization, New York, MIT Press, 1997, pp. 44.