Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Psychological and Physiological Acoustics
Session 4aPPb: Binaural Hearing (Poster Session)

4aPPb3. Head movement during horizontal and median sound localization experiments in which head-rotation is allowed

Daisuke Morikawa*, Yuki Toyoda and Tatsuya Hirahara

***Corresponding author's address: Toyama Prefectural University, Kurokawa 5180, Imizu, 939-0398, Toyama, Japan, t074001@st.pu-toyama.ac.jp**

We measured subjects' head movements during horizontal and median sound localization experiments in which head-rotation was allowed in order to know how they move their heads to localize sound in a head rotation condition. The head movements in a head-rotation condition were measured while localizing 500-Hz low-pass noise (LPN), 12-kHz high-pass noise (HPN), and white noise (WN). With regard to horizontal plane, sound localization became easier with head-rotation than head-still condition. All subjects turned their heads toward the presented sounds, yet they did not necessarily turn their heads to face the sound. The amount of head rotation was small for WN compared to that for LPN and HPN. Sound localization also became easier with head-rotation than head-still condition for the median plane. All subjects swung their heads right and left centering on 0°, no matter what the stimulus elevation angle was.

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4800148]
Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

We move our heads when we listen to sounds, because we subconsciously know that sound localization is more accurate when we move our head. The effect of head movement on sound localization has been known since its report by Wallach [1−11]. However, it is not clear how we move our heads when we localize sound.

Thurlow and Runge examined head movement during sound localization using a small number of sound sources: center, right front, lower right front, lower right, and upper left loudspeakers. By analyzing head movement during sound-localization experiments [2], they detected many horizontal head rotations when locating the sound direction. Based on Thurlow and Runge's result, Blauert concluded that people face the sound source because the minimum audible angle is smallest [3]. Iwaya *et al.* examined the effect of head movement on the front–back error in sound localization using 12 loudspeakers, which were placed along a horizontal circle at 30° intervals. They reported that listeners tended to move their heads more dynamically when the sound was presented away from the front (0°), trying to capture it directly in front of them [7]. By contrast, Nojima *et al.* examined head movement during sound localization using 5 or 6 loudspeakers which were placed along horizontal right half-circles at 45° intervals [12] or on a horizontal circle at 0°, 30°, 135°, 180°, 215°, and 315° [13]. They reported that listeners moved their heads toward the sound image in the lateral direction and could localize the sound even when they were not facing sound source. Toshima *et al.* examined the effectiveness of *TeleHead*, which is a steerable dummy head that tracks human head movement in real time. They carried out sound-localization experiments using *TeleHead* with 12 loudspeakers, which were placed on a horizontal circle at 30° intervals. They reported that listeners did not necessarily turn their heads to capture the sound in front of them [8].

In the experiments mentioned above, Thurlow and Runge, and Nojima *et al.* only clarified the rough tendency of head movement during sound localization, as they measured head movement during sound localization using a small number of loudspeakers. By contrast, Iwaya *et al.* and Toshima *et al.* measured head movement during sound localization with a large number of loudspeakers. However, the purpose of their experiments was not to clarify how we move our heads, but to clarify the effect of head movement on the front–back error in sound localization or the effectiveness of *TeleHead*. Therefore, they did not obtain a detailed analysis of head movement during sound localization.

In our study, to clarify how we move our heads when we localize sound, we measured the listener's head movements during horizontal- and median-plane sound-localization experiments with real sound sources under head-rotation conditions. We quantified the listeners' head movement when they localized sound when head movement was allowed for each stimulus. In addition, we discuss the head movement strategy for sound localization.

METHODS

Three-dimensional positions and Euler angles of the listener's head were measured using a motion sensor (Flock of Birds, Ascension Technology) placed on top of the listener's head. The motion sensor has static positional accuracy of 1.8 mm RMS and static angular accuracy of 0.5° RMS. Figure 1 depicts a block diagram of the experimental system. Two PCs were used: one for controlling the sound-localization experiment and the other for simultaneously recording the stimuli sound and head movement data to confirm their synchronization. The sampling frequency of the sound was 48 kHz and that of the motion sensor was 128 Hz.

The loudspeakers were placed around a chair in a horizontal circle as well as along the median plane of 1 m radius in the upper hemisphere at 30° intervals. The horizontal loudspeakers were mounted at a height of 1.1 m.

Listeners sat on a chair placed at the center of the speaker array without using any head fixing device. Stimuli were presented from one of the loudspeakers. In the horizontal-sound-localization experiment, each session consisted of 60 trials, and the stimuli were presented in random order from 12 loudspeakers. In the median-sound-localization experiment, each session consisted of 35 trials, and the stimuli were presented in random order from 7 loudspeakers. Each experiment consisted of four sessions. The listener's head movements were measured during the horizontal- and median-sound-localization experiments under head-rotation conditions.

White noise (WN), *i.e.*, Gaussian-distributed random noise, 500-Hz low-pass noise (LPN), and 12-kHz high-pass noise (HPN) were used as stimuli for sound localization. WN can be easily localized while LPN and HPN are difficult to localize because many front–back errors occur [14]. The durations of both the stimulus and the inter-stimulus intervals were 3 s. A 30-ms linear taper window was applied at the beginning and end of the stimuli. The sound pressure level of WN was 70 dB, whereas those of HPN and LPN were lower in accordance with the amount of filtering.

The listeners were instructed to close their eyes and were encouraged to freely rotate their heads while each stimulus was presented. Six and four males in their 20s participated in the horizontal and median experiment respectively.

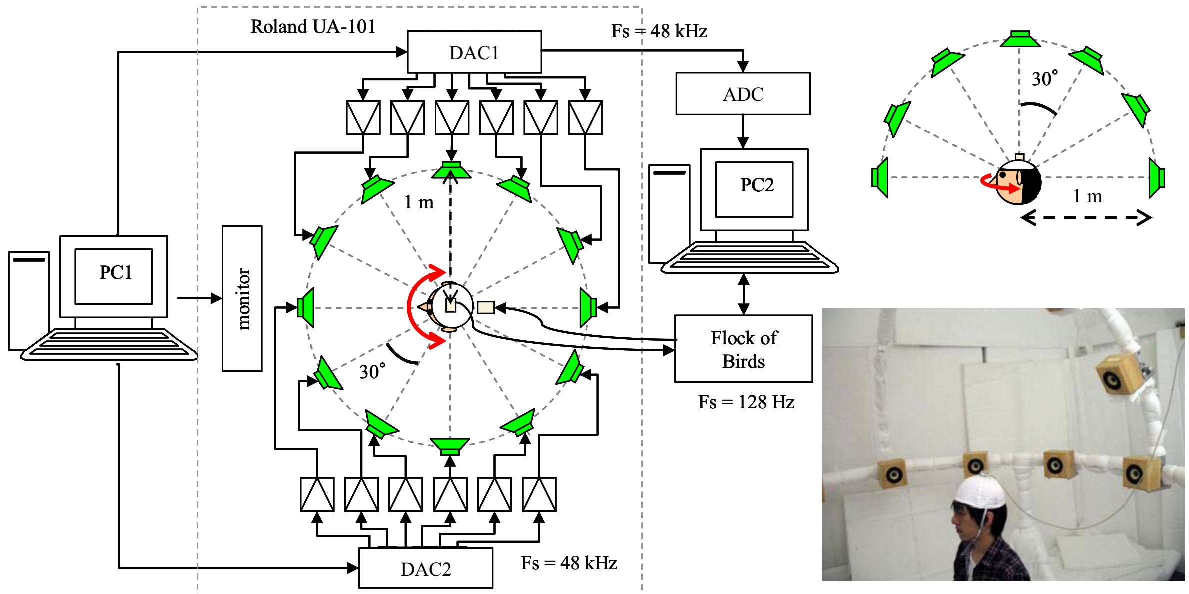

FIGURE 1. Block diagram of experimental system and photograph of the experiment room. PC1 was used for stimuli presentation for the sound-localization experiment. PC2 was used to simultaneously record the sound stimuli and the listener's head movement data.

RESULTS

Horizontal-plane sound localization

Figure 2 shows the typical head-yaw angle trajectory of the listener's head for LPN. The gray line represents the stimulus waveform and the red line represents the head-yaw angle trajectory. The blue and green circles represent the presented and perceived azimuthal angles of the stimulus, respectively. All the stimuli were localized correctly.

In most cases, the listener turned his head toward the presented stimulus direction, and then swung it back and forth. For the first stimulus, presented at the front right (60°), he turned his head right and swung it to right. For the third stimulus, presented at the right lateral direction (90°), he turned his head right once. For the 5th stimulus, presented at the rear left (–120°), he turned his head left and swung it left and right. For the 7th stimulus, presented at the front left (–30°), he turned his head left and swung it to left. For the 8th stimulus, presented directly in the front (0°), he turned his head right once. Usually, the maximum yaw angles at which he turned his head were smaller than the stimuli angles.

These head movements were classified into two patterns: single-side head swing and both-side head swing. The single-side head swing pattern is defined as the listener turning and swinging his head one or two times to either the right or left. The both-side head swing pattern is when the listener swings his head both to the right and left. Stimuli number 1, 3, 4, 6, 7, 8, 9, and 10 resulted in the single-side head swing pattern, and stimuli number 2 and 5 resulted in the both-side head swing pattern.

The $\mathrm{MEAN}_{i,j}(\theta)$, $\mathrm{MAX}_{i,j}(\theta)$, and $\mathrm{MIN}_{i,j}(\theta)$ define the mean, maximum, and minimum head-yaw angle during one stimulus that was recorded for each stimulus and angle. θ is the stimuli angle, i is the stimuli number of each angle, and j is the listener number. n represents the number of single-side or both-side head swings for each angle. The average values of the $\mathrm{MEAN}_{i,j}(\theta)$, $\mathrm{MAX}_{i,j}(\theta)$, and $\mathrm{MIN}_{i,j}(\theta)$ for each listener are given by Eqs. (1)–(3), respectively.

$$\overline{\mathrm{MEAN}}_j(\theta) = \frac{1}{\mathrm{n}} \sum_{i=1}^{\mathrm{n}} \mathrm{MEAN}_{i,j}(\theta) \tag{1}$$

$$\overline{\mathrm{MAX}}_j(\theta) = \frac{1}{\mathrm{n}}\sum_{i=1}^{\mathrm{n}} \mathrm{MAX}_{i,j}(\theta) \tag{2}$$

$$\overline{\mathrm{MIN}}_j(\theta) = \frac{1}{\mathrm{n}}\sum_{i=1}^{\mathrm{n}} \mathrm{MIN}_{i,j}(\theta) \tag{3}$$

Figure 3 shows the $\overline{\mathrm{MEAN}}_j(\theta)$, $\overline{\mathrm{MAX}}_j(\theta)$, and $\overline{\mathrm{MIN}}_j(\theta)$ for the single-side head swing (upper panel) and both-side head swing (lower panel) patterns for each angle of the presented stimuli, where N is the number of single-sided or both-sided head swings. The black, red, and blue circles are the $\overline{\mathrm{MEAN}}_j(\theta)$, $\overline{\mathrm{MAX}}_j(\theta)$, and $\overline{\mathrm{MIN}}_j(\theta)$, respectively. There are some points where data is not drawn for the single-side head swing, indicating the listener always swung his head to both sides. In the single-side head swing pattern, $\overline{\mathrm{MAX}}_j(\theta)$ or $\overline{\mathrm{MIN}}_j(\theta)$ was not necessarily 0° because the listener did not return his head to the initial position punctually when each stimulus presentation began. $\overline{\mathrm{MAX}}_j(\theta)$ was about 60° and $\overline{\mathrm{MIN}}_j(\theta)$ was about –60°.

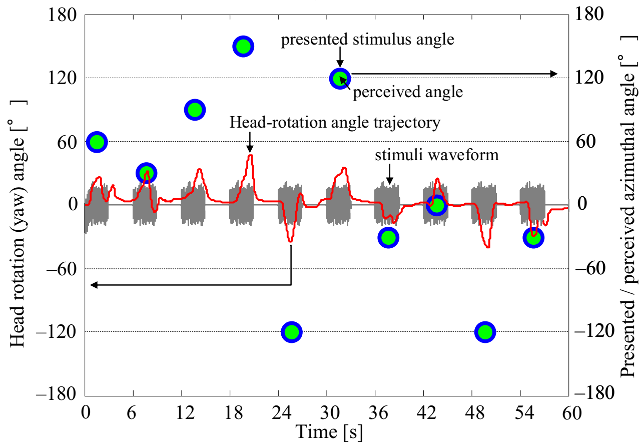

FIGURE 2. Example of typical head-yaw angle trajectory of a listener for LPN. Head movement can be classified into two patterns: single-side head swing and both-side head swing.

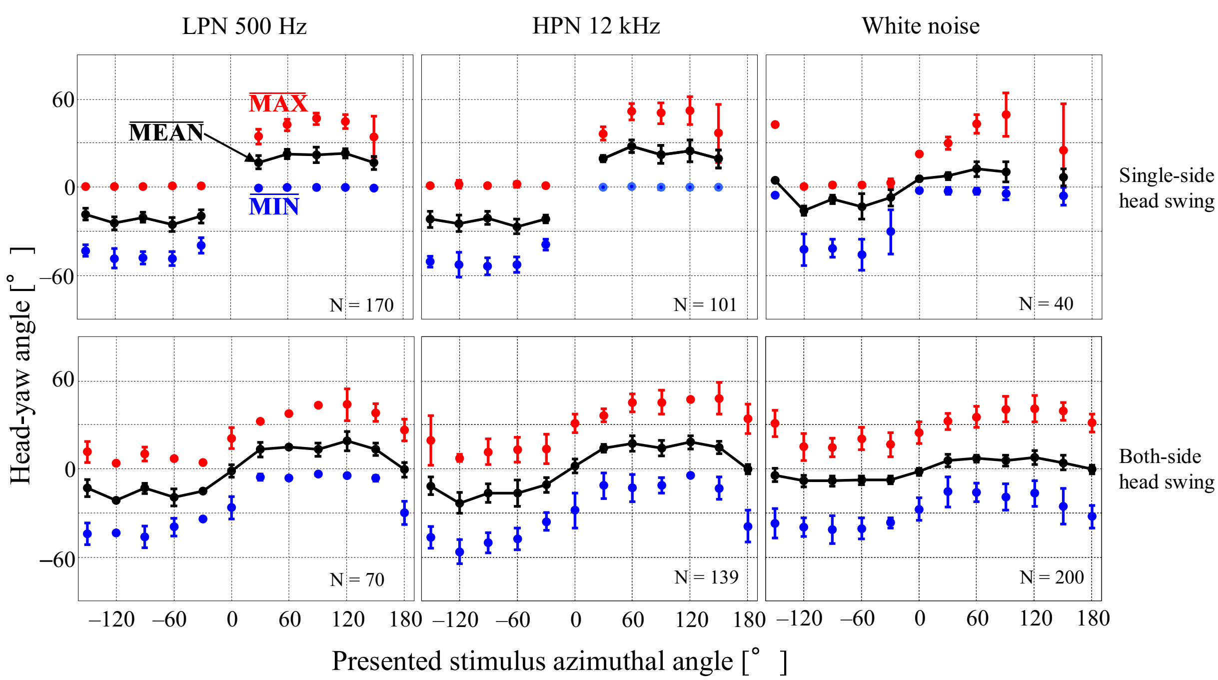

FIGURE 3. $\overline{\mathrm{MEAN}}_j(\theta)$, $\overline{\mathrm{MAX}}_j(\theta)$, and $\overline{\mathrm{MIN}}_j(\theta)$ of the listener for LPN, HPN, and WN. Upper panels show those of the single-side head swing pattern, and lower panels show those of the both-side head swing pattern.

$\overline{\text{AMAX}}_j(\theta)$ and $\overline{\text{AMIN}}_j(\theta)$ are the average values of the absolute value of $\text{MAX}_{i,j}(\theta)$, and $\text{MIN}_{i,j}(\theta)$ for each of listener, respectively. The means of $\overline{\text{AMAX}}_j(\theta)$ and $\overline{\text{AMIN}}_j(\theta)$ for all the listeners for 0°, ±30° and ±60°of the presented stimuli are shown in Figure 4(a) and are given by Eqs. (4)–(5). n represents the number of listeners who moved their heads toward the sound image on the lateral sides.

$$\overline{\text{AMAX}}_j(\theta) = \left| \frac{1}{\text{n}} \sum_{i=1}^{\text{n}} \text{MAX}_{i,j}(\theta) \right| \qquad (\theta = 0°,\ 30°,\ 60°) \tag{4}$$

$$\overline{\text{AMIN}}_j(\theta) = \left| \frac{1}{\text{n}} \sum_{i=1}^{\text{n}} \text{MIN}_{i,j}(\theta) \right| \qquad (\theta = 0°,\ \text{-}30°,\ \text{-}60°) \tag{5}$$

The listeners did not turn their heads to face the stimuli when the stimuli were presented over ±60°.

$\overline{\text{MMEAN}}(\theta)$ (Figure 4b) is the average value of $\text{MEAN}_{i,j}(\theta)$ and is given by Eq. (6), where n is the number of stimulus for each angle and s represents the number of listeners.

$$\overline{\text{MMEAN}}(\theta) = \frac{1}{\text{s}} \sum_{j=1}^{\text{s}} \left(\frac{1}{\text{n}} \sum_{i=1}^{\text{n}} \text{MEAN}_{i,j}(\theta) \right) \tag{6}$$

$\overline{\text{MMEAN}}(\theta)$ for each stimulus was the smallest for WN than for LPN and HPN. The detailed head movements differed among the listeners. For example, one listener always swung his head a certain amount even though the stimulus azimuthal angle was that of the single-side head swing pattern. Another listener preferred to use the single-side head swing pattern than the both-side head swing pattern. However, all six listeners showed similar head-movement strategies. They did not turn their heads to capture the sound stimuli in front of them, $\overline{\text{MAX}}_j(\theta)$ and $\overline{\text{MIN}}_j(\theta)$ were about ±60°, and the amount of head rotation was small for WN compared to that for LPN and HPN.

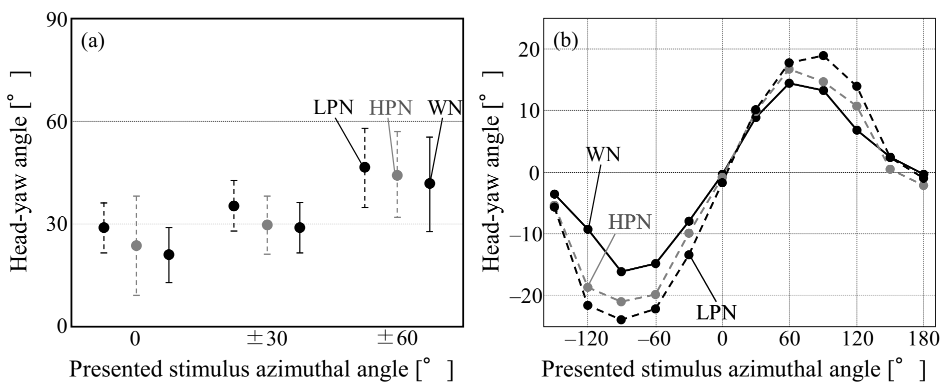

FIGURE 4. (a) Mean of $\overline{\text{AMAX}}_j(\theta)$ and $\overline{\text{AMIN}}_j(\theta)$ of all listeners for LPN, HPN, and WN at 0°, ±30° and ±60°. (b) $\overline{\text{MMEAN}}(\theta)$ for LPN, HPN, and WN for all azimuthal angles.

Median-plane sound localization

Figure 5 shows a typical head-yaw angle trajectory of the listener's head for LPN. The gray line represents the stimulus waveform, and the red line represents the head-yaw angle trajectory. The blue and green circles represent the presented and perceived elevation angles of the stimulus, respectively. All the stimuli were localized correctly except for stimuli number 2 and 7.

In all cases, the listener swung his head right and left centering on 0°, irrespective of the stimulus elevation angle. Thus, in median sound localization, there was almost no single-side head swing pattern but mostly the both-side head swing pattern.

Figure 6 shows the $\overline{\text{MEAN}}_j(\theta)$, $\overline{\text{MAX}}_j(\theta)$, and $\overline{\text{MIN}}_j(\theta)$ in the both-side head swing patterns for each angle of the presented stimuli. N represents the number of the both-side head swings. The black, red, and blue circles are the $\overline{\text{MEAN}}_j(\theta)$, $\overline{\text{MAX}}_j(\theta)$, and $\overline{\text{MIN}}_j(\theta)$, respectively.

The detailed head movement differed among the listeners. For example, among LPN, HPN, and WN, $\overline{\text{MAX}}_j(\theta)$ for each stimulus was the smallest, and $\overline{\text{MIN}}_j(\theta)$ for each stimulus was the largest for WN for this listener. Other listeners showed the same head movement regardless of the stimulus. However, all four listeners showed similar head-movement strategies. They mostly swung their heads right and left centering on 0°, irrespective of the stimulus elevation angle and stimulus.

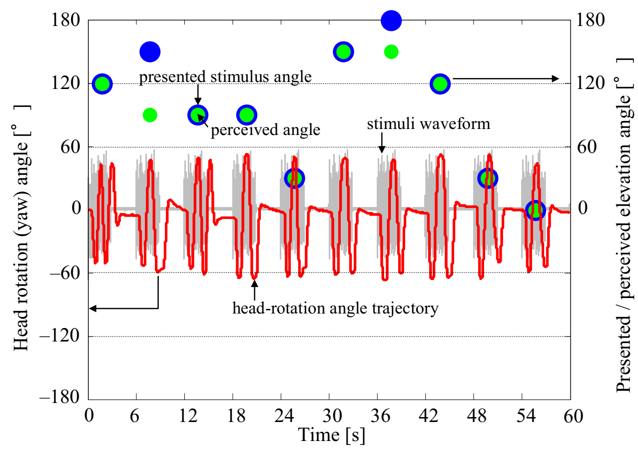

FIGURE 5. Example of typical head-yaw angle trajectory of a listener for LPN.

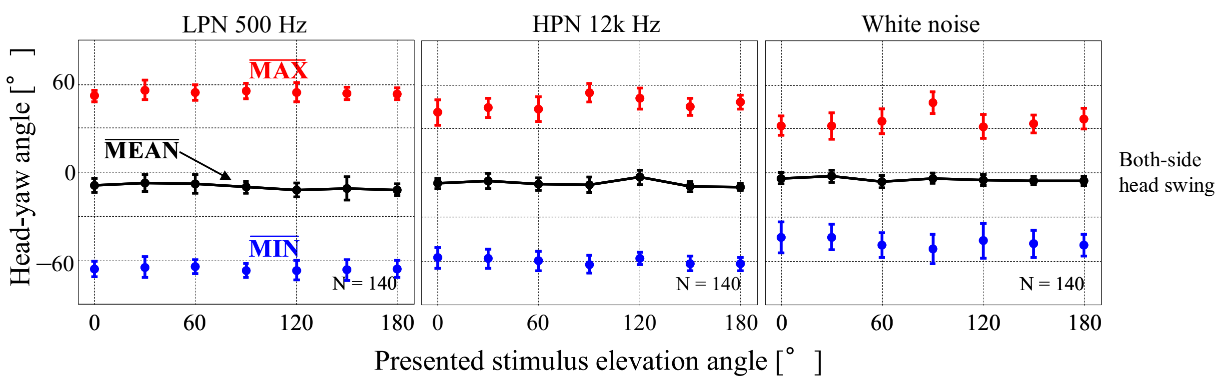

FIGURE 6. $\overline{\mathrm{MEAN}}_j(\theta)$, $\overline{\mathrm{MAX}}_j(\theta)$, and $\overline{\mathrm{MIN}}_j(\theta)$ of the listener for LPN, HPN, and WN for the both-side head swing pattern.

DISUCUSSION

In the horizontal-plane sound-localization experiments, WN was almost perfectly localized by all the listeners because it contains rich acoustic cues for localization. This was observed even in experiment under the head-still condition [14]. The dynamic acoustic cues provided by large head movement are probably not necessary in the WN case where small head movements are sufficient to confirm the sound direction. In contrast, the correct localization performances of LPN and HPN were less than 75% under the head-still condition [14]. The performances improved more than 20% when head movement was allowed because the front–back confusion disappeared due to the head rotation. It is clear that the dynamic cues from head movement play an important role in localizing LPN and HPN, which contain either an interaural time difference (ITD) or an interaural level difference (ILD). Thus, a larger amount of head movement is necessary for localizing LPN and HPN than for localizing WN.

In median-plane sound-localization, the performance of WN was about 70% under the head-still condition [14], because it contained only spectral cues (SC). This is similar to the performance of LPN and HPN under the head-still condition of horizontal sound localization. In addition, the performances of LPN and HPN were less than 40% [14]. The performances of WN, LPN, and HPN improved more than 20% when head movement was allowed. For median-plane sound localization, a large amount of head movement is necessary for WN as well as for LPN and HPN, unlike the case of horizontal-plane sound localization.

Our experiments demonstrated that the listeners did not turn their heads to capture the stimulus in front of them. The yaw angles at which they turned their heads were always smaller than the stimulus azimuthal angles. Even though the maximum yaw angle at which humans can turn their heads is about 70° [10], the listeners did not turn their heads to this extent, indicating that they can obtain sufficient information for sound localization with smaller head yaw angles. Namely, they did not face the sound stimuli, suggesting that they do not move their heads in order to use the smallest minimum audible angle direction. They all swung their heads while localizing the sound. The variation in ITD, ILD, and SC, *i.e.,* ΔITD, ΔILD, and ΔSC or the time derivative of ITD, ILD, and SC must be used as dynamic cues to localize sound. In fact, the sound image of the stimuli swings around the head when the head is rotated, resulting in a dramatic reduction in the ambiguity of the sound image location.

CONCLUSIONS

We carried out experiments to measure listeners' head movements during horizontal- and median-plane sound localization experiments under head-rotation conditions. In the horizontal-plane sound localization, listeners turned their heads toward the stimuli when each stimulus was presented. The head movement patterns were either single-side head swing or both-side head swing. The maximum head-yaw angle, however, was less than ±60° even though the stimulus direction. The listeners did not turn their heads to capture the stimulus in front of them when the stimulus was presented beyond ±60°. The amount of head rotation was small for WN compared to that for LPN and HPN, suggesting that if the acoustic cues included in the stimulus are less, the head movements necessary to localize the sound are more. In the median-plane sound localization, the listeners swung their heads right and left, centering on 0°, irrespective of the stimulus elevation angle and stimulus. The head movement pattern was mostly both-side head swing.

ACKNOWLEDGMENTS

Part of this work was supported by KAKENHI (22300061).

REFERENCES

1. H. Wallach, "On sound localization," J. Acoust. Soc. Am., **10**(1), 270-274 (1939).
2. W. Thurlow and S. Runge, "Head Movements During Sound Localization," J. Acoust. Soc. Am., **42**(2), 489-493 (1967).
3. J. Blauert, "*Spatial Hearing - The Psychophysics of Human Sound Localization* -," Revised Edition, The MIT Press, Cambridge, 178-191 (1997).
4. S. Perrett and W. Noble, "The contribution of head motion cues to localization of low-pass noise," Perception & Psychophysics, **59**(7), 1018-1026 (1997).
5. F. Wightman and D. Kistler, "Resolution of front-back ambiguity in spatial hearing by listener and source movement," J. Acoust. Soc. Am., **105**(5), 2841-2853 (1999).
6. D. Begault, E. Wenzel, A. Lee, and M. Anderson, "Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source," J. Audio Eng. Soc., **49**(10), 904-916 (2001).
7. Y. Iwaya, Y. Suzuki, and D. Kimura, "Effects of head movement on front-back error in sound localization," Acoust. Sci. & Tech. **24**(5), 322-324 (2003).
8. I. Toshima, H. Uematsu, and S. Aoki, T. Hirahara, "A steerable dummy head that tracks human head movement in real time: TeleHead," J. Acoust. Soc. Jpn, **61**(4), 197-207 (2005). (in Japanese)
9. T. Hirahara, H. Sagara, and M. Otani "Sound localization with scaled dummy-heads on a TeleHead," Proc. 19th International Congress on Acoustics, PPA-05-013, 1-4 (2007).
10. I. Toshima, S. Aoki, and T. Hirahara, "Sound Localization Using an Acoustical Telepresence Robot: TeleHead II," Presence, **17**(4), 365-375 (2008).
11. T. Hirahara, Y. Sawada, and D. Morikawa, "Impact of dynamic binaural signals on three dimensional sound reproduction," Proc. INTERNOISE (2011).
12. R. Nojima, M. Morimoto, and H. Sato, "Head movements during sound localization," Proc. Autumn Meet. Acoust. Soc. Jpn, 513-516 (2009). (in Japanese)
13. R. Nojima, M. Morimoto, and H. Sato, "Head movements during sound localization: Part2," Proc. Autumn Meet. Acoust. Soc. Jpn, 522-528 (2011). (in Japanese)
14. D. Morikawa and T. Hirahara, "Effect of head rotation on horizontal and median sound localization of band-limited noise," Acoust. Sci. & Tech. **34**(1), 56-58 (2013).