AES

Audio Engineering Society

Convention Paper

Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands



# The Accuracy of Localizing Virtual Sound Sources: Effects of Pointing Method and Visual Environment

PIOTR MAJDAK[1], BERNHARD LABACK, MATTHEW GOUPELL, and MICHAEL MIHOCIC

[1] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
piotr@majdak.com
bernhard.laback@oeaw.ac.at
matt.goupell@gmail.com
michael.mihocic@oeaw.ac.at

## ABSTRACT

The ability to localize sound sources in 3D-space was tested in humans. The subjects listened to noises filtered with subject-specific head-related transfer functions. In the experiment using naïve subjects, the conditions included the type of visual environment (darkness or structured virtual world) presented via head mounted display and pointing method (head and manual pointing). The results show that the errors in the horizontal dimension were smaller when head pointing was used. Manual pointing showed smaller errors in the vertical dimension. Generally, the effect of pointing method was significant but small. The presence of structured virtual visual environment significantly improved the localization accuracy in all conditions. This supports the benefit of using a visual virtual environment in acoustic tasks like sound localization.

## 1. INTRODUCTION

Testing the localization of sound sources requires accurate methods for the presentation of stimuli and the acquisition of subjects' responses. The acquisition method may affect the accuracy of the actual perceived sound direction and thus, change the efficiency of the localization task. Many methods have been used in localization tasks: verbal responses [1], rotating dial or drawing [2], pointing with the nose [3]-[5], pointing with the chest or finger [2], pointing with extensions of the body, like a stick, cane or gun [2, 6], pointing with a laser pointer [7, 8], or using sophisticated computer interfaces, e.g. [9]. In [2], Haber et al. compared the nine different methods listed above and showed the highest accuracy for pointing methods. However, they compared the methods in blind subjects using pure-tones and tested in the horizontal plane only. Additionally, the choice of the particular pointing method (head, finger, gun, or stick) remains unclear. For sound localization in 3D, the manual pointer promises a better ability to point to elevated directions, which can be difficult to access by turning the head and pointing with the nose. A direct comparison between head and manual-pointing methods has been never investigated for sound localization including the vertical planes. Thus, in this study we investigated the effect of head and manual hand pointing on sound localization ability in 3D space.

Many studies investigated the spatial coordination of the visual and auditory senses, e.g. [10, 11]. The general finding is that the addition of visual information improves the accuracy of sound localization when both channels provide congruent information. However, there are differences in the processing of information. In principle, the auditory system encodes a sound location primarily within a craniocentric frame of reference. This is defined for each subject individually by the position of the ears and acoustic properties of the head, torso, and pinna. In contrast, the visual system encodes positions within an oculocentric frame of reference, which, additionally, changes with eye movements [12]. The difference in the frames of reference can be a potential source of confusion while localizing sounds with visual feedback and is the reason for the investigations of visual effects and aftereffects on the sound localization (for review see [7]). Generally, when a subject has to point at a required position in the presence of visual feedback, visuomotor recalibration between the proprioceptive and the visual information can improve the pointing accuracy. For example, in [13], Redon and Hay found that using a visually structured background increases pointing accuracy to visual targets. Hence, a visual environment in sound localization tasks may help subjects to respond more accurately.

However, it is rather difficult to design a real visual environment (VE), which does not change from study to study and does not rely on the test facilities. Fortunately, virtual rendering techniques allow to design the required environment in software, which can be used in many experiments without any changes. However, using a virtual VE in an auditory localization task introduces many new issues [14]. Because of many direct and indirect health and safety effects on subjects [15], the improvements from the real VE do not directly imply the same advantages of the virtual VE. In [16], Zahorik et al. tested sound localization using a virtual VE, which was presented via head-mounted display. They showed that the accuracy in localizing virtual stimuli can be improved by training the subjects with visual feedback. However, the purpose of the study was different and they did not test the effect of the virtual VE. Additionally, they used generic head-related transfer functions (HRTF, [17]) to generate virtual sound sources. In this study we use individual HRTFs and directly compare sound localization accuracy between the condition with a virtual VE and the condition of testing in darkness.

## 2. METHODS

### 2.1. Subjects and Apparatus

Ten listeners participated in the experiments (six male and four female); the age range was 23 to 36. All subjects had normal hearing, normal or corrected-to-normal vision, and showed no evidence of any auditory or visual deficits. The listeners were naïve with respect to localization tests. All the listeners were right-handed.

The virtual acoustic stimuli were presented via headphones (HD 580, Sennheiser) in a semi-anechoic room. The A-weighted sound pressure level (SPL) of the background noise in this room was 18 dB re 20 µPa on a typical testing day. A digital audio interface (ADI-8, RME) with a sampling rate of 48 kHz and a resolution of 24 bit was used.

The visual environment was presented via head mounted display (HMD; 3-Scope, Trivisio). It was mounted on subject's head and provided two screens with a field of view of 32° x 24° (horizontal x vertical dimensions). The screen surrounding outside of the field of view was black. The HMD did not enclose the complete field of view and thus, it was necessary to darken the room for the tests in darkness (see later). In that experimental conditions the room provided no visual information. The visual environment was presented binocular; however the same picture was used for both eyes. The subjects adjusted the interpuppilary distance and the eye relief to achieve a focused and unvignetted image. The resolution of each screen was 800 x 600 (horizontal x vertical).

The position and orientation of subject's head were captured via electromagnetic tracker (Flock of Birds, Ascension) in real-time. One tracking sensor was mounted on the top of subject's head. In conditions were the manual hand pointer was used, the position and orientation of the pointer were captured with a second tracking sensor mounted on the pointer. The tracking device was capable to capture all 6-degrees of freedom at a rate of 100 measurements per second for each sensors. The tracking accuracy was 1.7 mm for the positions and 0.5° for the orientation.

Two personal computers were used to control the experiments in a client-server architecture. The machines communicated via Ethernet using TCP/IP. The client machine acquired the tracker data, created and presented the acoustic stimuli, and controlled the experimental procedure, while the server machine handled the 3D graphic rendering upon client's requests. This architecture allowed a balanced distribution of computational resources.

### 2.2. HRTF measurement

The HRTFs were measured for each subject individually. Twenty-two loudspeakers (custom-made boxes with VIFA 10 BGS as drivers; the variation in the frequency response was ±4 dB in the range from 200 to 16000 Hz) were mounted at fixed elevations from -30° to 80°. They were driven by amplifiers adapted from Edirol MA-5D active loudspeaker systems. The loudspeakers and the arc were covered with acoustic damping material to reduce the reflexion from the adjacent parts. The total harmonic distortion of the loudspeaker-amplifier systems was on average 0.19 % (at 63 dB SPL and 1 kHz). The subjects were seated in the center of the arc and had microphones (Sennheiser KE-4-211-2) placed in his/her ears, which were connected via pre-amplifiers (RDL FP-MP1) to the digital audio interface. An exponential sweep with a duration of 1728.8 ms had a frequency beginning at 50 Hz and ending at 20 kHz was used to measure each HRTF. The multiple exponential sweep method (MESM) was applied to measure HRTFs in an interleaved and overlapped order for one azimuth and all elevations at once [18]. Then, the subject was rotated by 2.5° to measure HRTFs for the next azimuth. In total, 1550 HRTFs were measured for one listener, where the positions were distributed with a constant spherical angle on the sphere. During the procedure the head position and orientation were monitored with the tracker. The subject's position was validated after the measurement of one azimuth. The valid ranges were set to 2.5 cm for the position, 2.5° for the azimuth, and 5° for the elevation and roll. In the out-of-range case, the measurement for that particular position was repeated immediately. On average, three measurements had to be repeated per subject. The measurement procedure lasted for approximately 20 minutes. The HRTFs were calculated from the recordings according to the system identification procedure for measurements with the MESM [18].

The effect of the equipment was removed by equalizing the HRTFs with the transfer functions of the equipment. A reference measurement was performed, in which in-ear microphones were placed in the center of the arc and the system identification procedure was performed for all loudspeakers. The equipment transfer functions were derived from the reference measurement.

Then, the directional transfer functions (DTF) were calculated according to the procedure of [5]. The magnitude of the common transfer function (CTF) was calculated by averaging the log-amplitude spectra of all HRTFs. The phase of the CTF was the minimum phase corresponding to its amplitude spectrum. The DTFs were the result of filtering the HRTFs with the inverse complex CTF. Finally, all DTFs were windowed (asymmetric Tukey window) to a 5.33 ms duration.

### 2.3. Stimuli

The stimuli were Gaussian white noises, which were filtered with the subject-specific DTFs and are referred to as acoustic targets. The duration was 500 ms. The targets were faded in and out using Tukey window with a taper corresponding to a fade of 10 ms.

The acoustic targets were uniformly distributed on the surface of a virtual sphere with the listener in the center of this sphere. All positions in the horizontal dimension were used. In the vertical dimension the position range was from -30° to +80°, relative to the eye-level of the listener.

The level of the stimuli was 50 dB re hearing level. The hearing level was estimated in a manual up-and-down procedure using a target positioned at azimuth and ele-

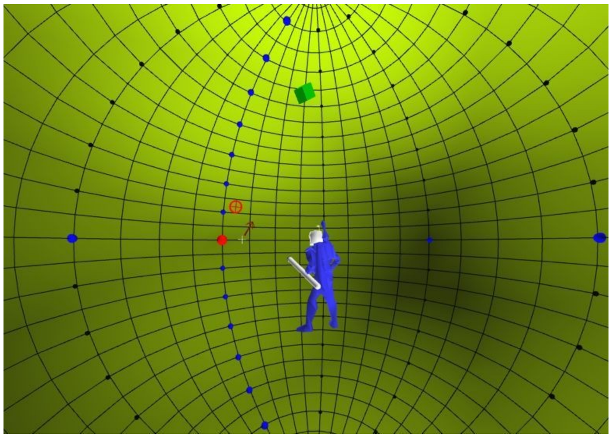

*Figure 1: Virtual VE used in this study.*

vation of 0°. In the experiment, the level for each presentation was randomly roved within the range of ±5 dB to reduce the possibility to localize spatial positions based on levels.

### 2.4. Procedure

Prior to the tests, subjects performed a procedural training, where they played a simplified version of a “first person shooter”. The subjects were immersed in a virtual environment (VE), finding themselves inside of a yellow, 5-m large sphere (see Fig. 1). Grid lines every 5° and 11.25° (horizontal and vertical, respectively) were used to improve the orientation in the sphere. The eye-level and the median were marked with small blue balls. The reference position (azimuth and elevation of 0°) was marked with a larger red ball. The lighting in the sphere was homogeneous; however, it was more light in the front and back than at the side. The subjects could not see their avatar; however, they could see the visualization of the hand pointer as shown in Fig. 2. The subjects were allowed to turn in the VE but could not to move in the VE.

At the beginning of each trial, the subjects were asked to find and look at a reference position. By clicking, a visual target in form of a red rotating cube was presented on the surface of the sphere at a random position. The subjects had to find and click at the target within four seconds to count the trial as “hit”. In this case, the subject heard a short sound, which indicated “hit” and the next trial began. In case of not being able to find the target within the alloted time, the next trial began. To avoid any auditory training effects, acoustic information about the target position was not provided during the procedural training. Blocks of 100 targets were tested for each head and manual pointing. The block order was balanced to avoid training differences in the pointing methods. The procedural training was continued until 95% of the targets were hit with an average accuracy better than 2°, measured in three succeeding blocks. On average, the subjects required 590 and 710 targets to reach the limits for the head and manual-pointing methods, respectively.

In the main experiment, the effect of the visual environment was studied under two different conditions. In the condition “HMD”, the subjects were immersed in the VE, as used in the procedural training. In the condition “Dark”, the subjects were tested in darkness. In this condition, the VE was turned black and only the reference position was shown for calibration purposes. Also, the room lights were switched off and thus, the subjects had no visual information regarding their orientation.

The effect of the response method was investigated by testing two methods: head and manual pointing. In the head-pointing method, the subjects were asked to turn their body and head to the perceived direction of the target. The orientation of the head was recorded as the perceived target position. In the manual-pointing method, subjects had a pointer in their right hand and were asked to point to the direction of perceived target. The projection of the pointer direction on the sphere's surface, calculated upon the position and orientation of subjects head and the pointer, was recorded as the perceived target position. The pointer was visualized whenever it was in subject's field of view.

At the begin of each trial, the subject was asked to look at the reference position. In the condition “Dark” this was the only moment where the HMD provided visual information. After confirmation, the acoustic target was presented. During the presentation the subject was not allowed to move. After the acoustic presentation, the


*Figure 2: VE from subject's point of view*

subject was asked to point to the perceived position with the head or the manual pointer, depending on the test condition.

The tests were performed in blocks, each block consisted of 100 acoustic targets with random positions and lasted for approximately 30 minutes. All combinations of the VE and pointing method were tested for all the subjects; however, the condition did not change within one block. The block order was random and different for each subject. In total, four blocks were tested per condition and subject.

### 2.5. Analysis

The horizontal-polar coordinate system [19], which uses the lateral and polar angles, was used in the analysis. The lateral angle ranged from -90° to 90°. The polar angle of the targets ranged from -30° (front, below eye-level) to 210° (rear, below eye-level). However, responses outside this polar range were allowed.

The measures of localization ability correspond to [5] who used lateral bias, RMS lateral errror, RMS local polar error, and quadrant errors as measures of localization ability. The lateral bias is the signed mean of the lateral difference between target and response. The RMS lateral error is referred to as lateral spread and is the root-mean-square (RMS) error of the lateral difference between target and response. The quadrant error is the percentage of front-back confusions. A confusion occurs when the difference between the response and target polar angle is higher than 90°. Quadrant errors are calculated for targets within the lateral range of ±30° only. The RMS polar error is referred to as raw polar spread and is the RMS error of the polar difference between target and response for targets within the lateral range of ±30°. Because many quadrant errors inflate the raw polar spread, correction was used to more accurately estimate the spread in the polar dimension. In the corrected polar spread, the quadrant errors were corrected before calculating the RMS error. This was done by flipping the response angle to the correct plane if necessary. Thus, the corrected polar spread corresponds to the precision of the vertical-plane localization ability despite the amount of the front-back confusions.

The spread and bias were compared between conditons using a multiway repeated-measures analysis of variance (RM ANOVA). The comparison of the quadrant errors was performed by fitting frequencies of occurrence to log-linear models and estimating the confidence intervals of log-ratios [20, 21].

## 3. RESULTS

Fig. 3 and Fig. 4 show the results for the manual-pointing method for the subject WK, who was a typical subject. These Figures show the results for the lateral and polar angle, respectively. For all panels, the target angles are shown on the x-axis, the perceived angles (responses) are shown on the y-axis. The top and bottom rows show the results for the conditions “Dark” and “HMD”, respectively. In Fig. 4, the quadrant errors are plotted as points, the responses in the correct quadrant are plotted as circles. Fig. 5 and Fig. 6 show the results for the head-pointing method obtained from the same subject. All other conventions are as in Fig. 3 and Fig. 4.

The average results over all subjects are summarized in Tab. 1. This table shows the different average errors and their standard errors calculated over all positions and subjects for each condition.

### 3.1. Lateral bias

Two RM ANOVAs with the factors pointer, VE, and position were performed on the data for the left and right hemispheres separately. The positions for the lateral angle were grouped in four groups: ±45° and ±15°. The effect of pointer was not significant ($p > 0.05$), the effect of the VE was significant ($p < 0.001$) showing a smaller bias for the HMD conditions. The effect of position was significant ($p < 0.001$) showing an overestimation of the central targets and an underestimation of the lateral targets. However, this pattern differed between VE and pointer as supported by the interaction between pointer and VE ($p = 0.006$ and $p = 0.049$ for the left and right hemisphere, respectively). The analysis of this interaction revealed that the pointer shows no effect in darkness (bias of 6.4°). However, for the condition HMD, the lateral bias was smaller for the head pointer (4.9°) than for the manual pointer (5.5°).

### 3.2. Lateral spread

The RM ANOVA with the factors pointer, VE, and positions was performed on the absolute differences between the target and response lateral angles. The positions for the lateral angle were grouped in six groups: ±75°, ±45° and ±15°. The effect of VE was significant ($p < 0.001$) showing a smaller error for the HMD condition (13.4°) than in the darkness (16.3°). Especially at lateral positions, the subjects showed more uncertainty in the darkness than with the HMD ($p < 0.0001$ for the interaction position x VE). The much lower spread at

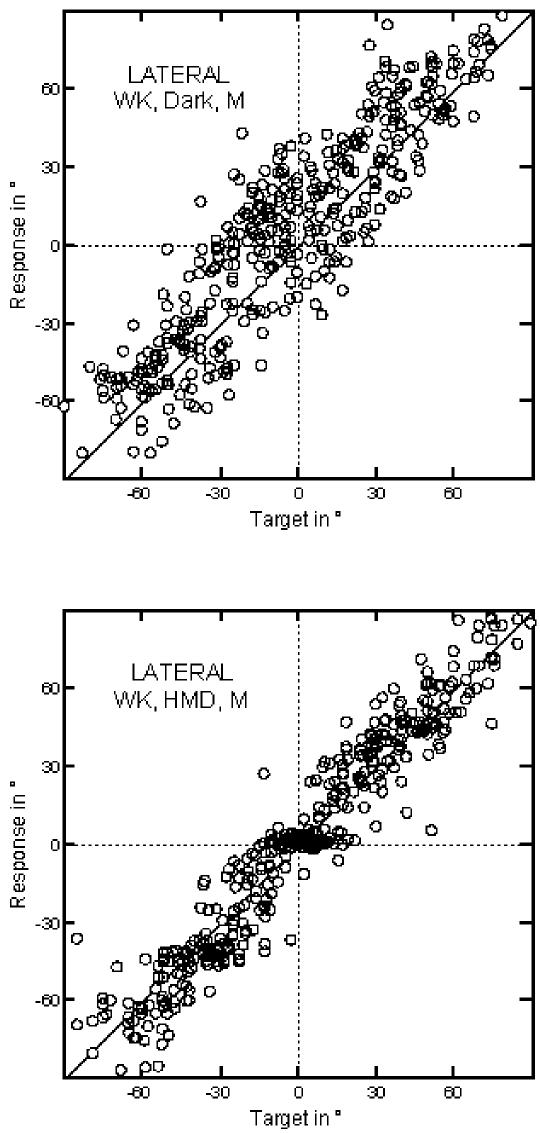


*Figure 3: Results in the lateral dimension for the subject WK using the manual pointer. The top panel shows the condition in the darkness. The bottom panel shows the condition with the VE.*

the median positions indicates a larger anchoring effect when the HMD is used.

The effect of the pointer was significant (p = 0.025), however, the average difference was very small (head: 14.58°; manual: 15.08°). The reason for the small difference despite high significance is an additional dependence on vision and positions, as supported by the significant interaction between positions, pointer, and VE (p = 0.018). In other words, the head pointing results in a lower spread only when HMD is used and only for median positions.

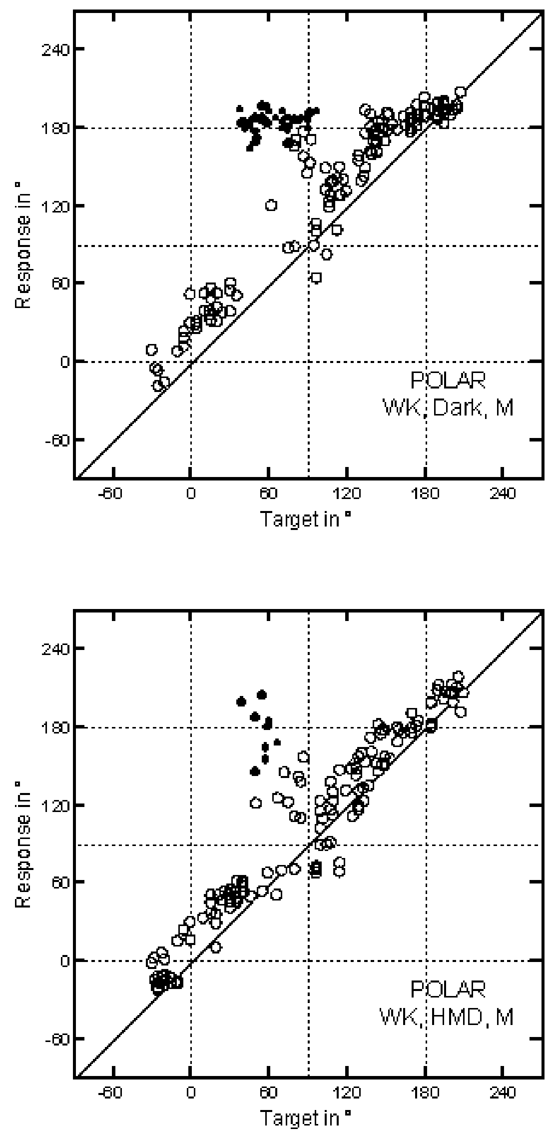


*Figure 4: Results in the polar dimension for the subject WK using the manual pointer. The dots represent the responses classified as front-back confusions. The circles represent the data in the correct quadrant. All other conventions are as in Fig. 3.*

### 3.3. Quadrant errors

The number of the front-back confusions for each condition were fit to log-linear models with factors position, pointer, and VE. The front and rear planes were analyzed separately as the visual inspection of the data

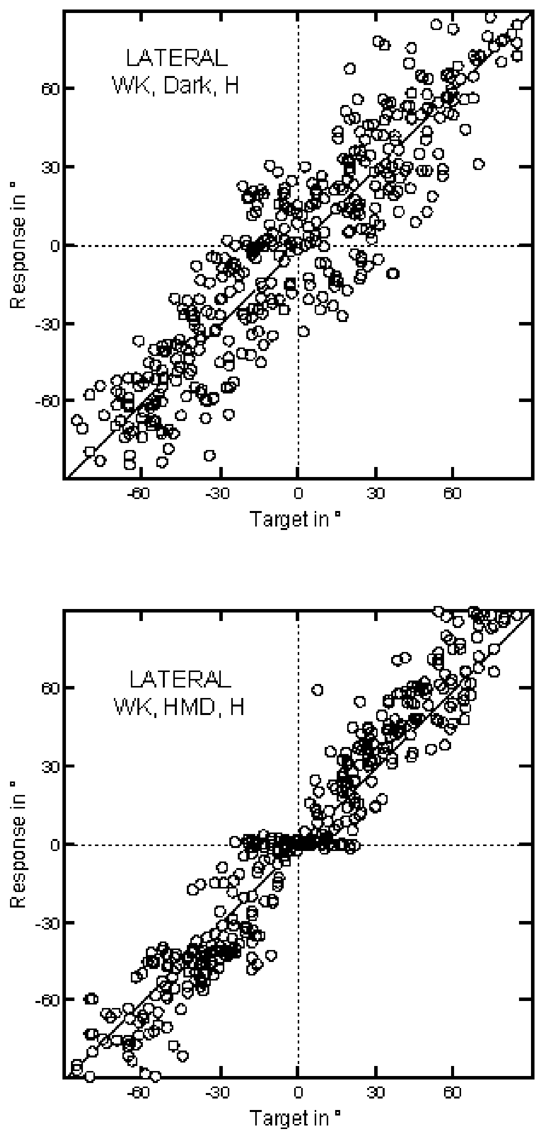


*Figure 5: Results in the lateral dimension for the subject WK using the head pointer. All other conventions are as in Fig. 3.*

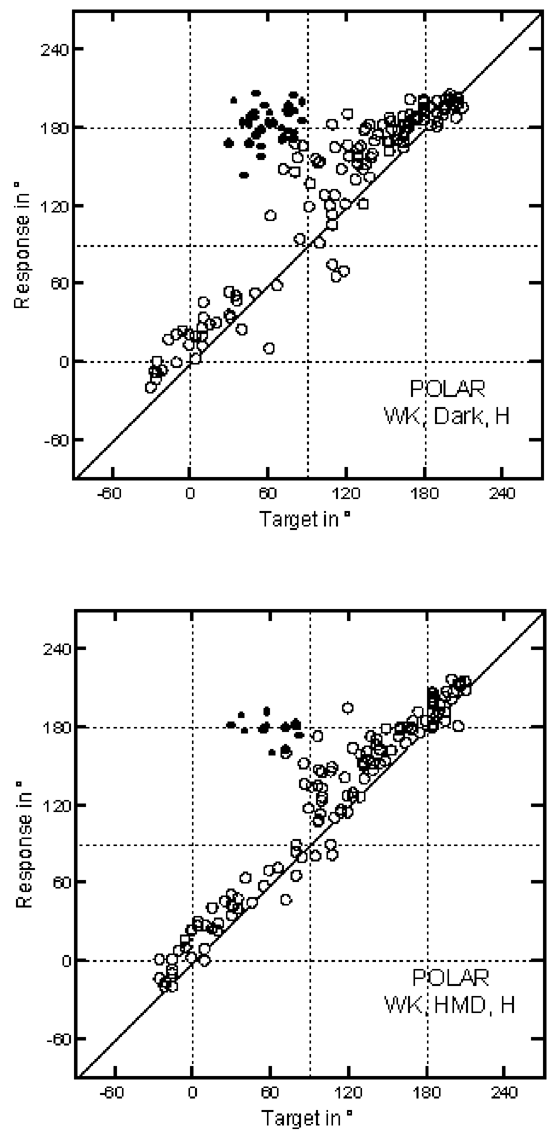


*Figure 6: Results in the polar dimension for the subject WK using the head pointer. All other conventions are as in Fig. 4.*

showed more front-to-back than back-to-front confusions (compare Fig. 4). The positions were placed to two groups: 0° (eye-level) and 60° (above the eye-level). Only targets and responses within the lateral range of ±30° were considered.

For the frontal plane, the model which included the factors position and VE was sufficient to represent the data (goodness-of-fit $G^2 = 0.677$). The effect of VE was significant ($p < 0.001$), showing an advantage of using HMD (26.2%), compared to the darkness (30.4%). The smallest quadrant errors were for the manual pointer and HMD (25.2% and 20.9% for the positions 0° and 60°, respectively).

For the rear plane, the best model included factors position, pointer, and the interaction position x pointer ($G^2 = 0.453$). The effect of pointer was significant ($p = 0.021$) showing smaller quadrant errors for manual pointer (14.4%) than for the head pointer (15.6%). However, this effect interacted with the position as indicated by the significant interaction between pointer and position ($p = 0.032$). Thus, the quadrant error was smaller when localizing rear-eye-level targets with the head

| | Dark | | HMD | |
|---|---|---|---|---|
| | Head | Manual | Head | Manual |
| **Lateral bias** | 6.5° | 6.4° | 4.9° | 5.5° |
| **Lateral spread** | 15.9° | 16.7° | 13.3° | 13.5° |
| **Quadrant errors (front)** | 30.3% | 30.3% | 26.2% | 26.2% |
| **Quadrant errors (back)** | 15.6% | 14.4% | 15.6% | 14.4% |
| **Raw polar spread** | 52.5° | 52.5° | 51.7° | 48.2° |
| **Corrected polar spread** | 32.3° | 32.3° | 30.5° | 30.5° |

*Table 1: Summary of localization errors as averages over all subjects. See text for details.*

pointer (16.4% and 14.8% for the eye-level and upper positions, respectively). When the manual pointer was used, the rear-upper targets showed smaller quadrant error (12.7% and 15.9% for the eye-level and upper positions, respectively). However, the differences were very small and are not congruent with the results of the frontal plane targets.

The significant effect of VE for the frontal plane indicates that the subjects were able to localize the acoustic targets more often in the frontal plane when the visual environment was present. For the rear plane the VE did not contribute.

The advantage of one of the pointing methods could be seen only in case of a small amount of front-back confusions, which was the case for the rear plane only.

### 3.4. Polar spread

The RM ANOVA with the factors pointer, VE, and positions was performed on the absolute differences between the target and response polar angles. The position groups for the polar angle were 0°, 60°, 120° and 180°. Only targets and responses within the lateral range of ±30° were considered. The results showed a significantly ($p = 0.004$) smaller spread for the HMD (49.9°) than in the darkness (52.5°). The effect of the pointer was weak ($p = 0.054$). However, there was a significant interaction between pointer and VE ($p=0.047$). In the darkness, both pointing methods showed similar errors (head: 52.46°; manual: 52.51°). Using HMD, the manual pointer yielded significantly smaller error (48.2°) than head pointer (51.7°). This indicates that subject were able to have an advantage of manual pointing when they had visual feedback.

In this analysis, we found a significant effect of position ($p < 0.0001$) revealing the highest errors for the frontal positions. This may be an effect of many quadrant errors for these positions.

Thus, the RM ANOVA was performed for the corrected polar angles where the front-back confusion were corrected. This analysis confirmed the significantly ($p = 0.001$) smaller spread for HMD (30.53°) compared to the darkness (32.32°). Also the effect of position was still significant ($p < 0.001$) with significantly lower error for the rear positions (35°) than the other positions (50° to 60°). This is probably an effect of many front-back confusions and indicates the limitations of this analysis. The difference between Dark and HMD was larger for the upper positions (6.2°) than for the eye-level positions (4.57°) as supported by the significant interaction between position and VE ($p = 0.001$). This indicates that the subjects were less precise while pointing to the top positions in darkness than while pointing to the top positions with HMD.

The significant pointer effect found for the raw polar spread was not found for the corrected polar spread ($p = 0.436$). This indicates that the resolution of quadrant errors removed the difference between head and manual pointing, which is probably an artifact of the correction algorithm and shows its limitations.

## 4. DISCUSSION

This study investigated the effect of sound localization in darkness and in a virtual visual environment (VE). Additionally, two response methods were studied: head and manual hand pointing.

The results show that the virtual VE significantly improved the localization ability of subjects in the horizontal and vertical dimensions. This supports the benefit of having virtual VE even in tasks which involve acoustic stimuli only.

The results for the pointing method require a more complex interpretation. In the lateral dimension, the pointing method showed only effect when the subjects had access to the VE. In such a case, the head pointing showed a smaller lateral bias for all positions and smaller lateral spread for median positions. Thus, in cases where the horizontal direction is of primary interest, the head-pointing method is a good choice. The front-back confusions depended marginally on the pointer method. In contrast, the pointing method had an impact on the polar spread, which indicates that the subject responded differently for the two pointing methods once they chose the plane. This seems to be reasonable because the pointing method has no effect on the perceived position of the sound, which means that the listener is able to identify the correct quadrant independent of the pointing method. However, the listener's response is affected by the pointing method, which affects the local accuracy of the response as indicated by the differences in bias and spread for the two pointing methods. According to our data, the manual pointing results in a higher precision in the vertical direction. Hence, in localization tasks, where high accuracy in the vertical dimension is required, manual pointing is the better choice. However, these results may be confounded by the high amount of front-back confusions, which is probably an effect of testing naïve subjects in this study.

## 5. ACKNOWLEDGMENTS

We would like to thank Johannes Zmölnig from the Institute for Electronic Music and Acoustics at the University of Music and Dramatic Arts in Graz for the technical support in the development of the virtual environment. We also would like to thank our subjects for their patience during the longsome tests. This study was supported by the Austrian Science Fund (FWF, P18401-B15).

## 6. REFERENCES

[1] Mason, R., Ford, N., Rumsey, F., and De Bruyn, B. (**2001**). "Verbal and Nonverbal Elicitation Techniques in the Subjective Assessment of Spatial Sound Reproduction," *J Audio Eng Soc*, **5**, 366-384.

[2] Haber, L., Haber, R. N., Penningroth, S., Novak, K., and Radgowski, H. (**1993**). "Comparison of nine methods of indicating the direction to objects: data from blind adults," *Perception*, **22**, 35-47.

[3] Bronkhorst, A. W. (**1995**). "Localization of real and virtual sound sources," *J Acoust Soc Am*, **98**, 2542-2553.

[4] Pinek, B., and Brouchon, M. (**1992**). "Head Turning versus Manual Pointing to Auditory Targets in Normal Hearing Subjects and in Subjects with Right Parietal Damage," *Brain and Cognition*, **18**, 1-11.

[5] Middlebrooks, J. C. (**1999**). "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," *J Acoust Soc Am*, **106**, 1493-510.

[6] Oldfield, S. R., and Parker, S. P. A. (**1984**). "Acuity of sound localization: a topography of auditory space. I. Normal hearing conditions," *Perception*, **13**, 581-600.

[7] Lewald, J., and Ehrenstein, W. H. (**1998**). "Auditory-visual spatial integration: A new psychophysical approach using laser pointing to acoustic targets," *J Acoust Soc Am*, **104**, 1586-1597.

[8] Seeber, B. (**2002**). "A new method for localization studies," *Acta Acoustica - Acustica*, **88**, 446-450.

[9] Begault, D. R., Wenzel, E. M., and Anderson, M. R. (**2001**). "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**, 904-916.

[10] Shelton, B. R., and Searle, C. L. (**1980**). "The influence of vision on the absolute identification of sound-source position," *Perception and Psychophysics*, **28**, 589-596.

[11] Lewald, J., Dörrscheidt, G. J., and Ehrenstein, W. H. (**2000**). "Sound localization with eccentric head position," *Behavioural Brain Research*, **108**, 105-125.

[12] Konishi, M. (**1986**). "Centrally synthesized maps of sensory space," *Trends Nerosci*, **9**, 163-168.

[13] Redon, C., and Hay, L. (**2005**). "Role of visual context and oculomotor conditions in pointing accuracy," *NeuroReport*, **16**, 2065-2067.

[14] Stanney, K. (**2002**). *Handbook of Virtual Environments*, (Lawrence, Mahwah, New Jersey)

[15] Viirre, E. and Bush, D. (**2002**). "Direct Effects of Virtual Environments on Users," in *Handbook of Virtual Environments*, edited by K. M. Stanney (Lawrence, Mahwah, New Jersey)

[16] Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., and Tam, C. (**2006**). "Perceptual recalibration in human sound localization: learning to remediate front-back reversals," *J Acoust Soc Am*, **120**, 343-59.

[17] Møller, H., Sørensen, M. F., Hammershøi, D., and Jensen, C. B. (**1995**). "Head-Related Transfer Functions of Human Subjects," *J. Audio Eng. Soc*, **43**, 300-321.

[18] Majdak, P., Balazs, P., and Laback, B. (**2007**). "Multiple Exponential Sweep Method for Fast Measurement of Head-Related Transfer Functions," *J Audio Eng Soc*, **55**, 623-637.

[19] Morimoto, M., and Aokata, H. (**1984**). "Localization cues in the upper hemisphere," *J Acoust Soc Jpn (E)*, **5**, 165-173.

[20] Lienert, G. A. (**1978**). *Verteilungsfreie Methoden in der Biostatistik (Non-parametric methods in biostatistics)*, (2nd edition, Verlag Anton Hain, Meisenheim am Glan)

[21] Majdak, P., Laback, B., and Baumgartner, W. (**2006**). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," *J Acoust Soc Am*, **120**, 2190-201.