Applied Acoustics 203 (2023) 109223

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Distance discrimination thresholds of proximal sound sources in a real anechoic environment ☆

Zhenyu Guo [a], Yuezhe Zhao [a], Liliang Wang [b], Yijing Chu [a], Guangzheng Yu [b,*]

[a] State Key Laboratory of Subtropical Building Science, School of Architecture, South China University of Technology, No.381, Wushan Road, Guangzhou 510641, Guangdong, China
[b] School of Physics and Optoelectronics, South China University of Technology, No.381, Wushan Road, Guangzhou 510641, Guangdong, China

ARTICLE INFO

Article history:
Received 1 November 2022
Received in revised form 20 December 2022
Accepted 9 January 2023

Keywords:
Distance localization
Just-noticeable difference
Proximal sound
Adaptive test

ABSTRACT

When a sound source off the midline is located close to the head, the interaural level difference changes dramatically with distance. Although distance discrimination thresholds for frontal sound sources have been comprehensively studied, the thresholds for real lateral sound sources in the proximal region have rarely been explored. The present study measured distance discrimination thresholds via a loudspeaker located at different distances and directions, both with fixed-emission and normalized intensities referring to the head center, to confirm the contribution of near-field binaural cues in a quantitative manner. The results demonstrated that the sound level is the main cue for distance discrimination, and binaural cues promote the ability to perceive differences in sound distances for lateral sources located at 0.5 m from the head center. The thresholds for lateral sources were significantly lower than those for sources in the median plane, both with and without level cues. However, when the sound sources were located at a distance of 1.0 m, the promotion of lateral thresholds was not significant. This study provides plausible just-noticeable differences for distance discrimination at various directions and distances from the head.
© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

Spatial hearing is an essential function of the human auditory system, enabling people to avoid unseen dangers and promoting the extraction of sound information (e.g., speech) from noise [1]. Spatial cues are also known to make sounds more realistic [2] and externalized [3], which is why state-of-the-art sound reproduction systems focus on representing spatial reality [4]. Moreover, exploring human spatial hearing could facilitate the design of better sound reproduction systems [5] as well as robotic localization systems [6].

Sound source localization involves the perception of two aspects of sound: its direction and its distance. Considerable research has focused on the directional localization of sound in the horizontal and median planes [7]. Studies have revealed that humans can perceive sound directions in a quite precise way; for example, the minimum audible angle in the horizontal plane can be as low as one degree for front sound sources [8].

In contrast, auditory distance perception (ADP) has received relatively little attention and distance discrimination thresholds are largely unknown. This lack of research may be because ADP is much fuzzier and complicated [9]. In particular, people struggle to estimate the distance of a sound source, and salient errors in sound distance estimation are commonly seen in distance localization experiments [10]. Generally, there is a tendency to overestimate the distance of nearby sources [11] and underestimate that of distant sources (i.e., acoustic horizon) [12]. The mechanism of ADP is also complex and susceptible to non-acoustic factors. Previous research has suggested that the localization of sound sources is a cross-modal procedure involving both vision [13] and our sense of motion [14]. For instance, Etchemendy et al. [15] showed that the acoustic horizon is significantly diminished when using a visual-assisted report mechanism in localization experiments.

1.1. Distance localization cues

It has been well-documented that the following distance localization cues are associated with distance perception:

(i) Binaural cues. Binaural cues include the interaural level difference (ILD) introduced by the head-shadow effect (i.e., the blocking of sound propagation by the head) and the interaural time difference (ITD) induced by the acoustic path difference

☆ Portions of this work were presented in "Discrimination experiment of sound distance perception for a real source in near-field," 1st EAA Spatial Audio Signal Processing Symposium, Paris, France, 2019

* Corresponding author.
E-mail address: scgzyu@scut.edu.cn (G. Yu).

https://doi.org/10.1016/j.apacoust.2023.109223
0003-682X/© 2023 Elsevier Ltd. All rights reserved.

between the two ears. The analysis of near-field head-related transfer functions (HRTFs) revealed that ITD is almost independent of distance, but ILD varies with the distance to a near-field sound source (⩽1.0 m) [16,17]. The HRTF magnitudes for both ears and the corresponding ILDs for various distances and azimuths are presented in Fig. 1. The near-field HRTFs were measured using the KEMAR manikin with the procedure proposed in previous work [18]. Note that HRTFs are normalized using the magnitude at the head center, which removes $1/r$ distance-related (referenced to the head center) level cues [4]. A comparison between levels with the $1/r$ distance-related factor removed and retained is shown in Fig. 1 (f) and Fig. 1 (g). For the front sound source [see Fig. 1 (a) and Fig. 1 (b)], the magnitude of the HRTFs only varies with distance in the high-frequency range due to the HRTF parallax. For lateral sound sources [see Fig. 1 (c) and Fig. 1 (d)], the magnitude of the HRTF decreases with distance in the ipsilateral ear. In contrast, the HRTF magnitude increases with distance in the contralateral ear. These opposite changes between the two ears lead to variations in ILD with distance [see Fig. 1 (d)]. One of the reasons for the opposite trends with distance in the two ears is that the paths from the sound source to the head center and the two ears differ. When the sound source is located in the proximal region, the distance ratio of the head center and ear deviates from 1 dramatically, especially for lateral sources. The variation of ILD in the near field is much larger than the just-noticeable difference (JND). Some studies have demonstrated that the accuracy of distance estimation is higher for sources located laterally than for those in the median plane [19,20]. In particular, Kopco et al. (2012) [21] observed ADP-related electrophysiological activities produced by binaural cues in auditory cortex areas.

(ii) Level cues. The level of sound, which varies in inverse proportion to the sound source distance (i.e., $1/r$ law for a point source in the free field), is a dominant cue [9,22]. The received level of sound depends on the distance and the intensity of the source itself. Hence, the level cue is a relative cue for distance localization. When the sound source is located in the far field, the difference in the sound source distance to the head center and to the ears is negligible. However, the difference could be non-negligible when the sound source is located in the near field [see Fig. 1]. In the present study, we used the head center as the reference point for the distance of the sound source, because the difference comes from the binaural structure of the human body and the reference point is in line with the HRTF.

Other cues contribute to distance localization in the free field, such as spectral cues [23]. In the reverberation space, the direct-to-reverberant ratio (DRR) is an essential cue for distance localization [9]. Additionally, the ADP relies on dynamic cues such as the Doppler frequency shift [24], acoustic tau [25], and motion parallax [26].

1.2. Psychoacoustic experiment using lateral sound sources

Among the abovementioned ADP localization cues, binaural cues have been extensively studied for several decades [27], and some psychoacoustic experiments have been conducted to validate their contribution to ADP. It is widely assumed that distance local-

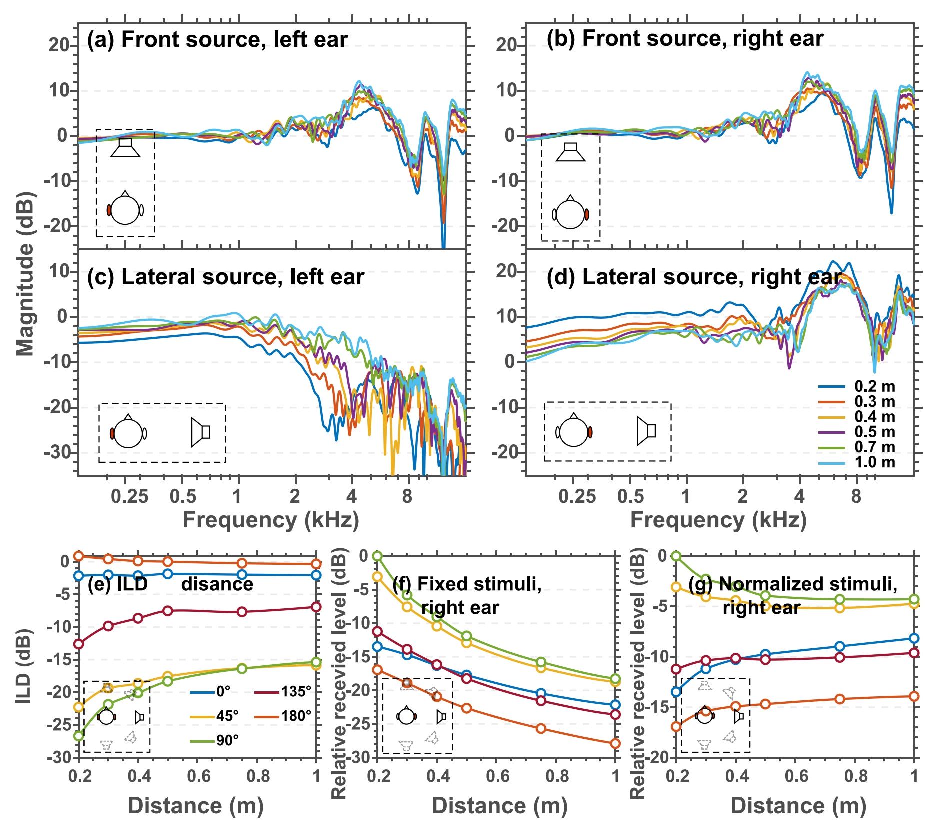

Fig. 1. Comparison of magnitudes and ILDs of near-field measurement KEMAR HRTFs. Subfigures (a) and (b) illustrate the HRTFs of the left and right ears in the 0° azimuth, respectively. Subfigures (c) and (d) show the HRTFs of the left and right ears in the 90° azimuth, respectively. HRTFs presented in four subfigures (a) to (d) were normalized with the magnitude at the head center. Subfigure (e) plots the ILDs under 3 kHz in the azimuth 0°, 45°, 90°, 135° and 180°. The circles indicate ILDs calculated with HRTFs; other ILDs were obtained through interpolation. Subfigures (f) and (g) present the overall received levels of the right ears at different distances and azimuths [the same as subfigure (e)] under fixed and normalized stimuli conditions, respectively. Subfigure (f) presents the HRTF level directly, while subfigure (g) compensates the $1/r$ level attenuation with distance to HRTF levels.

ization for lateral sound sources can be improved with binaural cues, especially in the proximal region where ILD changes dramatically.

To clarify the contribution of binaural cues to distance localization, two paradigms of distance localization have been adopted in experiments: (i) absolute distance localization and (ii) relative distance discrimination. The former requires subjects to estimate and report the distance of sound exactly, whereas the latter asks subjects to compare several stimuli and judge their relative distance.

The absolute distance localization paradigm is a direct way to determine whether certain cues help with distance localization, but large within-group variations were often recorded during such experiments since no reference was provided to the subjects. For instance, in the study of Brungart et al. (1999) [20], the average distance estimation error was 30–40% across different sound directions. This study also implied that the average slope between the estimated and actual distances was approximately 0.39, which is far lower than 1 (exact estimation) [28].

In terms of relative distance discrimination tasks, the discrimination thresholds are commonly measured through two methods, namely, the method of constant stimuli [21] and the adaptive method [29]. The former is more reliable, but takes a long time, while the latter can measure the threshold quickly.

Because ADP is largely dominated by the level of sound sources varying with distance, different methods have been utilized to isolate the impact of distance-related level cues. One method involves normalization of the level by compensating the $1/r$ attenuation [30]. Other studies have used loudness rather than the overall level to conduct the normalization [31]. When the sound source is located in the near field, these normalization methods cannot ensure that the received levels at both ears do not change with distance. For example, the HRTF levels shown in Fig. 1 (g) can be deemed as the received levels normalized by a factor of $1/r$ (with reference to the head center). The received level still changes with distance when the sound source is located within 0.5 m. Some studies randomly rove the level of sound in a small range to prevent participants from judging the distance using the sound level [20].

Despite a considerable number of distance localization experiments being conducted, two issues remain to be solved: (i) the JND thresholds for lateral sound sources measured with real sound sources in an anechoic chamber, which are intrinsically important for understanding the distance localization ability, are absent; and (ii) psychoacoustic experiments conducted by different researchers have obtained contradictory results regarding lateral distance localization.

For the first issue, Kopco et al. (2012) [21] measured the accuracy of distance discrimination of lateral stimuli at various distances with the interference of reverberation. Spagnol et al. (2017b) [30] measured the accuracy of distance discrimination, but did not consider the JND thresholds. In the research of Spagnol et al. (2017a) [29] and Spagnol et al. (2015) [32], the thresholds of lateral distance discrimination were measured by synthesizing a virtual sound via the distance variation function method. However, the lateral distance discrimination thresholds have not yet been determined using real sound sources.

Regarding the second issue, some studies support the assertion that near-field binaural cues promote lateral distance localization. An experiment conducted by Holt and Thurlow (1969) [33] indicated that subjects were capable of discriminating the distances of lateral sound sources, even when the sound level was normalized. In the absolute localization experiments of Brungart et al. (1999) [20] and Brungart (1999a) [19], subjects performed better with lateral sources than with medial sources. Kan et al. (2009) [34] contrasted the ADP performance under binaural synthesis with and without synthesized near-field ILD changes, and found that the accuracy of ADP was enhanced for a source located 40 cm away when near-field ILD changes were included.

In contrast, several studies have found that binaural cues are not strong enough to promote distance perception. In the study of Spagnol et al. (2017b) [30], who used a similar model as Kan et al. (2009) [34] in a relative distance localization experiment, no significant enhancement in distance perception was observed when using lateral sources. In the relative distance discrimination experiment conducted by Arend et al. (2021) [31], subjects could barely discriminate the distance in the absence of intensity cues, as the accuracy for all source directions was below the level of chance (i.e., 50% in that study). In another absolute distance localization study reported by Shinn-Cunningham et al. (2000) [35], the accuracy of distance localization for both monaural and binaural conditions was below the chance level. Moreover, Simpson and Stanton (1973) [36] and Arend et al. (2021) [31] found that subjects could not utilize binaural cues to promote the accuracy of distance localization by moving their heads.

1.3. Aim of present study

The present study aimed to measure the JND thresholds of lateral sound sources in proximal regions with real sound sources. Using these thresholds, the different contributions of binaural cues and $1/r$ distance-related level cues could then be interpreted quantitatively. The distance reference point was set in the head center, which is in line with the common definition of HRTFs [37]. Accordingly, the distance-related level variation was normalized by a factor of $1/r$. There remains a change of level after normalization, hence this study did not distinguish the contributions of the absolute received level and ILD cues. This consideration is specific to the current fashion of distance rendering of the statical virtual sound sources, hence the results may contribute to the understanding and the improvement of distance rendering.

A real sound source was used to ensure the fidelity of binaural cues, especially for extremely close sound sources. We did not use virtual sound sources because there are technical difficulties in recording individual HRTFs in the near field ($\leqslant$1.0 m), as the near-field HRTFs are strongly affected by the receiving position, and so large errors can be induced if the subjects move their heads during the measurement process [34,18]. Additionally, the degraded fidelity of virtual reproductions of proximal sound sources may further degrade the accuracy of distance localization [38].

The discrimination thresholds of distance under different conditions were measured using the adaptive method. Sound sources from different directions were considered in the experiments, given that binaural cues are only available for lateral sound sources. To provide further evidence of the effects of binaural cues alone, i.e., confirm whether binaural cues can be used by the auditory system when other stronger cues exist, the experiments considered both fixed and normalized sound levels [see Fig. 1(f) and Fig. 1(g)]. As both level cues and binaural cues may contribute differently at different distances, the thresholds were measured in two distance regions (less than 0.5 m and between 0.5 m and 1.0 m).

2. Methods

Experiments were conducted to determine the thresholds of sound distance discrimination. To avoid the distribution of reverberation and other non-acoustic cues from being used by the subjects, the experiments were performed in an anechoic chamber (background noise below −12.1 dBA) and the subjects were blindfolded during the entire period of the tests. The subjects' heads were fixed using a headrest during the experiment.

2.1. Subjects

Eight otologically normal subjects (four male, four female, aged 23–41 years, mean age 29 years, standard deviation 6 years) participated in the experiments. All subjects were recruited from the South China University of Technology. Some of them had prior exposure to psychoacoustic experiments, but none of them had previously participated in distance discrimination tests. For each subject, the experiment took about four hours over two days; the whole experimental process was finished in five days. Subjects were remunerated for their time.

2.2. Apparatus and stimuli

Previous studies commonly placed a loudspeaker on a rail and moved it manually [39,20]. However, manual devices make it difficult and time-consuming to carry out complicated experiments, especially for adaptive tests in which the position of the sound source is relevant to the result of previous trials. To overcome the inconvenience of a hand-moved loudspeaker system, a custom electric test platform was designed as follows.

Fig. 2 shows the apparatus and environment of the experiment. To control the location of the loudspeaker and the time interval precisely, the loudspeaker (Mission M30i) was mounted on an electric slide, which was driven by a two-phase stepper (57BYGH75) rather than being pushed manually. Both the location of the loudspeaker and the stimuli were manipulated by a computer. The maximum movement speed was 500 mm/s and the accumulated error was no more than 0.1 mm. Thus, the loudspeaker movement did not last more than 1.2 s during each test, given that the greatest distance of travel was 600 mm. The boundary of the slide was placed about 0.2 m from the edge of the head. The mechanical system was fixed on a stable aluminum bracket and covered with sound-absorbing cotton to diminish the reflections and movement noise of the electric slide. For each subject, the height of the ears was adjusted to be the same as the height of the center of the loudspeaker (1.3 m). In the central position of the head, the noise of speaker movement was less than 35 dBA (measured when no subject was present), and subjects reported that they could not judge the direction of movement from the movement noise.

In terms of the audio stream, an amplifier (ARCAM A65) cascaded with an external sound card (RME Fireface UC, sampling rate = 44.1 kHz) was used to drive the speaker. Broadband pink noise of 1 s duration with a 20 ms ramp-on and -off was used as the stimulus in the experiment.

Two categories of stimuli, i.e., normalized- and fixed-level stimuli, were used in the experiment. The normalized-level stimuli (denoted as normalized stimuli hereafter) had their levels normalized by a factor of $1/r$, so that the sound levels of sources at different distances were constant (75 dBA) in the central position of the head (measured when no subject was present) [26,30]. In contrast, the fixed-level stimuli (denoted fixed stimuli hereafter) were played at a fixed power (75 dBA for the source at 1 m from the center of the head), and so the level ascended naturally as the distance to the sound source decreased.

2.3. Procedures

Before being formally tested, each subject performed a brief training procedure consisting of several trials with feedback so that they could familiarize themselves with the task and the stimuli. No feedback was supplied to the subjects in the formal test.

In the formal test, each subject performed threshold measurement under a total of 20 conditions (2 reference distances $\times$ 5 sound azimuths $\times$ 2 stimulus types). The order of the test blocks was randomized across subjects. Two reference distances of

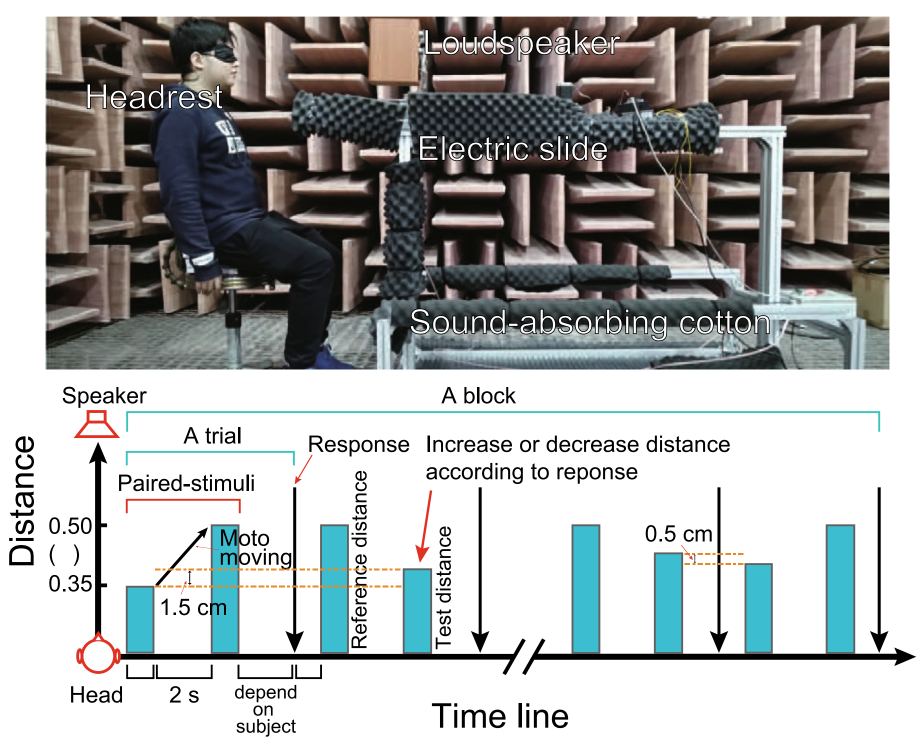

Fig. 2. Facilities and procedure of the experiment. The upper sub-figure presents the test platform and environment used in this study. The lower sub-figure illustrates an instance of the threshold measurement procedure when d_r = 0.5 m..

0.5 m and 1.0 m were considered in the experiment. Five azimuthal directions in the right plane were incorporated (i.e., 0°, 45°, 90°, 135°, and 180°, wherein 0° and 90° refer to the front and right, respectively). The thresholds at different azimuths were measured by altering the orientation of the subjects. The two stimulus types refer to the normalized and fixed stimuli.

Fig. 2 depicts the procedure of a specific test block in the formal test. During each test block, a pair of stimuli were played at two different distances with a 2-s interval (i.e., reference and test distance, denoted by d_r and d_t, respectively). The loudspeaker was moved to the midpoint of d_r and d_t in advance to shorten its travel distance.

In terms of the paired stimuli, the reference and test distance were ordered at random. The reference distance was consistent throughout a whole block, i.e., 0.5 m or 1.0 m depending on the test condition. The test distance was adjusted according to the response of the present subject in the previous trial during the test. Note that the test distance was always less than the reference distance.

After the paired stimuli had been played, a two-alternative forced-choice task was assigned to the subjects. Specifically, the subjects were instructed to report verbally which stimulus was closer to them, i.e., the first one or the second one. There was no time limit for subjects to give their answer.

To determine the distance discrimination threshold, a 2-down-1-up adaptive procedure was implemented. This method converges to the threshold with a 70.7% probability of a positive response [40]. The initial d_t in the first trial for each test block was 30% less than d_r (i.e., 0.35 m when d_r was equal to 0.5 m and 0.7 m when d_r was equal to 1.0 m) when the sound was located at azimuths of 0°, 45°, 90°, and 135°. A larger initial d_t, i.e., 40% less than d_r, was set for the rear sound source (azimuth of 180°), because a pilot experiment revealed that distance discrimination was worse when the source was located to the rear [17]. For subsequent trials, the sound source was moved farther from or closer to the reference distance in steps of 1% (proportion of the reference distance) when the subject made an incorrect judgment or two successive correct judgments, respectively. Before the first incorrect judgment, the step size was set to 3% rather than 1% to reduce the duration of the experiment. The adaptive procedure terminated after 12 reversals (a reversal refers to a shift from an increment to a decrement of d_t, or vice versa). The average distance of the last five reversals (denoted by d_t') was determined as the threshold distance. The deviation between this distance and the reference distance can be regarded as the JND for the 70.7% correct performance.

3. Results

3.1. Overview

In total, 160 thresholds (8 subjects × 20 conditions) were obtained. The distance thresholds are presented in Fig. 3 for subjects S1–S8. With the 1.0 m reference distance, the blue markers (i.e., fixed sound sources) are closer to the reference distance than the red markers (i.e., normalized sources). This means that the subjects were more sensitive to differences in the distance when the sound level cues were preserved. A similar phenomenon can be observed for the 0.5 m reference distance, although the deviation between the two types of stimuli was more subtle in this case.

Given that the sensitivity of ADP to distance changes is roughly constant with a fixed distance ratio [21], the JND threshold was calculated as follows:

$$JND_r = \frac{d_r - d_t'}{d_r}, \quad (1)$$

where JND_r denotes the distance discrimination threshold of reference distance r; this is referred to as the 'threshold' in the remainder of this paper. The average JND for stimuli with and without levels were 11.7% (standard deviation, SD = 6.0%) and 24.8% (SD = 11.4%) for the 0.5 m reference distance and 5.9% (SD = 2.6%) and 28.3% (SD = 8.0%) for the 1.0 m reference distance. Furthermore, for the 0.5 m reference distance, the JND thresholds of the two types of stimuli were fairly similar for a lateral source, except for subjects S7 and S8. Unexpectedly, in two instances, subjects performed slightly better with the normalized sound level—S5 under the 45° condition and S3 under the 90° condition. These results might indicate that level cues are less important for discriminating distance differences of laterally proximal sources within 0.5 m.

3.2. Statistical analysis of JND results

Fig. 3 and Fig. 4 (b) show the average JNDs across subjects obtained with fixed (i.e., with the level cue) and normalized (i.e., without the level cue) stimuli, respectively. These figures intu-

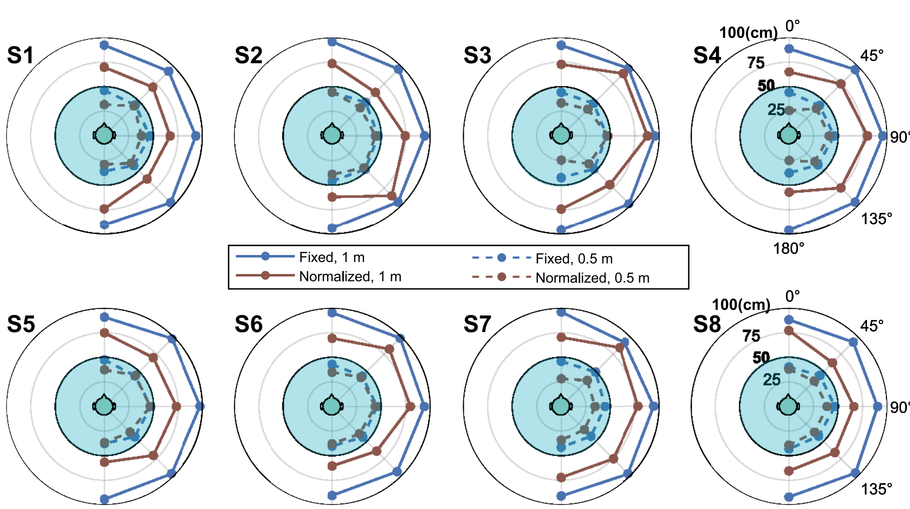

Fig. 3. Threshold results $d_{t'}$ (in cm) for 70.7% correct performance of eight subjects. Red and blue markers represent thresholds of normalized and fixed stimuli, respectively. The dashed lines in the blue region and the solid lines in the white region represent thresholds under the 0.5 m and 1.0 m reference distances, respectively. The reference distances are highlighted by black solid lines.

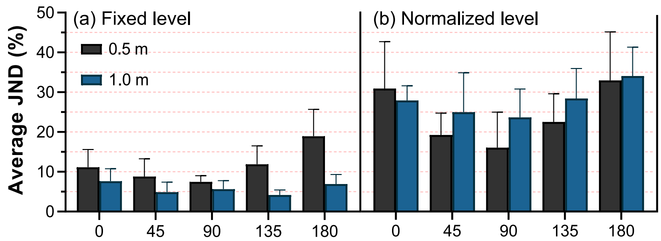

Fig. 4. Across-subject average JNDs under (a) fixed stimuli and (b) normalized stimuli condition. The bars and vertical lines illustrate the average JNDs and corresponding standard deviations across the eight subjects, respectively.

itively illustrate two points: (i) normalizing the level of the stimulus induced a clear declination in the distance discrimination ability and a dramatic fluctuation among subjects—SDs across all conditions under normalized and fixed stimuli were 9.9 and 5.4, respectively; (ii) under both normalized and fixed stimuli, subjects performed better for the lateral sound sources compared with those located in the median plane. Compared with a frontal source, a lateral source at 90° (reference distance of 0.5 m) reduced the average JNDs from 11.2% (SD = 4.5) to 7.5% (SD = 1.5) and from 31.5% (SD = 12.0) to 16.4% (SD = 9.1) in the fixed and normalized cases, respectively.

The JNDs of each condition conformed to a normal distribution according to the Anderson–Darling test. A three-way repeated-measures analysis of variance (rmANOVA) was performed on the JNDs to uncover the effects and interactions of the three within-subject factors, i.e., reference distance, source azimuth, and stimulus type. The Geisser–Greenhouse correction was applied to correct the violation of the sphericity assumption. The rmANOVA results presented in Table 1 indicate that the direction $[F(4,28) = 12.97, p = 0.0003]$ and type of stimulus $[F(1,28) = 551.3, p = 0.0002]$ significantly affected the JND, whereas the reference distance did not have a significant effect $[F(1,28) = 0.49, p = 0.41]$. The level cue of sound accounted for 55.6% of the total variation ($p = 0.0002$), making it the main factor in determining the JNDs. The binaural cue, which can be interpreted from the difference among performance across various orientations of the sound source, contributed 9.3% of the total variation ($p = 0.0003$). Furthermore, the rmANOVA results indicated that both the reference distance $\times$source azimuth $[F(4,28) = 3.23, p = 0.027]$ and reference distance $\times$stimulus type $[F(1,28) = 9.64, p = 0.041]$ were significant interactions. The three-way interaction was not significant. Note that the homogeneity of variances was not fulfilled, so the statistical results of three-way rmANOVA may be biased. Thus, the detailed two-way rmANOVA analyses were conducted as follows.

Table 1
Results of three-way ANOVA analysis.

Factor	*Var* (%)	*SS*	*df*	*F*	*p*
Reference Distance (RD)	0.22	51.2	1	0.49	.409
Stimulus Azimuth (SA)	9.28***	2113	4	12.97	<.001***
Stimulus Type (ST)	55.62***	12671	1	551.3	<.001***
RD × SA	1.56*	354.7	4	3.29	.027*
RD × ST	3.81*	868.8	1	9.64	.042*
SA × ST	1.43	325.6	4	1.84	.149
RD × SA × ST	1.00	228.0	4	2.43	.150
Error		655.9	28		

Caps * and *** indicate significance levels of 0.05 and 0.001, respectively.

A two-way rmANOVA (reference distance and source azimuth) was conducted on the JNDs of two types of stimulus types separately. Only for fixed-level stimuli, the reference distance had a significant main effect on the JNDs $[F(1,28) = 48.44, p = 0.0002]$. For normalized stimuli, the reference distance made no significant difference. The azimuth of the sound source influenced the JNDs significantly with both the fixed sound $[F(4,28) = 11.04, p = 0.0028]$ and the normalized sound $[F(4,28) = 6.68, p = 0.0063]$. These results revealed that the azimuth factor contributed to the JNDs. Additionally, a significant interaction effect for the reference distance $\times$azimuth was only found for the fixed-level stimuli. Specifically, according to a Tukey posthoc multiple comparisons for fixed-level stimuli, the subjects performed worse with the 0.5 m reference distance than with the 1.0 m reference distance when the sound source was located at 135° ($p = 0.010$) and 180° ($p = 0.009$). These results are counterintuitive, because the level of sound forms a constant ratio with distance (i.e., it obeys the $1/r$ law), and thus provides an invariable physical cue.

Next, a two-way (reference distance and stimuli type) rm ANOVA was applied to the JNDs of the five azimuths separately and corresponding posthoc tests were implemented. For all azimuths, the stimuli type has a significant effect on the JNDs. For azimuths of 90°, 135°, and 180°, the stimuli type × reference distance has a significant interaction effect. The JNDs of the fixed stimuli were significantly lower than those of the normalized stimuli in almost all conditions. However, there was no significant difference between the two types of stimuli when the sound is located in 90° azimuths at 0.5 m reference distance.

Since an interaction effect was found, a one-way (source azimuth) rmANOVA and multiple comparisons were applied to all four combined conditions, i.e., 2 reference distances × 2 stimulus types, to compare the JNDs of different azimuths The results indicated that, with a reference distance of 0.5 m, there was a significant difference among azimuths under both the normalized stimuli $[F(4,28) = 8.73, p = 0.0043]$ and the fixed stimuli $[F(4,28) = 7.24, p = 0.0048]$. Post-hoc multiple comparisons indicated that 45° versus 180° ($p = 0.006$), 90° versus 180° ($p = 0.012$), and 135° versus 180° ($p = 0.028$) had significant JND diversities under the 0.5 m reference distance with the fixed stimuli. For the normalized stimuli, there was a significant difference between 0° and 90° ($p = 0.036$) under the 0.5 m reference distance. Although 0° and 45°, 180° and 45°, 180° and 90° also exhibit considerable average differences over 10%, i.e., 11.9%, 14.0%, and 17.21%, respectively, these differences were not significant due to the variances. No significant effect of sound direction was observed under the reference distance of 1.0 m. In summary, for both types of stimuli, the JNDs of lateral sources were lower than those of front or rear within 0.5 m. When the reference distance was 1.0 m, the difference between the JNDs of median and lateral sources was not significant for any type of stimulus.

4. Discussion

In the present study, the capacity to perceive differences in proximal sound distances has been evaluated through the adaptive method. The JNDs with a 70.7% detection probability were measured under various conditions, including sound sources with fixed and normalized levels, five directions, and two reference distances in a free field. According to the experimental results, the JND for perceiving distance differences is heavily dependent on the level cues and direction of the sound, and changes to some degree with the distance of the sound source. Theoretically, these results suggest that both binaural cues and level cues contribute to ADP. The influence of both types of cues are discussed in the following subsections.

4.1. Binaural cues

The benefit of binaural cues can be extracted through the JND differences between lateral sound sources (i.e., 45°, 90°, and 135°) and median sources (i.e., 0° and 180°), since ILD only exists when the sound source is oriented laterally to the subjects. According to the results presented in this paper, the average JNDs across subjects were lower for lateral sources than for median sources (see Table 2). The average JNDs decreased as the source approached the interaural axis, and reached a minimum when the source was located to the side of the subjects. This enhanced ADP for lateral sources has been reported in some previous studies [41,20], which suggested that binaural cues might provide additional information for ADP. However, in our experiments, significant improvements in the discrimination of distance were only found when the reference distance was 0.5 m. This is somewhat expected, as Brungart and Rabinowitz (1999) [16] showed that the slope of ILD with respect to distance was not steep around 1.0 m ([see also Fig. 1 (e) to (g)]). The relative invariance of ILD and received level under normalized stimuli around 1.0 m may account for the less significant results with sound sources around 1.0 m from the subjects.

The results indicate that the azimuth factor accounted for 18.5% and 25.8% of the total variation under normalized and fixed stimuli, respectively. Especially for lateral sources within 0.5 m, significant improvements were found. Both Spagnol et al. (2017a) [29] and Spagnol et al. (2015) [32] used a similar test procedure (i.e., adaptive method), but found no significant improvement in lateral distance discrimination. One possible reason is that a real sound source produced by a loudspeaker, rather than a binaural virtual source via near-field HRTFs, was used in the present study. The near-field HRTFs used by Spagnol et al. (2017a) [29] and Spagnol et al. (2015) [32] were synthesized by the distance variation function method; another study using a similar synthetic method also did not observe any significant difference between the results for lateral and median sources [30]. Binaural cues restored with higher fidelity could benefit the auditory discrimination of sources at different distances.

Binaural cues may not be as robust for ADP as level cues, as mentioned in previous research [3,34]. Compared with fixed stimuli, the variability between subjects in the case of normalized stimuli was clearly larger (average standard deviation of JND increased by 4.91% across reference distances and azimuths). Minor improvements in stability can be observed when comparing the SDs of the JNDs of median and lateral sources, because more distance localization cues were available. For example, when the reference distance was at the 0.5 cm reference distance, the SDs of the 90° sources were 2.75% and 2.70% lower than those of the 0° azimuth sources for fixed and normalized stimuli, respectively.

4.2. Level cues

With regard to the effect of distance-related level cues, our results show that these are the main cues that enable the auditory system to perceive differences in sound distance. When the level of the stimuli was normalized, the JND of distance discrimination increased from 8.8% (SD = 5.4%) to 26.6% (SD = 10.0%) when averaged across all directions and reference distances.

There is a hypothesis that if the variation in sound level with distance provides a sufficiently strong cue for ADP, then the JND of distance discrimination in the median plane is equal to the JND of intensity discrimination of sound, which theoretically corresponds to a 5% change in distance [42]. The average JNDs for fixed stimuli at 1.0 m reference distance range from 4.2% (135°) to 7.7% (0°), which corresponds to 0.37 and 0.70 dB intensity difference, respectively. These results were in line with the intensity discrimination thresholds (0.41 dB) of the white noise reported in Miller (1947) [43] and thresholds (average 0.51 dB across frequencies) of the 80 dB pulsed sinusoids reported in Jesteadt et al. (1977) [44]. Hence, the results endorse that the distance-related level change is the dominant cue for ADP when the sound source is located around 1 m.

For the front and rear sound source conditions, where the binaural cues are irrelevant, the average JNDs for the fixed stimuli were 11.1% and 18.9% for the 0.5 m reference distance and 7.7% and 6.9% for the 1.0 m reference distance, respectively. The results indicate that ADP cannot benefit from rapid variations in level with respect to distance when the source is located in the proximal area. It is worth noticing that the minimum discriminable distances (i.e., $d_r - d_t'$) for the fixed stimuli at 0.5 m and 1.0 m reference distances exhibit no significant difference. The same change of distance will induce a larger change in the received level for closer sources when considering the $1/r$ attenuation law [see Fig. 1 (f)]. In other words, the ADP is less sensitive to the distance-related change of level, hence performs worse in the proximal area. A similar phenomenon has been observed in studies using relative distance discrimination tests [45], and this result agrees with the widely observed degradation in nearby sound localization (i.e., overestimation trend) to some extent.

The results for the rear sound source at the 1.0 m reference distance are in line with those of a previous study [42], where the JNDs were between 3.5% and 6.0%; Spagnol et al. (2017a) [29] obtained JNDs of around 7% for receding stimuli at 1.0 m. In research using the 2-down-1-up method, the JNDs were found to

Table 2
Comparison of JND threshold results obtained in previous research.

Condition		JND (%)						
Stimuli	Azimuth	This study	Spagnol 2017			Spagnol 2015		
Fixed			Mean	Approach	Recede	Mean	Approach	Recede
0.5 m	0	11.1	12	6	18	19	8	18
	90	7.5	12	2.5	22	12.5	5	20
1.0 m	0	7.7	13	20	6	15	23	7
	90	5.6	13.5	17	10	14	20	8
Normalized								
0.5 m	0	31.5	17	16	18	/	/	/
	90	16.4	16	12	20	/	/	/
1.0 m	0	28.4	17	18	16	/	/	/
	90	24.1	17.5	15	20	/	/	/

The JNDs obtained by Spagnol et al. (2015) [32] and Spagnol et al. (2017a) [29] and were extracted from the figures in these papers. 'Approach' and 'recede' denote whether the test stimuli appear after or before the reference stimuli, respectively. 'Mean' denotes the average of the 'approach' and 'recede' JND.

be 7.3% and 5.7% for a frontal source at 1.0 m when subjects performed the experiment the first time and when they were familiar with the stimuli, respectively [45]. Note that a 70.7% correct proportion, rather than 50%, was required in the present study; the JNDs under a 50% chance level are presumed to be closer to 5%. The JNDs obtained in this study at 1.0 m were all close to 5% (7.67%, 4.90%, 5.65%, 4.17%, and 6.94% for 0°, 45°, 90°, 135°, and 180°, respectively), which implies that the level cue is dominant for ADP.

4.3. General discussion

In Table 2, the JNDs measured in this study are compared with those obtained in other research using a similar paradigm. In previous studies, significant improvement was only observed for lateral sources when the sound source approached the subject (i.e., from far to near). For receding stimuli, no lateral advantage was found. In terms of normalized stimuli, the JNDs of front stimuli in Spagnol's study were lower than those in the present study, e.g., for front normalized stimuli, the JNDs were 31.5% and about 17% in our study and Spagnol et al. (2017a) [29]. However, the JNDs of lateral sound sources were quite close for both 0.5 m and 1.0 m reference distances. For fixed stimuli, the JNDs for approaching and receding stimuli have a larger gap in Spagnols' studies. Our JNDs have a similar trend with those reported by Spagnol under the approaching condition. The effect of virtual sources versus real sources and different normalization methods may account for the disparities. Besides, the studies of Spagnol roved the level in a small range at random, while the present study normalized the level with respect to the head center.

In summary, the current experimental results show that binaural cues only benefited the perception of distance differences when the sound was located close to the head (0.5 m in this study). The enhanced lateral discrimination of proximal sources is in line with the steep slope of ILD with respect to distance in this region. The binaural advantage exists for both normalized and fixed stimuli when the source is within 0.5 m, which contradicts some of the abovementioned research. Level cues are still the primary and most robust cues for distance discrimination. When the sound source was located around 1.0 m away, the JND for discriminating differences of distances approached the corresponding JND threshold for perceiving sound level differences (i.e., 5%). Since the free field is distinct to the reverberation environment and the reverberations contribute to ADP greatly, these conclusions only apply to the free field currently.

5. Limitations of the study

The present study used a loudspeaker rather than a binaural synthesis system to perform the experiments, and broadband pink noise was used as the sound source. Thus, the spectral cues were theoretically not isolated in the experiment. However, when considering that the spectral cues are related to the head and the spherical wavefront, they should make a similar contribution when the sound source is located in different directions This means that the contribution of binaural cues could still be verified by comparing the results of different azimuths.

Although the facilities and procedures used in this study were subtly controlled, several drawbacks still exist. The flooring of the speaker distance was about 0.2 m from the head, which limits the measurement of JNDs at closer reference distances. Besides, more participants are needed in future research, especially for the JNDs of the normalized stimuli.

The received levels at the ear still varied with distances slightly, even when normalization was imposed (see Fig. 1). The residual level comes from the position difference between the head center and the two ears. This was incorporated into a part of the binaural contribution in the present study. Hence, this study did not identify the contribution of absolute received levels.

6. Conclusions

This study conducted a psychoacoustic experiment with eight subjects via a specially designed test platform in an anechoic chamber. The JNDs for perceiving differences in the distance of a real sound source were measured under various conditions through the adaptive method. The experimental results can be summarized as follows:

(1) Although binaural cues are weaker and less robust than level cues, they are beneficial for distance discrimination with lateral sound sources, especially when the source is located within 0.5 m from the subject. Considering that ADP is less sensitive to level variations due to changes in distance, the importance of binaural cues for distance discrimination is non-negligible. Additionally, binaural cues make a significant contribution regardless of whether level cues are present.
(2) Level cues are the most essential and robust component of sound for ADP in the free field, especially for sound sources located about 1.0 m from the head. For frontal sound sources, the JND threshold of distance perception approaches that of the corresponding sound intensity (i.e., 5%) at 1.0 m. When the sound is closer to the head (i.e., less than 0.5 m), the JND increased significantly. The minimum discriminable distances were found to be similar when the source is located at 0.5 and 1.0 m, which means that level cues are less reliable for ADP in the proximal area.

The present study provides conclusive and quantitative evidence that binaural cues can assist the human auditory system to discriminate the distance of a sound source. Only two reference distances 0.5 m and 1.0 m, were tested in the present study. Future works measuring JNDs at more reference-distance conditions are warranted, especially for reference distances below 0.5 m.

CRediT authorship contribution statement

Zhenyu Guo: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft. **Yuezhe Zhao:** Resources, Writing - review & editing, Supervision. **Liliang Wang:** Methodology, Software, Investigation. **Yijing Chu:** Writing - review & editing, Resources. **Guangzheng Yu:** Conceptualization, Funding acquisition, Project administration, Supervision, Project administration, Resources.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No. 12074129 and 11574090) and the Fundamental Research Funds for the Central Universities (No.

2022ZYGXZR104). The authors are grateful to Huali Zhou, Jun Zhu, and Yamin Wang for their comments, which helped to improve the manuscript. We express our gratitude to all subjects who participated in the study.

References

[1] Freyman RL, Balakrishnan U, Helfer KS. Spatial release from informational masking in speech recognition. J Acoust Soc Am 2001;109(5 Pt 1):2112–22. https://doi.org/10.1121/1.1354984.
[2] Xu C, Kang J. Soundscape evaluation: Binaural or monaural? J Acoust Soc Am 2019;145(5):3208. https://doi.org/10.1121/1.5102164.
[3] Prud'homme L, Lavandier M. Do we need two ears to perceive the distance of a virtual frontal sound source? J Acoust Soc Am 2020;148(3):1614–23. https://doi.org/10.1121/10.0001954.
[4] Xie B. Head-related transfer function and virtual auditory display. Plantation, FL: J. Ross Publishing; 2013.
[5] Kang D-S, Choi J-W, Martens WL. Distance perception of a virtual sound source synthesized near the listener position. Multimed Tools Appl 2015;75(9):5161–82. https://doi.org/10.1007/s11042-015-2878-9.
[6] Magassouba A, Bertin N, Chaumette F. Exploiting the distance information of the interaural level difference for binaural robot motion control. IEEE Robot Autom Lett 2018;3(3):2048–55. https://doi.org/10.1109/LRA.2018.2806560.
[7] Middlebrooks JC. Sound localization. Handb Clin Neurol 2015;129:99–116. https://doi.org/10.1016/B978-0-444-62630-1.00006-8.
[8] Mills AW. On the minimum audible angle. J Acoust Soc Am 1958;30(4):237–46. https://doi.org/10.1121/1.1909553.
[9] Zahorik P, Brungart DS, Bronkhorst AW. Auditory distance perception in humans: A summary of past and present research. Acta Acustica united with Acustica 2005;91(3):409–20.
[10] Brungart DS, Rabinowitz WM, Durlach NI. Evaluation of response methods for the localization of nearby objects. Percept Psychophys 2000;62(1):48–65. https://doi.org/10.3758/bf03212060.
[11] Brungart DS, Scott KR. The effects of production and presentation level on the auditory distance perception of speech. J Acoust Soc Am 2001;110(1):425–40. https://doi.org/10.1121/1.1379730.
[12] Bronkhorst AW, Houtgast T. Auditory distance perception in rooms. Nature 1999;397(6719):517–20. https://doi.org/10.1038/17374.
[13] Kolarik AJ, Raman R, Moore BCJ, Cirstea S, Gopalakrishnan S, Pardhan S. The accuracy of auditory spatial judgments in the visually impaired is dependent on sound source distance. Sci Rep 2020;10(1). https://doi.org/10.1038/s41598-020-64306-8.
[14] Hug MX, Vergara RO, Tommasini FC, Etchemendy PE, Bermejo F, Fernandez LG. Reaching measures and feedback effects in auditory peripersonal space. Sci Rep 2019;9(1):9476. https://doi.org/10.1038/s41598-019-45755-2.
[15] Etchemendy PE, Spiousas I, Calcagno ER, Abregu E, Eguia MC, Vergara RO. Direct-location versus verbal report methods for measuring auditory distance perception in the far field. Behav Res Methods 2018;50(3):1234–47. https://doi.org/10.3758/s13428-017-0939-x.
[16] Brungart DS, Rabinowitz WM. Auditory localization of nearby sources. Head-related transfer functions. J Acoust Soc Am 1999;106(3 Pt 1):1465–79. https://doi.org/10.1121/1.427180.
[17] Yu G, Wang L. Effect of individualized head-related transfer functions on distance perception in virtual reproduction for a nearby sound source. Arch Acoust 2019;251–258.
[18] Yu G, Wu R, Liu Y, Xie B. Near-field head-related transfer-function measurement and database of human subjects. J Acoust Soc Am 2018;143(3):EL194. https://doi.org/10.1121/1.5027019.
[19] Brungart DS. Auditory localization of nearby sources. III. Stimulus effects. J Acoust Soc Am 1999;106(6):3589–602. https://doi.org/10.1121/1.428212.
[20] Brungart DS, Durlach NI, Rabinowitz WM. Auditory localization of nearby sources. II. Localization of a broadband source. J Acoust Soc Am 1999;106(4 Pt 1):1956–68. https://doi.org/10.1121/1.427943.
[21] Kopco N, Huang S, Belliveau JW, Raij T, Tengshe C, Ahveninen J. Neuronal representations of distance in human auditory cortex. Proc Natl Acad Sci 2012;109(27):11019–24. https://doi.org/10.1073/pnas.1119496109.
[22] Kolarik AJ, Moore BC, Zahorik P, Cirstea S, Pardhan S. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. Atten Percept Psychophys 2016;78(2):373–95. https://doi.org/10.3758/s13414-015-1015-1.
[23] Coleman PD. Dual role of frequency spectrum in determination of auditory distance. J Acoust Soc Am 1968;44(2):631–2. https://doi.org/10.1121/1.1911132.
[24] Rosenblum LD, Carello C, Pastore RE. Relative effectiveness of three stimulus variables for locating a moving sound source. Perception 1987;16(2):175–86. https://doi.org/10.1068/p160175.
[25] Ashmead DH, Davis DL, Northington A. Contribution of listeners' approaching motion to auditory distance perception. J Exp Psychol: Hum Percept Perform 1995;21(2):239–56. https://doi.org/10.1037/0096-1523.21.2.239.
[26] Genzel D, Schutte M, Brimijoin WO, MacNeilage PR, Wiegrebe L. Psychophysical evidence for auditory motion parallax. Proc Natl Acad Sci USA 2018;115(16):4264–9. https://doi.org/10.1073/pnas.1712058115.
[27] Coleman PD. An analysis of cues to auditory depth perception in free space. Psychol Bull 1963;60:302–15. https://doi.org/10.1037/h0045716.
[28] Zahorik P. Assessing auditory distance perception using virtual acoustics. J Acoust Soc Am 2002;111(4):1832–46. https://doi.org/10.1121/1.1458027.
[29] Spagnol S, Hoffmann R, Kristjansson A, Avanzini F. Effects of stimulus order on auditory distance discrimination of virtual nearby sound sources. J Acoust Soc Am 2017;141(4):EL375. https://doi.org/10.1121/1.4979842.
[30] Spagnol S, Tavazzi E, Avanzini F. Distance rendering and perception of nearby virtual sound sources with a near-field filter model. Appl Acoust 2017;115:61–73. https://doi.org/10.1016/j.apacoust.2016.08.015.
[31] Arend JM, Liesefeld HR, Pörschmann C. On the influence of non-individual binaural cues and the impact of level normalization on auditory distance estimation of nearby sound sources. Acta Acustica 2021;5:10. https://doi.org/10.1051/aacus/2021001.
[32] S. Spagnol, E. Tavazzi, F. Avanzini, Relative auditory distance discrimination with virtual nearby sound sources, in: 8th Int. Conference on Digital Audio Effects (DAFx-15), 2015, pp. 1–6.
[33] Holt RE, Thurlow WR. Subject orientation and judgment of distance of a sound source. J Acoust Soc Am 1969;46(6):1584–5. https://doi.org/10.1121/1.1911909.
[34] Kan A, Jin C, van Schaik A. A psychophysical evaluation of near-field head-related transfer functions synthesized using a distance variation function. J Acoust Soc Am 2009;125(4):2233–42. https://doi.org/10.1121/1.3081395.
[35] B. Shinn-Cunningham, S. Santarelli, N. Kopco, Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural vs. monaural cues (poster), in: Assoc. Res. Otolaryn. Meeting, Vol. 23, 2000.
[36] Simpson W, Stanton LD. Head movement does not facilitate perception of the distance of a source of sound. Am J Psychol 1973;86(1):151–9.
[37] Møller H. Fundamentals of binaural technology. Appl Acoust 1992;36:171–218. https://doi.org/10.1016/0003-682X(92)90046-U.
[38] D.S. Brungart, B.D. Simpson, Auditory localization of nearby sources in a virtual audio display, in: Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575), 2001, pp. 107–110. doi:10.1109/ASPAA.2001.969554.
[39] Spiousas I, Etchemendy PE, Eguia MC, Calcagno ER, Abregu E, Vergara RO. Sound spectrum influences auditory distance perception of sound sources located in a room environment. Front Psychol 2017;8:969. https://doi.org/10.3389/fpsyg.2017.00969.
[40] H. Levitt, Transformed up-down methods in psychoacoustics, J. Acoust. Soc. Am. 49 (2) (1971) Suppl 2:467+.
[41] Kopco N, Shinn-Cunningham BG. Effect of stimulus spectrum on distance perception for nearby sources. J Acoust Soc Am 2011;130(3):1530–41. https://doi.org/10.1121/1.3613705.
[42] Strybel TZ, Perrott DR. Discrimination of relative distance in the auditory modality: the success and failure of the loudness discrimination hypothesis. J Acoust Soc Am 1984;76(1):318–20. https://doi.org/10.1121/1.391064.
[43] Miller GA. Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness. J Acoust Soc Am 1947;19(4):609–19. https://doi.org/10.1121/1.1916528.
[44] Jesteadt W, Wier CC, Green DM. Intensity discrimination as a function of frequency and sensation level. J Acoust Soc Am 1977;61(1):169–77. https://doi.org/10.1121/1.381278.
[45] Ashmead DH, LeRoy D, Odom RD. Perception of the relative distances of nearby sound sources. Percept Psychophys 1990;47(4):326–31. https://doi.org/10.3758/bf03210871.