Effect of stimulus spectrum on distance perception for nearby sources[a)]

Norbert Kopčo[b)]
Department of Cybernetics and Artificial Intelligence, Technical University of Košice, Košice, Slovakia

Barbara G. Shinn-Cunningham
Department of Cognitive and Neural Systems and the Hearing Research Center, Boston University, Boston, Massachusetts 02215

(Received 30 April 2010; revised 1 June 2011; accepted 24 June 2011)

The effects of stimulus frequency and bandwidth on distance perception were examined for nearby sources in simulated reverberant space. Sources to the side [containing reverberation-related cues and interaural level difference (ILD) cues] and to the front (without ILDs) were simulated. Listeners judged the distance of noise bursts presented at a randomly roving level from simulated distances ranging from 0.15 to 1.7 m. Six stimuli were tested, varying in center frequency (300–5700 Hz) and bandwidth (200–5400 Hz). Performance, measured as the correlation between simulated and response distances, was worse for frontal than for lateral sources. For both simulated directions, performance was inversely proportional to the low-frequency stimulus cutoff, independent of stimulus bandwidth. The dependence of performance on frequency was stronger for frontal sources. These correlation results were well summarized by considering how mean response, as opposed to response variance, changed with stimulus direction and spectrum: (1) little bias was observed for lateral sources, but listeners consistently overestimated distance for frontal nearby sources; (2) for both directions, increasing the low-frequency cut-off reduced the range of responses. These results are consistent with the hypothesis that listeners used a direction-independent but frequency-dependent mapping of a reverberation-related cue, not the ILD cue, to judge source distance.
 [DOI: 10.1121/1.3613705]

PACS number(s): 43.66.Qp, 43.66.Ba, 43.66.Pn [RYL] Pages: 1530–1541

I. INTRODUCTION

Auditory distance perception depends on a combination of multiple acoustic and non-acoustic cues. Factors known to influence distance perception include, for example, the stimulus spectral content, *a priori* knowledge of the stimulus presentation level, azimuthal location of the source, sound reflections from the environment, and visual information about candidate sound sources in the environment (for a review, see Zahorik *et al*., 2005).

The current study examined distance perception for nearby sources in a simulated reverberant environment. For such sources, at least two robust cues for distance are available, each of which can provide some distance information even without *a priori* knowledge of the stimulus level. Specifically, interaural level differences (ILDs) arise for nearby sources located to the side of the listener (Brungart, 1999; Shinn-Cunningham *et al*., 2005), while in reverberant space, the direct-to-reverberant power ratio[1](D/R) provides distance information for sources from all directions (Mershon and King, 1975; Hartmann, 1983; Zahorik, 2002a).

Several previous studies examined nearby-source distance perception in reverberation (Santarelli *et al*., 1999; Santarelli, 2000; Shinn-Cunningham *et al*., 2000). However, no previous study examined the effect of stimulus spectral content or bandwidth on performance, even though there are reasons to expect that these factors will influence perception (a review of the effects of stimulus spectral content on distance perception for distant sources is available in Blauert, 1996). First, both the low-frequency stimulus cutoff as well as stimulus bandwidth influence distance perception of sounds near the listener in anechoic space, where performance is likely to be based on ILDs (Brungart, 1999). Moreover, perceptual sensitivity to ILD is better for broadband stimuli than for narrowband stimuli (Hartmann and Constan, 2002). Similarly, even though D/R sensitivity has been examined in only a few studies, available data suggest that listeners are better at discriminating changes in D/R for stimuli containing low frequencies than for stimuli with only high-frequency content; in addition, D/R sensitivity is greater for broadband stimuli than for narrowband stimuli (Zahorik, 2002b; Larsen *et al*., 2008).

The current study explored how well listeners could utilize level-invariant distance cues in a simulated reverberant space. The main goals were (1) to measure how stimulus spectrum and bandwidth affect distance perception for nearby sources coming from different simulated directions and (2) to compare observed performance to the physical distance cues available to the listener and to explore which cues influenced distance judgments. An analysis of the effect of the overall stimulus level on responses was also performed (see the Appendix).

Frontal and lateral source directions were tested. Comparison of results for these two directions was undertaken in

[a)]Portions of this work were presented at the 27th meeting of the Association for Research in Otolaryngology.
[b)]Author to whom correspondence should be addressed. Electronic mail: kopco@tuke.sk

order to explore the degree to which, in reverberation, listeners (1) only rely on D/R, which varies with distance for both frontal and lateral sources, (2) only rely on ILD, which varies with distance for lateral but not for frontal sources, or (3) rely on D/R and ILD to perceive distance of sound sources.

II. METHODS

Stimuli, generated in virtual auditory space (Carlile, 1996; Zahorik, 2002a), varied in simulated direction (frontal vs lateral), low-frequency cutoff frequency (low, medium, and high), and stimulus bandwidth (narrow and wide). In order to encourage listeners to rely on distance cues other than loudness, the stimulus presentation levels were roved and the listeners were informed of this fact.

A. Subjects

Six subjects (four female and two male, including author NK) participated in this study. All subjects had normal hearing (confirmed by an audiometric screening) defined as thresholds no greater larger than 15 dB hearing level in the range of 125 Hz to 8 000 Hz. Subjects' ages ranged from 25 to 31 yr.

B. Stimuli and setup

All stimuli used in this study were generated using individually measured binaural room impulse responses (BRIRs) that included realistic room reverberation. Unless specified otherwise, all details of the measurement procedures, including the microphone, speaker, and room characteristics, the BRIR measurement technique used, placement procedures, etc. were identical to those described in Shinn-Cunningham *et al.* (2005).

Briefly, individualized BRIRs were measured in a small classroom (3.4 m $\times$ 3.6 m $\times$ 2.9 m height) using the Bose Acoustimass cube speaker (radiation characteristics of this speaker are described in Shinn-Cunningham *et al.*, 2005). The room was carpeted, with hard walls and acoustic tiles covering the ceiling. The room reverberation times, T_{60}, were estimated from 10 recorded room impulse responses using the method formulated by Schroeder (Schroeder, 1965), as implemented by Brown in a MATLAB function available at the Mathworks web site (Brown, 2002). The mean ($\pm$standard deviation) values of T_{60} in octave bands centered at 500, 1 000, 2 000, and 4 000 Hz were 613 ($\pm$175), 508 ($\pm$60), 512 ($\pm$57), and 478 ($\pm$76) ms, respectively. The BRIRs were measured by placing miniature microphones at the blocked entrances of the listeners' ear canals. The loudspeaker was set to face the listener at various distances (15, 19, 25, 38, 50, 75, 100, 140, or 170 cm) from the center of the listener's head, either directly in front of or directly to the right of the listener along the interaural axis, at the level of the listener's ears.

After the BRIR measurement, an individual set of virtual stimuli was generated for each subject. The target stimulus was a 300-ms-long sample of white noise. Specifically, a set of 10 independent noise burst tokens was generated for each subject (i.e., the noise tokens differed between subjects). To manipulate the stimulus frequency content, noises were filtered using linear-phase bandpass filters with a 50-dB stop-band attenuation, designed using the frequency sampling method to create finite impulse responses, FIRs, of length 501 in the MATLAB signal processing toolbox (function fir2). Noise stimuli were then convolved in the time domain with the resulting FIRs to generate the signals presented. Six stimulus types were used, differing in their bandwidths and center frequencies. Table I gives the passband and stop-band cutoffs of the stimuli, as well as the number of ERBs covered by each stimulus (Glasberg and Moore, 1990).

Following the bandpass filtering, stimuli were convolved with individual BRIRs, generating a set of 1080 stimuli for each subject (18 locations $\times$ 6 stimulus spectrum types $\times$ 10 tokens). The overall sound pressure level (SPL) at the right ear was always approximately equal to or greater than the level at the left ear, since sources either came from in front or to the right of the listener. Because of this, in the remainder of the article, the right and left ears are referred to as the near and far ears, respectively. All stimuli were normalized so that the overall SPL at the right, near ear was constant at 74 dB SPL.[2] Without the overall level normalization we imposed, the direct-sound level varied with distance as much as 25 dB at the near ear and by as much as 15 dB at the far ear (for broadband stimuli presented from the side). The normalization eliminated the monaural overall level cue at the near ear; however, it did not eliminate the monaural overall level cue at the far ear (distant sounds could be up to 10 dB louder than near sounds in the far-ear signal). To reduce the reliability of this far-ear monaural overall level distance cue, the overall stimulus level was roved following the normalization at the near ear, with presentation levels

TABLE I. Spectral content of the stimuli used in the experiment. The bandpass filters used to generate the stimuli were defined by their pass-bands (in which attenuation was near 0 dB) and stop-bands (in which the attenuation was 50 dB or more). The number of equivalent rectangular bandwidths (ERBs) (Glasberg and Moore, 1990) is given in the final column.

Stimulus type	Center frequency F_C (kHz)	Pass band (kHz)	Stop bands (kHz)	Number of ERBs
Broadband	1.3	0.3–5.7	$< 0.05, > 5.84$	22.5
Wideband low	0.95	0.3–3.0	$< 0.05, > 3.1$	16.8
Narrowband low	0.39	0.3–0.5	$< 0.05, > 0.75$	3.0
Wideband medium	4.1	3.0–5.7	$< 2.77, > 5.84$	5.6
Narrowband medium	3.0	2.9–3.1	$< 2.68, > 3.26$	0.6
Narrowband high	5.6	5.5–5.7	$< 5.18, > 5.84$	0.3

chosen from a uniform distribution ranging over ± 5 dB (as described below, the Appendix presents analysis showing that all but one subject were able to follow the instructions to disregard overall level when judging distance).[3]

The stimulus files, generated at a sampling rate of 44.1 kHz, were stored on the hard disk of the control computer (IBM PC compatible). On each trial, one of the tokens of the target stimulus was selected and presented through TDT hardware consisting of TDT PD1 D/A converters, a TDT PA4 programmable attenuator to set the level, a TDT HB6 headphone buffer, and Sennheiser HD 580 headphones. No compensation for the headphone transfer function was done. Note that such compensation would change the spectro-temporal characteristics of all stimuli identically. Given that this experiment used relatively arbitrary, unstructured noise stimuli in all trials, there was no reason to expect that listeners would be influenced by the headphone-transfer-function filtering, which imposed a much smaller variation in the absolute spectral content than the variations resulting from the experimental bandpass filtering.

During the experiment, the listener was seated in a single-walled, sound-treated booth. A computer screen showed a top-down view of the listener with his/her arms stretched out, a circle indicating the distance of the farthest stimulus (170 cm), and two radial lines corresponding to the stimulus directions (in front of and to the right of the listener). The screen was used during the experiment to present written instructions to the listener. After each stimulus presentation, the listener indicated its perceived location by clicking at the desired response location in the visible plane using a computer mouse pointer.

C. Experimental procedure

The experiment was divided into four one-hour-long sessions, each performed on a different day. The first session was a practice session (i.e., the data from this session were discarded). Each session was divided into 12 runs, one for each combination of six stimulus spectral conditions, described above, and two directions (frontal and lateral). The order of runs within a session was random and differed both from session to session and from subject to subject. At the start of each run, the subject was informed as to what type of stimulus would be presented and the direction from which it would be simulated. The stimulus spectral content and direction were kept constant within a run. Two sample stimuli were then presented at a fixed level of 74 dB SPL (near ear), equal to the average presented in the run, one from the nearest and one from the farthest location along the simulated direction for that run. These samples were played to enhance the consistency of listeners' responses by presenting them with exemplars of both the closest and farthest stimuli to be tested in the run.

Each run consisted of 45 trials (5 repeats for each of 9 distances), presented in random order. Each trial consisted of a presentation of a stimulus followed by a single response from the subject. The experiment was self-paced: a 300-ms silent pause occurred after each response before the next trial was presented. No feedback was provided.

D. Data analysis

Responses were recorded as the x and y coordinates of the mouse pointer. This information was used to compute the Euclidean distance from the center of the listener's head to the response location based on the distance scale established by the graphical user interface image. Three statistics were used to evaluate the results: (1) Pearson's correlation coefficients between the simulated source distance and the response distance within a run (as was done in Brungart, 1999), (2) the means of the response distance for each simulated distance, and (3) the standard deviations in the response distances, computed as a function of stimulus distance for each run. These statistics were computed on a log-distance scale, separately for data from each run, corresponding to 45 stimulus-response pairs. These results were then averaged across the identical-type runs from the three different sessions (i.e., the data were not pooled across runs from different sessions prior to computing the statistics).

Note that because of the response interface used in this study the range of available responses was fixed, which may have led to floor and ceiling effects in the results. Moreover, because the listeners had to respond by clicking on the top-down view of the experimental scene on a computer screen, they had to mentally transform the perceived location into a response location on the screen. Thus, the response method used here might have introduced some distortions in responses. However, our conclusions are all based on comparisons across conditions, all of which should be affected similarly by the response method. Therefore, it is unlikely that different conclusions would be reached if the responses had been gathered using a different method.

III. BEHAVIORAL RESULTS

A. Correlation coefficient r

1. Results

Figure 1 plots the correlation coefficients between simulated distance and response distance as a function of the stimulus type. Small symbols represent individual subject data; large squares show across-subject averages. For each stimulus type, filled symbols offset to the left represent the frontal source results; open symbols offset to the right represent the lateral source results. The condition groups are ordered by increasing low-frequency cutoff (increasing left to right); within each group, conditions are ordered by stimulus bandwidth (see Table I).

Performance of the individual subjects was generally similar to the subject average data. One exception was subject S6, represented by circles, who had the lowest correlation coefficient in 11 out of the 12 conditions, and whose correlation coefficients were especially low for lateral sources (open symbols). This subject reported great difficulty in performing the task. Moreover, responses for this subject were significantly correlated with the stimulus presentation level (see the Appendix), which was not true for the results for any other subject. Therefore, in the rest of this paper, data from subject S6 are plotted when possible, but are

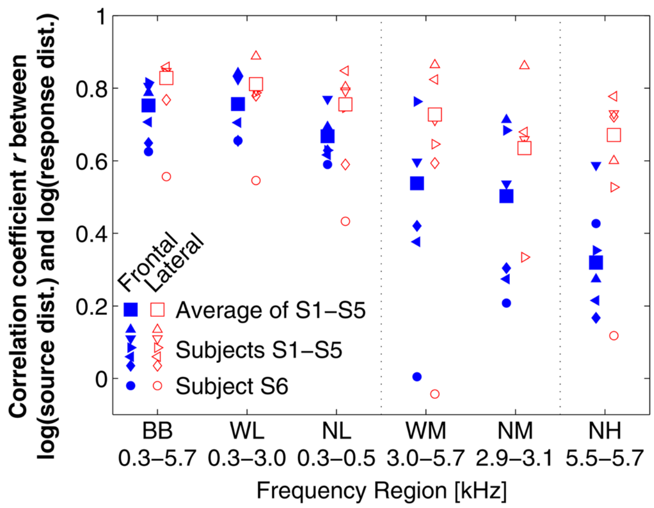

FIG. 1. (Color online) Correlation coefficient r between the logarithm of simulated and response distance as a function of the stimulus type. Vertical dotted lines separate spectral conditions in which frontal performance (filled symbols) differs significantly ($p < 0.05$). Results of pairwise analysis for both frontal and lateral sources are summarized in Table II. BB – broadband, W – wideband, N – narrowband, L – low, M – medium, H – high. Note: The narrowband medium lateral datum for subject S6 (circles, correlation of −0.54) falls below the range of values shown in the figure.

excluded from any across-subject analysis, including plots of across-subject averages and statistical analyses.

A three-way, repeated-measures analysis of variance (ANOVA) with the factors of direction (frontal vs lateral), spectral content (six levels), and repeat (three levels) was performed using the Huynh–Feldt epsilon correction for non-homogeneous data. With this correction, the main effect of spectral content was significant ($F_{5,20} = 3.37$; $p = 0.0001$); both the interaction of direction and spectral content and the main effect of direction revealed a trend, but did not reach significance ($F_{5,20} = 3.37$; $p = 0.051$ and $F_{1,4} = 6.58$; $p = 0.062$, respectively). No other main effects or interactions approached significance.

Given the marginally significant interaction of direction × spectral content, two additional one-way ANOVAs were performed separately on the lateral and the frontal data. Both ANOVAs revealed a significant effect of spectral content ($p < 0.05$). Bonferroni-corrected *post hoc* pairwise *t*-tests (which account for the heterogeneity of variances, as recommended by Ury and Wiggins, 1971, and implemented in the CLEAVE package, Herron, 2005) were used to test for significant differences (Table II). For the frontal data (filled symbols), the analysis revealed three groups of conditions. Within each group, no conditions differed significantly from any other condition ($p > 0.05$); however, all pairwise comparisons of conditions from two different groups revealed significant differences ($p < 0.05$). These three groups were (BB, WL, and NL), (WM and NM), and (NH) (see groups separated by vertical dotted line in Fig. 1). For the lateral data (open symbols), there were only two statistically different groups: (BB and WL) vs (NM and NH) (the two left-most vs the two right-most conditions in Fig. 1). Results for the NL and WM conditions fell in between the groups (however, note that the NL and NM conditions differed significantly from one another).

TABLE II. Pairwise comparisons of the correlation coefficients, *r*, for different stimulus conditions. Table shows condition pairs in which performance significantly differed ($p < 0.05$) for frontal (F) or lateral (L) sources, based on Bonferoni-corrected, post-hoc pairwise t-tests performed separately for the frontal and the lateral data.

Stimulus type	BB	WL	NL	WM	NM	NH
Broadband (BB)	· ·	· ·	· ·	F ·	F L	F L
Wideband low (WL)	· ·	· ·	· ·	F ·	F L	F L
Narrowband low (NL)	· ·	· ·	· ·	F ·	F L	F ·
Wideband medium (WM)	· ·	· ·	· ·	· ·	· ·	F ·
Narrowband medium (NM)	· ·	· ·	· ·	· ·	· ·	F ·
Narrowband high (NH)	· ·	· ·	· ·	· ·	· ·	· ·

Given that the stimulus bandwidth had little effect, a final ANOVA was performed on the full dataset after averaging the data across bandwidths, considering only the low-frequency cut-off as an important factor. Specifically, the data within each of the groups (BB, WL, and NL), (WM and NM), and (NH) were averaged. A three-way, repeated-measures ANOVA with the factors of direction (frontal vs lateral), spectral content (three levels), and repeat (three levels) was performed using the Huynh–Feldt epsilon correction for non-homogeneous data. With this correction, the main effect of spectral content ($F_{2,8} = 19.02$; $p = 0.001$), the main effect of direction ($F_{1,4} = 8.09$; $p = 0.047$), and the direction × spectral content interaction ($F_{2,8} = 6.13$; $p = 0.024$) all reached significance. No other main effects or interactions approached significance.

2. *Interim summary and discussion*

Overall, performance measured as correlation between simulated distance and response distance was better for lateral sources than for frontal sources (open symbols fall above filled symbols). This difference was small for broadband stimuli containing low frequencies (left-most three stimulus conditions in Fig. 1) but grew as the low-frequency cutoff increased (group to group).

For frontal sources (filled symbols), performance decreased as the lowest frequency present in the stimulus increased from low (0.4 kHz) to medium (3 kHz) to high (5.6 kHz; groups in Fig. 1). In contrast, for lateral sources (open symbols), there was a decrease in performance as the cutoff frequency increased; however, this change was smaller than for frontal sources. In particular, there was no drop in performance from the medium- to the high-frequency stimuli [(WM, NM) to NH].

There was no significant effect of the stimulus bandwidth on performance for either frontal or lateral sources (within groups separated by vertical dashed lines, filled symbols are similar to each other and open symbols are similar to each other). Finally, performance was generally more consistent across subjects (i.e., the spread of the individual data points for a given condition is smaller) for the stimuli containing low frequencies (BB, WL, NL) than for the stimuli containing no low frequencies (WM, NM, NH).

B. Mean and standard deviation in responses

If it is assumed that a noisy linear relationship exists between the simulated distances and responses, the size of the correlation coefficient between the two variables depends on both the variability in responses and on the size of the change in the mean response with simulated distance. Specifically, for the fixed range of simulated input distances used here, correlations will be smaller either if response variability increases, even if the range of response means remains fixed, and/or if the range of mean distance responses decreases, even if the response variability remains unchanged. To determine whether the dependence of the correlations on stimulus properties described in the previous section arose due to changes in mean response, changes in response variability, or changes in both, these values were analyzed.

1. Results

Figure 2 shows the across-subject mean of the logarithm of response distance [for frontal sources in panel (A) and for lateral sources in panel (B)] and the average of the standard deviations in the logarithm of the response distances [for frontal sources in panel (D) and for lateral sources in panel (E)] as a function of the simulated source distance. Error bars are omitted from the plots for clarity; however, the average of the standard errors of the across-subject means is shown in each panel, computed by calculating the standard error separately for each distance and stimulus condition and then averaging across the distances and stimulus conditions shown in a given panel. Results in panels (A) and (B) are roughly linear; to summarize results, the slopes of the individual subject's results were estimated by finding a linear fit to the mean log response data as a function of the log of the simulated distance. Panel (C) plots the average of these slope estimates. Similarly, panel (F) plots the mean standard deviation in the responses averaged across subjects and distances.

For frontal sources, listeners overestimated distance for nearby locations [see left portion of panel (A), where all data fall well above the diagonal, shown by the dotted thin line]. In contrast, for lateral sources [panel (B)], listeners were relatively accurate in mean distance judgments and used nearly the entire response range. Stimulus spectral content had a larger effect on responses for frontal sources than for lateral sources [lines differ more from one another in panel (A) than panel (B)]. The mean responses for stimuli containing low frequencies (solid lines) changed the most with source distance, while the responses for the stimuli not containing low frequencies (dashed and dotted lines) were relatively flatter in both panels (A) and (B).

Panel (C) confirms these trends in the mean responses. Specifically, for both the open and the filled symbols, the three left-most data points (BB, WL, NL) are above the WM and NM data points, which, for the filled symbols, are above the NH points. Thus, the slopes of the responses as a function of distance became shallower for both frontal and lateral sources as the low-frequency energy was removed. Moreover, the filled symbols fall below the open symbols, showing that the slopes were shallower for frontal sources than for lateral sources. This dependence of response slope on stimulus type helps explain the variations of the correlation coefficients with source direction and frequency content: a low

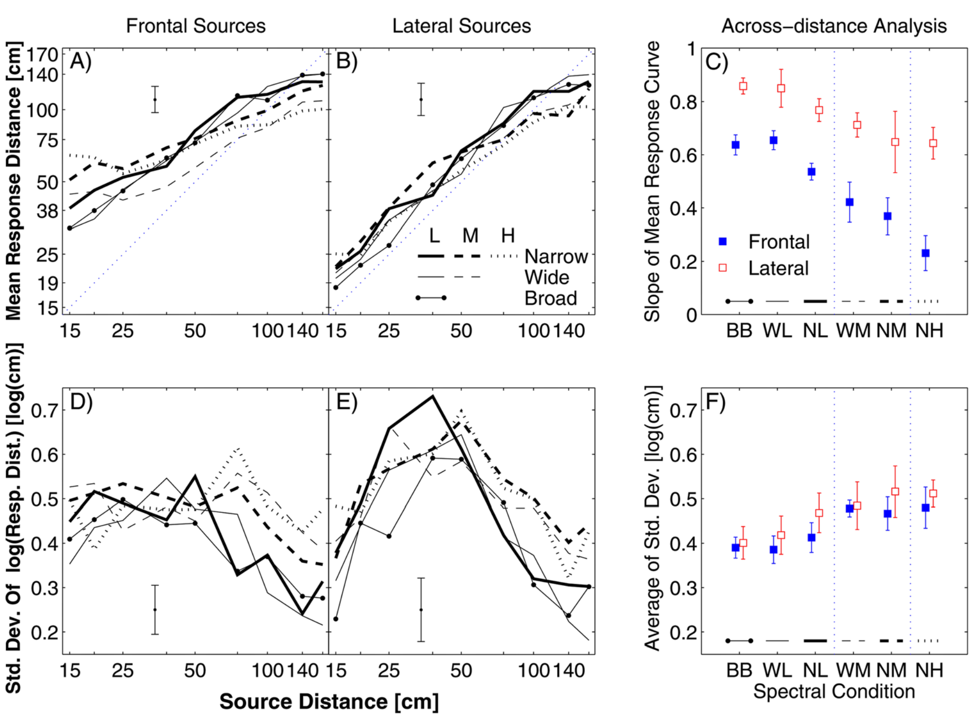

FIG. 2. (Color online) Mean (panels A and B) and standard deviation (panels D and E) of the logarithm of response distance as a function of the simulated source distance for the frontal (A, D) and lateral (B, E) source locations. Graphs show averages across subjects S1–S5. Standard deviations were computed separately for each run, and then averaged across the runs and across subjects. Error bars were omitted from individual data points for clarity; however, a representative error bar in each panel shows the standard error of the across-subject mean, averaged across the data in that panel. (C) Across-subject average (± SEM) in the slope of the mean response distance curves shown in panels (A) and (B). (F) Across-subject average (± SEM) in the standard deviations from panels (D) and (E) collapsed across source distance. Line segments along the lower edge of panels (C) and (F) indicate the line style of the corresponding graph in panel (A), (B), (D), and (E).

correlation was found for conditions in which the slope relating simulated and response distance was shallow. The only major discrepancy between the trends in Fig. 2(C) and Fig. 1 is that the differences in mean response slope between the frontal and lateral data are greater in Fig. 2(C) than the differences in the stimulus-response correlations shown in Fig. 1.

Standard deviations in responses to frontal sources were roughly constant with distance for distances below about 50 cm, and then decreased for greater distances, especially for the BB, WL, and NL conditions, as shown by the solid lines of different widths in Fig. 2(D). In contrast, the standard deviations in responses to lateral sources were smallest at both shortest and largest distances, peaking for distances between 25 and 50 cm [Fig. 2(E)]. Given that the mean localization judgments for the nearby, frontal sources were near the center of the allowable range, response variance for these sources was unlikely to be influenced by any floor effect [left side of Fig. 2(D)]. However, for the most distant frontal sources [right side of Fig. 2(D)] and for the nearest and most distant lateral sources [left and right edges of Fig. 2(E)], floor and ceiling effects in the responses may account for the low response variability that is seen. Thus, to a first order approximation, any dependence of response variability on distance of sources from a particular direction appears to have arisen from floor and ceiling effects. In the middle range of distances, where floor and ceiling effects were negligible for both frontal and lateral sources, response variability for lateral sources was larger than for frontal sources. This difference in variability goes in the opposite direction of the correlation coefficient results: if response ranges were comparable across the two source directions, these differences in variability would yield higher correlations for frontal sources than for lateral sources.

The average standard deviations collapsed across source distance confirm these trends [Fig. 2(F)]: (1) the standard deviations for frontal sources tended to be, if anything, smaller than the standard deviations in lateral sources [the filled symbols have a slight tendency to fall below the open symbols in Fig. 2(F)], and (2) the frontal source standard deviations were similar for the NM and NH conditions [compare the two rightmost filled symbols in Fig. 2(F)], even though the response correlations were higher for the lateral NM stimuli than for the frontal NH stimuli.

2. Interim summary and discussion

Both means and standard deviations were affected by the stimulus spectrum. However, the differences in the stimulus-response correlations across the spectral conditions and source directions shown in Fig. 1 appear to have been dominated by the differences in the mean responses/response ranges. For both directions, the mean response slope was shallower, i.e., response range was smaller, for sources with no low frequency content than for those containing low frequency energy, which helps account for the lower correlations between simulated and response distance for sources without low frequencies. For frontal sources, responses for nearby sources were biased, with listeners reporting that the sources were farther away than their simulated distances. This bias compressed the range of responses for frontal sources, an effect that can explain why the correlation between simulated and response distance was smaller for frontal sources than lateral sources. However, the difference between lateral and frontal correlations was smaller than the corresponding differences in response slopes. This small inconsistency can be accounted for by the modest increase in response standard deviations for lateral sources compared to frontal sources.

The increase in variability seen with the observed increase in response range is similar to the effect that stimulus context has in a broad range of perceptual tasks (e.g., see Parducci and Wedell, 1986; Gescheider, 1988; Parker *et al.*, 2002; Petrov and Anderson, 2005). Specifically, increases in the response range can result in increases in response variability if memory noise, rather than perceptual noise, limits performance (e.g., see Durlach and Braida, 1969; Braida *et al.*, 1984). Here, listeners may have adopted perceptual anchors at the minimum and maximum response distances, which results in increased response "noise" as the response range increases (e.g., see Braida *et al.*, 1984); moreover, the limited range of responses imposed by our response method may have influenced results. Additional studies are needed to determine whether response range directly impacts response variability in tasks like those presented here.

IV. ANALYSIS OF ACOUSTICAL CUES FOR DISTANCE

The primary level-independent distance cues for nearby sources in natural environments are ILD and D/R (Brungart, 1999; Shinn-Cunningham *et al.*, 2005). This section presents an analysis of the ILD and D/R cues contained in the current experimental stimuli. Stimulus bandwidth had only a modest impact on performance in the current study, making it unnecessary to consider how distance cues were integrated across frequencies. Therefore, this analysis focuses on narrowband stimuli, which were at most a few ERBs wide (see Table I). First, the mean D/R and ILD were computed as a function of stimulus distance, both for frontal and lateral directions. To explore the degree to which distance-dependent variations in these cues could explain how listeners judged distance, regression analysis was performed on the mean responses, looking to see how well the mean D/R and ILD values account for mean distance judgments.

Perceptual sensitivity to ILD and D/R cues is likely to affect response variability. Moreover, past studies show that sensitivity to changes in D/R depends on the reference D/R value (e.g., see Larsen *et al.*, 2008). However, as noted above, response variability in the current study seems to have scaled with the perceived range of distances in the stimulus set, which suggests that memory noise, rather than perceptual sensitivity, is the limiting factor in the current results (see Durlach and Braida, 1969; Braida et al., 1984). For this reason, the remaining analysis focused on the degree to which mean distance judgments can be explained by changes in the mean values of D/R and ILD.

A. Direct-to-reverberant power ratio

1. Results

Figure 3 plots the across-subject average D/R as a function of the source distance, separately for frontal and lateral

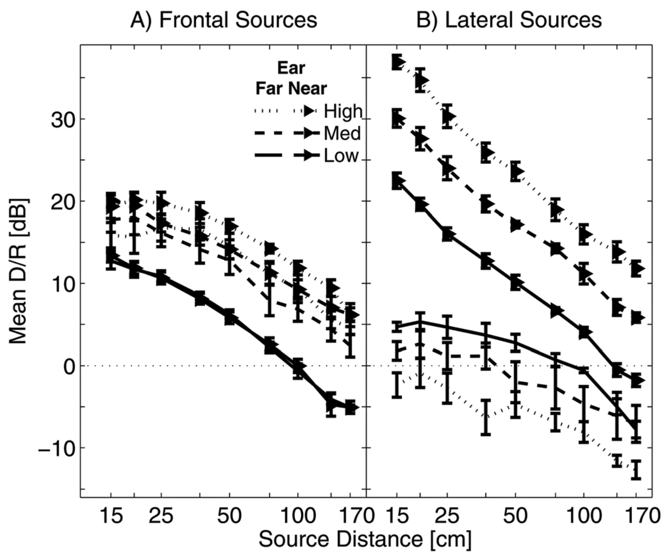

FIG. 3. Mean D/Rs for the narrowband stimuli used in this study, plotted separately for the far (left) and near (right) ears as a function of source distance. Across-subject averages (± SEM) in the measured D/Rs (averaged across the ten noise tokens) are plotted for each combination of stimulus spectral condition (line style) and direction [panel (A) vs (B)].

sources [panels (A) and (B), respectively] and for the near and far ears (▶ symbols mark the near-, right-ear data). The D/Rs were determined by separating the direct and reverberant portions of the individual BRIRs, convolving each portion of the BRIR with the stimulus, and then computing the difference in the total power contained in the two portions.

For both frontal and lateral sources, D/R decreased with increasing source distance. However, the mean values corresponding to a particular distance depended on source spectral content, source direction, and, for lateral sources, on whether the near- or far-ear signal was considered. For frontal sources, the far-ear and near-ear D/Rs were approximately equal, as expected, and only varied with spectral content [the pairs of lines of the same type, corresponding to the far and near ear D/Rs for a given frequency, lie near each other in Fig. 3(A)]. For lateral sources, the near-ear D/Rs always were larger than the far-ear D/Rs and decreased more with source distance than the far-ear D/Rs [in Fig. 3(B), the lines without symbols all fall below those with symbols]. For both frontal and lateral sources and for both far- and near-ear D/Rs, the D/R changed similarly with distance for all three frequency bands (in both panels, the D/Rs in a given ear are roughly parallel); however, for frontal sources the D/R variation was slightly larger at the lower frequencies (∼18 dB; solid line) than at the high frequencies (∼13 dB; dotted line). For frontal sources (regardless of ear) and for the near ear for lateral sources, the mean D/R was largest for the high-frequency bands (dotted line) and smallest for the low-frequency band (solid line). For frontal sources, the mean D/R varied over a range of approximately 15 dB across the measured distances [Fig. 3(A)]. For lateral sources, the near-ear D/Rs spanned roughly a 25 dB range, while the far-ear D/Rs span only a 12 dB range [Fig. 3(B)].

2. Interim summary and discussion

Results in Fig. 3 show that it is not possible to make accurate distance judgments using a fixed map from D/R to distance, independent of source direction and source spectrum. For example, if listeners heard two low-frequency sources with a D/R of 10 dB in the near ear and assigned them the same distance of about 50 cm, they would be relatively accurate if judging distance for a lateral source. However, they would overestimate the distance of a frontal source: a D/R of 10 dB for a frontal source would truly signify a distance of about 20 cm. Thus, if listeners judged distance by mapping the stimulus D/R to a response without taking into account source direction, they would be likely to either overestimate the distances of frontal sources and/or underestimate the distance of lateral sources. Similarly, if listeners heard two lateral sources with a D/R of 20 dB in the near ear and perceived them both at 75 cm, they would be relatively accurate for a narrowband source centered around 5.6 kHz; however, they would overestimate the distance of a narrowband source centered at 400 Hz. A D/R of 10 dB for this low-frequency source would truly signify a distance of about 20 cm. Thus, if listeners judged distance by mapping the stimulus D/R to a response without taking into account source spectral content, they would be likely to either overestimate the distances of low-frequency sources and/or underestimate the distance of high-frequency sources.

B. Interaural level difference

1. Results

The difference in the SPLs in 1/3-octave-filtered far- and near-ear signals (F_C's shown in Table I) was computed separately for the first 300-ms of each of the ten pre-generated noise tokens for each stimulus type, location, and subject (note that the 300 ms averaging window, which corresponds to the portion of the stimulus in which the direct sound energy was present, is comparable to the behavioral estimates of the integration window for ILD sensitivity; Hartmann and Constan, 2002).

The mean ILDs averaged across tokens are shown in Fig. 4. For lateral sources, ILD decreased with increasing distance for stimuli in all frequency bands (see lines without symbols in Fig. 4). ILDs for lateral sources had the largest magnitudes and spanned the largest range for high-frequency sources, and had the smallest magnitudes and spanned the smallest range for low-frequency sources (the dotted line without symbols falls above the solid line without symbols, with the dashed line without symbols falling in between). In contrast, as expected, ILDs were near zero for all frontal sources, regardless of spectral content.

2. Interim summary and discussion

Results in Fig. 4 show that it was not possible to make accurate distance judgments using a fixed mapping from ILD to distance, independent of source direction and source spectrum. ILD did not vary with distance for frontal sources, so it could not provide accurate distance cues for these sources. Moreover, the mapping from ILD to distance also varied with source spectrum for lateral sources. Specifically, if listeners heard two lateral sources with ILDs of 10 dB and responded that these sources were at 20 cm, they would have

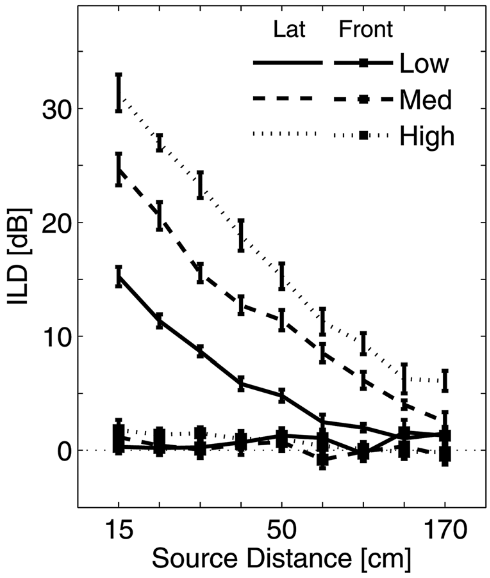

FIG. 4. Mean ILD for the narrowband stimuli as a function of source distance. Across-subject averages (± SEM) of the measured ILDs (averaged across the ten noise tokens used in each condition) are plotted separately for each combination of stimulus spectral condition and direction.

been relatively accurate for a source centered around 400 Hz; however, they would have underestimated the distance of a source centered at 3 kHz, since an ILD of 10 dB for this mid-frequency source would truly signify a distance of about 60 cm. Thus, if listeners judged distance by mapping the stimulus ILD to a response without taking into account source spectral content, they would tend to either overestimate the distances of low-frequency sources and/or underestimate the distance of high-frequency sources.

C. Relating acoustic distance cues to responses

1. Results

Figure 5 shows the mean response distances as a function of the individual cues extracted from the stimuli [near-ear D/R cue in panels (A)–(C); far-ear D/R in panels (D)–(F); ILD in panels (G)–(I)], separately for each stimulus type (left-most, central, and right-most columns for low, mid, and high frequencies, respectively) and source azimuth (solid vs dotted lines within each panel). The mean cue and response values shown in Figs. 2, 3, and 4 were combined to generate this figure. The inset in each panel shows the correlation coefficients between the distance-dependent mean values of the cues and the mean distance responses, computed separately for each subject and then averaged across the subjects.

The top row of panels in Fig. 5 shows the relationship between the near-ear D/R and the mean response distances. Within each panel (i.e., for a given source frequency), results

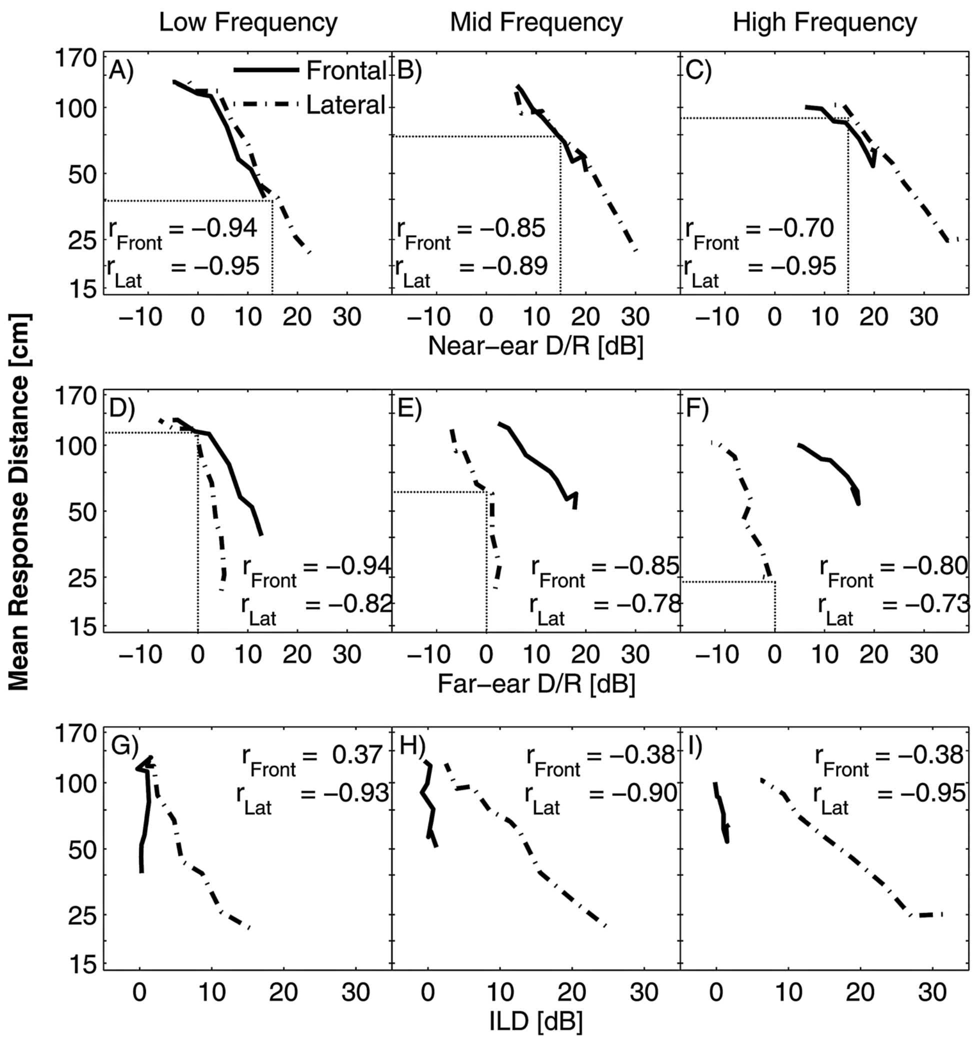

FIG. 5. Mean responses as a function of three different distance-dependent acoustic cues: the near-ear (right-ear) D/Rs [panels (A)–(C)], the far-ear (left-ear) D/Rs [panels (D)–(F)], and the ILD [panels (G)–(I)]. Each column of panels corresponds to a different stimulus frequency. The solid and dotted lines represent the frontal and the lateral data. The correlation coefficients inset in each panel show the across-subject mean of the correlation coefficient computed for each subject between the mean value of a given cue (at each distance) and the mean response (for the same distance).

for frontal and lateral sources fell along the same mapping [the solid and dashed lines lie approximately on top of each other in Figs. 5(A), 5(B), and 5(C)]. The mapping was approximately linear with a slope that was very similar for both frontal and lateral sources and across all three frequencies [the solid and dashed lines in Figs. 5(A), 5(B), and 5(C) are approximately parallel to each other across the three panels]. The only consistent difference across the panels was that the lines corresponding to the lower frequencies were offset downward relative to the lines corresponding to higher frequencies. For example, a D/R of 15 dB corresponds to a response distance of approximately 40 cm in panel (A), but to 80 cm in panel (C) (compare the thin dotted lines across panels). The correlations between the responses and the near-ear D/Rs were all relatively high [see the r_{lat} and r_{front} values in Figs. 5(A)–5(C)]. However, they were larger for lateral sources, for which both the cue values and the responses covered larger ranges, than for frontal sources, for which the ranges were smaller. Also, for frontal sources, the correlations were larger for the low-frequency stimuli than for the high-frequency stimuli, an effect that may be due to the decreased range of near-ear D/Rs in the high-frequency stimuli and/or to a slight change in the slope of the mapping from the cue to the response. Specifically, averaged across the near and far ears, the slopes of the linear fit to the individual subject low- and high-frequency data were -0.0667 log(cm)/dB and -0.0416 log(cm)/dB, respectively (this difference in the slopes is weakly significant; student's *t*-test, $t_4 = 2.93, p < 0.05$).

The middle row of panels in Fig. 5 shows the relationship between the far-ear D/R and the distance responses. All mappings were again approximately linear. For frontal sources, the mappings from D/R to response as well as the correlations were similar to the near-ear D/Rs, as expected, given that the two ears received similar signals [solid lines and the r_{front} values in Figs. 5(D)–5(F) are similar to the respective lines and values in Figs. 5(A)–5(C)]. However, the lateral-source and frontal-source results did not overlap for sources with the same frequency content. Although the slopes of the best-fit lines relating far-ear D/R to response distance were comparable across the three panels, their offsets differed greatly. Specifically, the low-frequency graphs are offset upward relative to the high-frequency graphs [compare the thick dotted lines in panels (D), (E), and (F); the offsets are marked by the thin dotted lines]. Finally, for the lateral sources, the correlations were much smaller for the far-ear D/Rs than for the near-ear D/Rs [compare the values of r_{lat} between the respective panels of Figs. 5(D)–5(F) and Figs. 5(A)–5(C)].

The bottom row of panels in Fig. 5 shows the relationship between the ILD and the distance responses. For frontal sources, the ILDs provided no distance information [the solid lines are nearly vertical and the r_{front} values very low in Figs. 5(G)–5(I)]. However, there was a strong linear relationship between the ILD cue and the responses for lateral sources (the correlations between ILD and response distances, r_{lat}, were as large as the correlations between the near-ear D/Rs and response distances). The relationship between ILD and distance for lateral sources was frequency dependent: the slopes of the lateral-source graphs were steeper at low frequencies than at high frequencies [compare the dotted lines in Figs. 5(G) and 5(I); this change in slope with frequency was highly significant; student's t-test, $t_4 = 5.51, p < 0.01$].

2. Interim summary and discussion

The near-ear D/Rs in Fig. 5 can account for most of the observed trends in how listeners judge source distance. Specifically, performance can be explained by assuming that listeners map the near-ear D/R to perceived distance if there is a D/R-to-distance map that is frequency specific, but independent of source direction. Such a strategy accounts for biases in the distance judgments for frontal sources: for these frontal sources, the D/R-to-distance map that is accurate for lateral sources causes systematic overestimation of distance for frontal sources, particularly those close to the listener, consistent with what we observed.

Figure 5 also shows that the far-ear D/R cue provides some distance information in all conditions. However, the far-ear D/Rs always change less with the source distance than the corresponding near-ear D/Rs; thus, the far-ear D/R conveys less distance information than the near-ear D/R. Moreover, the far-ear D/R cannot explain the biases observed for frontal sources in the same simple way that the near-ear D/R can: the perceptual mapping based on the far-ear-D/R cue would have to be both direction- and frequency-dependent to account for results (let alone to provide accurate judgments). Thus, given the availability of the near-ear D/R, it is parsimonious to conclude that the listeners did not use the far-ear D/R when judging distance.

Finally, Fig. 5 shows that the ILD cues may provide distance information for lateral sources, even though they provide no information about frontal sources. However, just as with the far-ear D/R, the ILD provides no simple, direction-independent mapping from cue to response distance. Indeed, for high frequency sources, the near-zero ILDs that arise for frontal sources are never observed for lateral sources, even though the distance responses for sources from the two directions overlap. Thus, ILDs do not help explain the biases in the judgments for the nearby frontal sources, whereas both response biases and the correlations between simulated and responded distance are fully explained by assuming a simple direction-independent but frequency-dependent mapping between the near-ear D/R and response. Of course, while these results are consistent with listeners essentially ignoring all distance cues except the near-ear D/R, we cannot conclude that judgments are not influenced by ILDs or other cues. Further experiments are needed to test the idea that, in the presence of strong reverberation-related cues, listeners give little perceptual weight to other potential cues to source distance, including the ILD distance cues that have been shown to matter in anechoic space.

V. GENERAL DISCUSSION

The current study showed that, in reverberation, distance perception for nearby sources roved in level depends on the stimulus spectral content. The dependence was strong for frontal sources, with correlation coefficients ranging from 0.3 to 0.75, and much weaker for lateral sources, with correlations ranging from 0.65 to 0.85. The spectral characteristic that most strongly influenced performance was the lowest frequency present in the stimuli: distances of stimuli that contained energy at frequencies around 300 Hz were judged relatively accurately for both frontal and lateral directions, while for stimuli that only contained energy at 5.7 kHz, distance perception was less accurate, particularly for sources in front of the listener. While the presence of low-frequency energy influenced distance perception, stimulus bandwidth did not.

A detailed analysis of the listener responses showed that mean responses could account for the changes in the correlations between simulated and response distance with changes in source direction and spectral content (changes in response variances could not explain these changes). Specifically, for all stimuli, listeners overestimated the distance of nearby frontal sources. This response bias for nearby frontal sources was most pronounced for stimuli containing only high frequencies. In contrast, listeners responded relatively accurately for stimuli coming from distant frontal locations and for all stimuli simulated from the side.

A comparison of the changes in the mean responses to the distance-dependent changes in the candidate acoustic cues found a simple relationship between the near-ear D/R and the mean distance judgments for the stimuli and conditions examined in the current study. Specifically, a direction-independent mapping from the near-ear D/R to the perceived distance with a frequency-dependent offset could explain these results.

The analysis also found that changes in the ILD cue were strongly correlated with mean response distances for lateral sources, but not for frontal sources where ILDs are near zero. Thus, it is possible that either the ILD contributed to the distance judgments for lateral sources or that the lateral-source judgments were based solely on the ILD cue. Past studies show that ILD is used for distance judgments in anechoic space, particularly for low-frequency sources (Brungart, 1999). In the current study, the ILD cue varied most strongly with distance for high-frequency sources. Coincidentally, D/R sensitivity is lower for high-frequency than for low-frequency sources (Larsen *et al.*, 2008). In addition, D/R varied less for the current high-frequency stimuli than it did for low-frequency stimuli. Based on these facts, if ILD did contribute to distance judgments in the current study, its effect was likely to be greatest for lateral, high-frequency sources, where ILD could convey distance information but information in D/R was relatively weak.

To further explore whether the ILD cue contributed to judgments here, the correlation coefficients for lateral sources in the current study were compared to correlations from a comparable anechoic experiment (Brungart, 1999), where ILDs were the main distance cue available. Specifically, the current broadband (BB), wideband low (WL), and wideband medium (WM) stimuli are grossly similar to the stimuli used in that study. Consistent with the current results, Brungart observed performance that became worse as the low-frequency stimulus cutoff increased (Brungart's r's are 0.83, 0.87, and 0.58 for the stimuli corresponding to the current BB, WL, and WM stimuli, respectively). However, performance in this anechoic study dropped off more steeply with increasing low-frequency cutoff than we found in the current study. Specifically, the anechoic and reverberant correlations in the two studies are comparable for the conditions corresponding to BB and WL; however, for the WM stimuli, the correlation coefficient of 0.58 in the anechoic study is considerably lower than the 0.72 correlation of the current reverberant study. The fact that our results are comparable to the results in anechoic space for BB and WL conditions, but better in the WM condition, suggest that D/R contributes to distance judgments even at higher frequencies, where D/R information is relatively weak.

Even though the current results cannot rule out that listeners use ILD information when judging the distance of lateral sources, there is no direct evidence that ILD cues affected distance judgments in the current study. Instead, we found that lateral and frontal sources of a given frequency with the same D/R were *incorrectly* assigned the same response distance. These results are consistent with the hypothesis that listeners used the D/R to estimate distance the same way both for sources from the side, where ILD cues are available to the listener, and for frontal sources, where ILD cues are negligible. Thus, although listeners may combine ILD and D/R information when judging distance of nearby sources in reverberant space, the current data are parsimoniously explained by assuming they use only D/R.

In general, however, other mechanisms could also be envisioned. For instance, listeners could optimally combine D/R and ILD information. With such a scheme, distance judgments for lateral sources, for which both cues are informative, would be more precise than for the frontal sources, for which only D/R provides information. Of course, the observed bias, whereby nearby frontal sources are judged as more distant than their simulated distance, cannot be directly explained with such a mechanism. This kind of bias, however, could be due to other factors, like a visually induced bias related to the fact that no sound source is visible in front of the listener (Gardner, 1968).

The current study analyzed D/R, showing that this physical attribute of the stimuli correlated well with distance judgments in a given frequency. However, it is very unlikely that the D/R is directly computed by the auditory system. In order to compute D/R directly, listeners would have to separate the direct sound from the reflected sound in the total signal, which is an under-constrained computational problem. Indeed, listeners cannot even effectively separate spectral cues for elevation from source spectral content to judge elevation, which is arguably a simpler problem than that of separating direct from reverberant sound (Rakerd *et al.*, 1999).

Although listeners may not be able to separate direct from reverberant sound explicitly, there are a number of physical characteristics of the received signal that co-vary with D/R, and that listeners might be able to estimate to create a reverberation-based distance cue. Three such cues that have been proposed in previous studies are the early-to-late power ratio (Bronkhorst and Houtgast, 1999), the interaural coherence (Bronkhorst, 2001), and monaural changes in the spectral centroid or in frequency-to-frequency variability in the signal (Larsen *et al.*, 2008). These cues are next considered in the context of the current study.

In general, as D/R decreases, the interaural coherence decreases, providing a simple-to-compute cue that co-varies with D/R. However, the interaural coherence in the stimuli presented to our subjects cannot, by itself, predict observed behavior. In our stimuli, the interaural coherence for frontal sources was both greater and varied more with source distance than the interaural coherence for lateral sources (data not shown). This pattern suggests that distance performance based only on interaural coherence should be better for frontal sources than for lateral sources, rather than the other way around, as we observed. In addition, binaural and monaural D/R sensitivities are similar, suggesting that binaural cues play no role in D/R perception (Larsen, 2008; however, at least one past study did find a relationship between interaural coherence and distance judgments; see Bronkhorst, 2001).

The spectral centroid changes with D/R, as well. However, Larsen showed that bandwidth strongly affects the reliability of such monaural spectral cues. If listeners used monaural spectral cues to judge distance in the current study, we should have seen performance deteriorate with decreasing bandwidth, yet we found that stimulus bandwidth had little effect on performance.

Thus, of the three previously considered cues that vary with D/R, this leaves the early-to-late power ratio as the most likely candidate to explain how performance varies with D/R. However, other cues (such as short-term fluctuations in binaural cues, Goupell and Hartmann, 2007, or monaural cues like the change in the stimulus autocorrelation or modulation structure) could also play a role (also see footnote 3). Future studies are needed to evaluate which, if any, of these physical cues best accounts for these results.

Past distance studies came to different conclusions about what acoustic cues affect listener judgments. Given that these studies differed in the specific conditions tested (e.g., availability of overall level as a cue, choice of spatial region, etc.), these apparent discrepancies are most likely due to the fact that listeners weight different potential distance cues differently, depending on the listening situation. Such fluid strategies could, for instance, allow listeners in anechoic space to use ILD cues to judge source distance, but to focus on sound features that vary with D/R to estimate source distance in reverberant space. Such computational flexibility makes it difficult to formulate a general model of distance perception without further studies to tease apart how the characteristics of the environment influence distance judgments. Therefore it is not surprising that, for nearby sources roving in level and presented in a simulated reverberant setting, the near-ear D/R cue alone can predict all major trends in how performance depends on source direction and stimulus spectral content.

The results presented here are based on individualized simulations of spatial cues using a fairly small directional sound source in a small, moderately reverberant classroom. Although our results may not predict what happens for sources far from the listener, where D/R will be lower than what we tested, it is still likely that they will generalize to localization of nearby sources in other rooms. For nearby sources, the D/R is likely to be very large even in rooms whose reverberation levels are greater than or whose volumes differ from the volume of the classroom simulated here. Moreover, the directionality of the loudspeaker used here is similar to that of many real sound sources, including humans producing speech (see Shinn-Cunningham *et al.*, 2005). Of course, for less reverberant rooms, the acoustic cues would become more like those in an anechoic chamber, in which the ILD cue is very likely to be the dominant cue (Brungart, 1999).

VI. CONCLUSIONS

In simulated reverberant space, listeners are able to judge source distance; however, their judgments vary with source direction. For lateral sources, judgments are relatively accurate and strongly correlated with simulated distance. For frontal sources, judgments of the distance of nearby sources are farther away than simulated distances, compressing the range of responses. This compression results can explain the observed reduction in the correlation between simulated distance and responses found for frontal sources. Analyses of acoustic cues that vary with distance show that both the near-ear D/R and the ILD change with distance and provide potential cues for source distance. However, for these reverberant results, the pattern of responses can be explained fully by assuming that listeners judged distance using a fixed D/R-to-distance mapping that varied with frequency. Future experiments are necessary to determine whether listeners only use ILD distance cues when there are no strong reverberation-related distance cues present (e.g., in anechoic space), or whether ILDs have some influence on distance judgments even when stimuli contain reverberant energy.

ACKNOWLEDGMENTS

This work was supported by the NSF Science of Learning Center program (CELEST Grant No. SBE-0354378). N.K. was supported by Grant No. NIH R03 TW007640 (to BGSC), and by the European Community under a Marie Curie International Research Staff Exchange Scheme (Grant No. 247543). Matt Schoolmaster helped with preparation of the experimental code and with data collection.

APPENDIX: OVERALL STIMULUS LEVEL AS A DISTANCE CUE

Despite the clear importance of the overall level cue for distance perception (Ashmead *et al.*, 1990; Litovsky and Clifton, 1992), it is likely that in everyday conditions listeners can ignore stimulus level when judging distance. For example, good performance was observed in a past anechoic study in which overall intensity was roved (Brungart, 1999). However, in contrast to past work, the current study was performed in a simulated reverberant space rather than being performed in an anechoic chamber or a real reverberant space. Moreover, the current study used randomly selected noise tokens, while some previous studies used frozen noise (e.g., Larsen *et al.*, 2008). As a result, listeners may have difficulty ignoring changes in loudness when judging distance in the current task. Here, the influence of stimulus level is analyzed explicitly to verify that listeners could ignore overall level when instructed that it was uninformative.

All stimuli used in this study were normalized to a constant SPL at the near ear and then roved by $\pm$ 5 dB to minimize the usefulness of overall level as a distance cue. To test whether the subjects failed to ignore this cue, even though they were informed it was not useful, the correlation coefficient between the presentation levels at the near ear and the response distances was computed for each run.[3]

Correlations between stimulus level and the distance response were near zero (across-subject averages fell between -0.2 and $+0.2$) for all listeners except subject S6, who had the worst overall ability to judge stimulus distance (circles in Fig. 1). This subject's correlations with presentation level had consistently large negative values (up to -0.7), showing that the subject interpreted louder sounds as being near and quieter sounds as being far (as mentioned in the main text, subject S6 was excluded from all across-subject analyses for this reason). On the other hand, the weak correlations between the presentation level and perceived distance response for the remaining subjects indicate that they were able to ignore overall level even in the conditions in which they could not judge the source distance accurately using whatever cues they were attending (for example, in the frontal NH condition; see Fig. 1).

[1]The D/R is often measured to quantify the acoustic effect of reverberation on sounds in a room (Zahorik, 2002b). Given that D/R is inversely proportional to the square of source distance, it is often talked of as a primary reverberation-related distance cue. However, it is unlikely that the brain actually can compute D/R directly. Instead, several candidate characteristics that correlate with the D/R have been proposed (see Sec. V).

[2]Note that since the overall dB SPL was fixed, the spectrum level of the narrowband stimuli was higher than the spectrum level of the broadband stimuli.

[3]Since the overall RMS level was normalized, the average levels of the direct sound and of the reverberant sound changed systematically with distance. Thus, if subjects could extract these characteristics from the stimuli, any of these cues could also provide some information about the source distance (see also footnote 1).

Ashmead, D. H., Leroy, D., and Odom, R. D. (**1990**). "Perception of the relative distances of nearby sound sources," Percept. Psychophys. **47**, 326–331.

Blauert, J. (**1996**). *Spatial Hearing* (MIT Press, Cambridge, MA), pp. 93–137.

Braida, L. D., Lim, J. S., Berliner, J. E., Durlach, N. I., Rabinowitz, W. M., and Purks, S. R. (**1984**). "Intensity perception. XIII. Perceptual anchor model of context-coding," J. Acoust. Soc. Am. **76**, 722–731.

Bronkhorst, A. W. (**2001**). "Effect of stimulus properties on auditory distance perception in rooms," in *Physiological and Psychological Bases of Auditory Function*, edited by D. J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht, The Netherlands), pp. 184–191.

Bronkhorst, A. W., and Houtgast, T. (**1999**). "Auditory distance perception in rooms," Nature **397**, 517–520.

Brown, C. (**2002**). "T60 Matlab function."

Brungart, D. S. (**1999**). "Auditory localization of nearby sources III: Stimulus effects," J. Acoust. Soc. Am. **106**, 3589–3602.

Carlile, S. (**1996**). *Virtual Auditory Space: Generation and Applications* (RG Landes, New York), pp. 109–142.

Durlach, N. I., and Braida, L. D. (**1969**). "Intensity perception. I. Preliminary theory of intensity resolution," J. Acoust. Soc. Am. **46**, 372–383.

Gardner, M. B. (**1968**). "Proximity image effect in sound localization (L)," J. Acoust. Soc. Am. **43**, 163.

Gescheider, G. A. (**1988**). "Psychophysical scaling," Annu. Rev. Psychol. **39**, 169–200.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notchednoise data," Hear. Res. **47**, 103–147.

Goupell, M. J., and Hartmann, W. M. (**2007**). "Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models," J. Acoust. Soc. Am. **122**, 1029–1045.

Hartmann, W. M. (**1983**). "Localization of sound in rooms," J. Acoust. Soc. Am. **74**, 1380–1391.

Hartmann, W. M., and Constan, Z. A. (**2002**). "Interaural level differences and the level-meter model," J. Acoust. Soc. Am. **112**, 1037–1045.

Herron, T. (**2005**). "C language exploratory analysis of variance with enhancements," http://www.ebire.org/hcnlab/software/cleave.html (Last viewed Janunary 30, 2005).

Larsen, E., Iyer, N., Lansing, C. R., and Feng, A. S. (**2008**). "On the minimum audible difference in direct-to-reverberant energy ratio," J. Acoust. Soc. Am. **124**, 450–461.

Litovsky, R. Y., and Clifton, R. K. (**1992**). "Use of sound-pressure level in auditory distance discrimination by 6-month-old infants and adults," J. Acoust. Soc. Am. **92**, 794–802.

Mershon, D. H., and King, L. E. (**1975**). "Intensity and reverberation as factors in the auditory perception of egocentric distance," Percept. Psychophys. **18**, 409–415.

Parducci, A., and Wedell, D. H. (**1986**). "The category effect with rating-scales-number of categories, number of stimuli, and method of presentation," J. Exp. Psychol. Human Percept. Perform. **12**, 496–516.

Parker, S., Murphy, D. R., and Schneider, B. A. (**2002**). "Top-down gain control in the auditory system: evidence from identification and discrimination experiments," Percept. Psychophys. **64**, 598–615.

Petrov, A. A., and Anderson, J. R. (**2005**). "The dynamics of scaling: a memory-based anchor model of category rating and absolute identification," Psychol. Rev. **112**, 383–416.

Rakerd, B., Hartmann, W. M., and McCaskey, T. L. (**1999**). "Identification and localization of sound sources in the median sagittal plane," J. Acoust. Soc. Am. **106**, 2812–2820.

Santarelli, S. (**2000**). "Auditory localization of nearby sources in anechoic and reverberant environments," Ph.D. thesis, Cognitive and Neural Systems Departmentment, Boston University, Boston, MA, pp. 142–187.

Santarelli, S., Kopco, N., and Shinn-Cunningham, B. G. (**1999**). "Localization of near-field sources in a reverberant room," in *22nd mid-Winter meeting of the Association for Research in Otolaryngology* (St. Petersburg Beach, FL), p. 26.

Schroeder, M. R. (**1965**). "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**, 409–412.

Shinn-Cunningham, B. G., Kopco, N., and Martin, T. J. (**2005**). "Localizing nearby sound sources in a classroom: Binaural room impulse resonses," J. Acoust. Soc. Am. **117**, 3100–3115.

Shinn-Cunningham, B. G., Santarelli, S., and Kopco, N. (**2000**). "Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural vs. monaural cues," in *23rd mid-Winter meeting of the Association for Research in Otolaryngology* (St. Petersburg Beach, FL), p. 88.

Ury, H. K., and Wiggins, A. D. (**1971**). "Large sample and other multiple comparisons among means," Br. J. Math. Stat. Psychol. **24**, 174–194.

Zahorik, P. (**2002a**). "Assessing auditory distance perception using virtual acoustics," J. Acoust. Soc. Am. **111**, 1832–1846.

Zahorik, P. (**2002b**). "Direct-to-reverberant energy ratio sensitivity," J. Acoust. Soc. Am. **112**, 2110–2117.

Zahorik, P., Brungart, D. S., and Bronkhorst, A. W. (**2005**). "Auditory distance perception in humans: A summary of past and present research," Acta. Acust. Acust. **91**, 409–420.