

Localizing nearby sound sources in a classroom: Binaural room impulse responses^{a)}

Barbara G. Shinn-Cunningham^{b)}

Boston University Hearing Research Center and Departments of Cognitive and Neural Systems and Biomedical Engineering, 677 Beacon Street, Boston, Massachusetts 02215

Norbert Kopco

Boston University Hearing Research Center and Department of Cognitive and Neural Systems, 677 Beacon Street, Boston, Massachusetts 02215

Tara J. Martin

Boston University Hearing Research Center, 677 Beacon Street, Boston, Massachusetts 02215

(Received 15 December 2003; revised 26 January 2005; accepted 27 January 2005)

Binaural room impulse responses (BRIRs) were measured in a classroom for sources at different azimuths and distances (up to 1 m) relative to a manikin located in four positions in a classroom. When the listener is far from all walls, reverberant energy distorts signal magnitude and phase independently at each frequency, altering monaural spectral cues, interaural phase differences, and interaural level differences. For the tested conditions, systematic distortion (comb-filtering) from an early intense reflection is only evident when a listener is very close to a wall, and then only in the ear facing the wall. Especially for a nearby source, interaural cues grow less reliable with increasing source laterality and monaural spectral cues are less reliable in the ear farther from the sound source. Reverberation reduces the magnitude of interaural level differences at all frequencies; however, the direct-sound interaural time difference can still be recovered from the BRIRs measured in these experiments. Results suggest that bias and variability in sound localization behavior may vary systematically with listener location in a room as well as source location relative to the listener, even for nearby sources where there is relatively little reverberant energy. © 2005 Acoustical Society of America. [DOI: 10.1121/1.1872572]

PACS numbers: 43.66.Qp, 43.66.Pn, 43.55.-n, 43.55.Br [AK]

Pages: 3100–3115

I. INTRODUCTION

Spatial acoustic cues are important for many tasks, ranging from locating a sound source (e.g., see Middlebrooks and Green, 1991) to detecting and understanding one source in the presence of competing sources from other locations (e.g., see Bronkhorst, 2000; Ebata, 2003). A great deal of research effort in the field of psychoacoustics has gone into measuring and analyzing head-related impulse responses (HRIRs; the impulse response from source to the listener's ears for sources presented in anechoic space) to gain insight into the acoustic cues underlying these important behavioral functions (e.g., see Wightman and Kistler, 1989; Shaw, 1997; Algazi *et al.*, 1999; Brungart and Rabinowitz, 1999; Kulkarni and Colburn, 2004) and to allow simulation of realistic binaural signals (e.g., Middlebrooks, 1999; Begault *et al.*, 2001; Macpherson and Middlebrooks, 2002; Culling *et al.*, 2003; Kidd *et al.*, 2005).

The reverberation present in everyday settings influences auditory perception and behavior in both positive and negative ways. Reverberation can provide listeners with a cue for sound source distance (e.g., see Mershon *et al.*, 1989; Naguib, 1995; Bronkhorst and Houtgast, 1999; Shinn-Cunningham, 2000a; Zahorik, 2002a, b), allow listeners to

estimate properties of the environment (e.g., see Bradley and Soudre, 1995; Bech, 1998b, a; Torres *et al.*, 2001; Shinn-Cunningham and Ram, 2003), and improve the subjective realism and externalization of virtual auditory space simulations (e.g., see Durlach *et al.*, 1992; Begault *et al.*, 2001). Reverberant energy also influences speech intelligibility; early reflections (occurring within 50–80 ms of the direct sound) generally increase the audibility of a speech source without degrading intelligibility (e.g., see Bradley *et al.*, 1999, 2003), whereas later reflections smear out the temporal information in the speech waveform and decrease intelligibility (e.g., see Bradley, 1986; Nabelek *et al.*, 1989; Bradley *et al.*, 1999; Bistafa and Bradley, 2000). Both early and late reflections cause interaural decorrelation that can degrade the accuracy of directional localization (Hartmann, 1983; Hartmann and Rakerd, 1999; Hartmann *et al.*, 1999; Shinn-Cunningham, 2000b; Kopco and Shinn-Cunningham, 2002) and the ability to detect and understand one source in the presence of a statistically stationary competing sound source from another location (Plomp, 1976; Zurek, 1993; Culling *et al.*, 1994; Shinn-Cunningham, 2002; Devore and Shinn-Cunningham, 2003; Zurek *et al.*, 2004). Similarly, early and late reflections are known to have a large impact on the impressions of source width and envelopment (e.g., see Barron, 2001). Despite the importance of reverberant energy on nearly all aspects of auditory perception, few studies have

^{a)}Portions of this work were presented at the 2001 Mid-Winter Meeting of the Association for Research in Otolaryngology.

^{b)}Electronic mail: shinn@cns.bu.edu

analyzed how reverberant energy affects the spatial acoustic cues reaching a listener.

In the field of architectural acoustics, there have been a number of studies of how auditoria and other large echoic spaces influence various aspects of the signals reaching a listener's ears (e.g., see Kleiner *et al.*, 1993; Bradley, 1996; Hidaka and Beranek, 2000; Nishihara *et al.*, 2001; Torres *et al.*, 2001; Okano, 2002). Similarly, there are some studies of the behavioral impact of reverberation (e.g., see Berkley, 1980; Bradley, 1986; Nabelek *et al.*, 1989; Culling *et al.*, 1994; Hartmann and Rakerd, 1999; Hodgson, 1999; Darwin and Hukin, 2000; Culling *et al.*, 2003). Statistics comparing the relative energy in the early and late portions of room impulse responses (e.g., the clarity index C_{80}) summarize how reverberant energy affects the subjective experience of listening to music or understanding speech in a particular space (e.g., see Hidaka and Beranek, 2000). However, there are relatively few studies of the effects of room reflections on spatial and binaural acoustic cues or how these effects vary with listener location in the room (although see de Vries *et al.*, 2001; Hartmann *et al.*, 2005). The current study analyzes some of the properties of acoustic spatial cues present in binaural room impulse responses (BRIRs; the impulse responses from a sound source to the ears of a listener located in a room) measured in an ordinary classroom. The goal of this study is to begin to quantify how acoustic spatial cues are affected by reverberant energy as a function of the listener's location in a room and the source location relative to the listener.

Most previous studies of HRIRs have focused on measurements made at distances of a meter or more, where the only effect of source distance is an overall change in gain. For these cases, interaural time differences (ITDs) and interaural level differences (ILDs) in the signals reaching the ears resolve source location to within a "cone of confusion." However, for nearby sources (within a meter of the listener), interaural level differences increase and ILDs provide distance information (Duda and Martens, 1998; Brungart and Rabinowitz, 1999). For nearby sources, interaural differences resolve source location to within a "torus of confusion" (Shinn-Cunningham *et al.*, 2000) and spectral content can resolve source position within a torus of confusion (i.e., in the up/down, front/back dimensions; Asano *et al.*, 1990; Butler and Humanski, 1992; Wightman and Kistler, 1997; Vliegen and van Opstal, 2004). Because the acoustics of such situations differ from those normally studied, sound localization, source detection, and signal understanding differ when sources are nearby compared to the more commonly studied "distant source" conditions (e.g., see Brungart and Durlach, 1999; Shinn-Cunningham, 2000b; Shinn-Cunningham *et al.*, 2001; Shinn-Cunningham, 2002; Kopco and Shinn-Cunningham, 2003). Given that many important, everyday events (e.g., a personal conversation) involve sources relatively near the listener, it is important to understand what spatial acoustic cues arise for nearby sources in reverberant environments.

For nearby sources, small changes in the source location relative to the listener can cause considerably larger changes in the direct-sound energy reaching the ears than for more

distant sources and the direct-sound level at the two ears can be very different (e.g., see Brungart and Rabinowitz, 1999; Shinn-Cunningham *et al.*, 2000; Kopco and Shinn-Cunningham, 2003). As a result, the interaction between source location and the effects of reverberation will be maximized when sources are near the listener, even though the overall influence of reverberant energy will be smaller than for more distant sources. Furthermore, analysis of how the effect of room reflections varies with source location for nearby sources can provide insights into what will happen for more distant sources, given that the main effect of increasing source distance beyond one meter is to reduce the direct sound level and increase the relative strength of reverberant energy at the ears.

The current analyses examine how listener location in a room and source location of nearby sources influence BRIRs in one example classroom. The specific room, source locations relative to the listener, and listener locations considered were chosen for two reasons: to (1) explore how the acoustics within a specific room can vary as the listener is moved within the room (e.g., ranging from a listener positioned far from any reflecting surfaces to the most extreme situation, where the listener is seated in the corner of the room) and (2) gain insight into results of related behavioral studies examining the effects of reverberation on perception for sources near a listener that were performed in or simulated the same space analyzed here.¹ While future papers will address how the acoustics of the classroom influence perception and behavior, the current analysis focuses on the effects reverberation can have on different aspects of the spatial cues present in BRIRs. The long-term goal of these efforts is to tease apart which aspects of the perceptual consequences of room reverberation are predicted directly from acoustic effects and which arise from interactions between properties of the signals at the ears and auditory processing.

The following analyses show that in the frequency domain, anechoic head-related transfer functions (HRTFs) vary relatively smoothly with frequency in both phase and magnitude compared to the frequency-to-frequency variability of binaural room transfer functions (BRTFs), a fact that reflects the complex interactions between the direct sound and reflected energy that arise in a room. Whereas anechoic head-related impulse responses are nearly equal to the impulse responses obtained by minimum-phase reconstruction from the magnitude of the HRTF plus a delay related to the travel time to the ear (i.e., phase information in the HRTF is relatively unimportant for reproducing the proper HRIR except for an overall group delay; e.g., see Kistler and Wightman, 1991), the phase information in a BRTF is critical. In a room, each reflection may boost or reduce the energy in a particular frequency of the direct sound, depending upon the exact timing of the reflection and whether it adds in or out of phase with the preceding energy at a given frequency. In addition, if the frequency content of a sound source fluctuates with time, the interaction between direct and reflected energy in a room will also vary over time in ways that depend in detail on the spectro-temporal structure of the source signal and of the BRTF. This complexity is reflected in the frequency-to-

frequency fluctuations in the BRTF magnitude and phase, which can be quite large.

Given the relatively smoothly varying nature of anechoic HRTFs with frequency, small short-term spectral fluctuations in a broadband source signal (e.g., over the course of tens to hundreds of milliseconds, such as are commonly found in speech and music) have only modest effects on the spatial cues reaching a listener in anechoic space. Furthermore, analyzing the spatial cues in a frequency-smoothed anechoic HRTF (e.g., smoothing HRTFs over a critical band of frequencies) provides a reasonable summary of the spatial cues that an arbitrary HRTF-filtered source is likely to contain.

In contrast, and as a direct result of the spectro-temporal complexity of a typical BRIR, even small fluctuations in a sound source's frequency content with time can cause large temporal variations in the spatial cues observed in the BRTF-filtered signal. The exact manner in which the reflected energy in the BRIR will affect spatial cues depends very strongly on both the spectral and temporal properties of a stimulus. For instance, a pure tone signal that is turned on in a room will rapidly converge to a steady-state magnitude and phase value at each ear, with the magnitudes and phases depending only on the left- and right-ear BRTF magnitudes and phases at the tone frequency. Thus, the interaural differences caused by a pure tone in a room will rapidly converge to values that are constant over time. However, a signal with more bandwidth will contain fluctuations in its short-term spectrum. As the short-term signal frequency content varies over time, the way in which the short-term magnitude and phase spectra of the signals reaching each ear are affected by room reverberation will change and evolve dynamically, depending on the spectro-temporal content of the source (including the relative phases of different frequency components of the input) and how it interacts with the reflected energy. Because this interaction is different at the two ears, interaural parameters will also fluctuate over time. This complexity makes it impossible to predict what spatial cues will be present in a room unless the source is specified.

Rather than analyzing what spatial cues will be present in the signals reaching a listener's ears for a specific stimulus, the current analysis examines aspects of BRIRs/BRTFs that are related to the ways in which short-term spatial cues will be influenced by reverberant room energy. For instance, the size of frequency-to-frequency fluctuations in the BRTF magnitude and phase is related to how strongly the spatial cues reaching the ears of a listener in the room depend on the short-term spectro-temporal content of the source. Thus, frequency-to-frequency variability in BRTFs is related to the variability in spatial cues likely to be observed (e.g., at the output of a peripheral auditory nerve) across samples for independent tokens of finite-length broadband noise signals and across time for an ongoing broadband signal. Conversely, the expected value of different spatial cues in the signals reaching a listener's ears (averaged across independent signal tokens or across time) is closely related to properties of the BRTF averaged across frequency (e.g., across a critical band). Thus, in order to gain insight into how reverberant energy will tend to affect across-time variability in

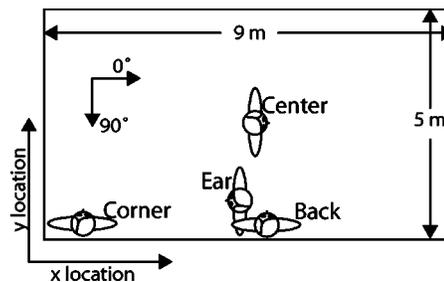


FIG. 1. Schematic showing the rough orientation and location of the KEMAR manikin in the classroom when measurements were made for the four listener locations (see Table I for more detailed descriptions of different listener locations).

spatial cues for broadband sources, we consider frequency-to-frequency variations in the BRTFs as a function of listener location and source position, a property that (as discussed below) depends primarily on the energy ratio of late reverberation to direct sound and early reflections in the BRTFs. In order to gain insight into how reverberant energy influences average spatial properties of the signals reaching the ears, we compare frequency-smoothed BRTFs to comparable anechoic transfer functions as a function of listener and source position in the room.

II. METHODS

BRIRs were measured on a Knowles Electronics Manikin for Acoustic Research (KEMAR) for 21 different source locations relative to KEMAR, consisting of all combinations of seven source azimuths (0° , 15° , 30° , 45° , 60° , 75° , and 90°) to the right of the listener and three source distances (0.15, 0.40, and 1 m). All measurements were taken with the sound source in the horizontal plane containing the ears [e.g., at zero elevation relative to the listener; see Brown (2000) for more detail]. Measurements were taken with KEMAR seated on a tall wooden stool with his ears approximately 1.5 m above the floor at four different locations in a classroom (shown schematically in Fig. 1 and described in detail in Table I), henceforth called *center*, *back*, *ear*, and *corner*. To isolate the influence of reverberant energy on acoustic spatial cues, *pseudo-anechoic* measurements were generated by time-windowing the *center* BRIRs² and analyzing only the portion corresponding to the direct sound impulse response (see below).

In order to estimate measurement reliability, all measurements (21 source locations \times 4 listener locations) were repeated on three separate occasions, with the equipment

TABLE I. Approximate orientation and location of the KEMAR manikin in the classroom when measurements were made for the four listener locations (see Fig. 1 for definition of orientation angle and x-y axes).

Listener location	Orientation	X location (m)	Y location (m)
Center/ Anechoic	0°	4.5	2.5
Back	-90°	4.5	0.5
Ear	180°	4.5	0.5
Corner	-90°	0.5	0.5

TABLE II. Number of usable measurement repetitions (of three measurements performed) obtained in each source-listener configuration. Bold numbers highlight configurations for which fewer than three repetitions were usable. (Note that fewer reliable measures were obtained in the *corner* configuration because measurements were repeated three times in the order *center*, *back*, *ear*, and *corner*, and one of the two microphones malfunctioned at the end of the final measurement sequence.)

Listener location	Source distance (m)	Source azimuth						
		0°	15°	30°	45°	60°	75°	90°
Center/ Anechoic	0.15	3	3	3	3	3	2	3
	0.40	3	3	3	3	3	3	3
	1	3	3	3	3	3	3	3
Back	0.15	3	3	3	3	3	3	3
	0.40	3	3	3	3	3	3	3
	1	3	3	3	3	3	3	3
Ear	0.15	3	3	3	3	3	3	3
	0.40	3	3	3	3	3	3	3
	1	3	3	3	3	3	3	3
Corner	0.15	3	3	2	2	3	3	3
	0.40	2	2	2	2	2	2	1
	1	2	2	2	2	2	2	2

taken down and reassembled in between measurement sessions. For 18 of the 84 source-listener configurations, technical problems rendered some of the measurements unusable; Table II details the number of useful repetitions available for each source-listener configuration.

The classroom dimensions were roughly 5×9×3.5 m. The room was carpeted and had hard concrete walls on three sides; the remaining (9-m-long) wall was constructed from a sound-absorptive partition. Acoustic tiles covered the ceiling. Few acoustically hard objects were in the room during the measurements (two small tables along the short wall and a collapsible ping-pong table, which was folded in half and oriented vertically, parallel to and near the long hard wall). The broadband T_{60} of the room was estimated from the measured BRIRs using the method formulated by Schroeder (Schroeder, 1965), as implemented by Brown in a Matlab function available at the Mathworks web site (Brown, 2002). For the *center* location, these estimates did not vary dramatically with source location or distance. The mean of the estimates in the *center* condition (across both ears, all source directions and all source distances) was 565 ms (standard deviation 24 ms). The means (and standard deviations) of the estimates of T_{60} from BRIRs were relatively robust with changes in listener location with mean (and standard deviation) values of 581 ms (28 ms), 585 ms (17 ms), and 619 ms (33 ms) for the *back*, *left*, and *corner* locations, respectively.

BRIRs were measured by concatenating two identical 32 767-long maximum length sequences (MLS; Rife and Vanderkooy, 1989; Vanderkooy, 1994) generated using a 44.1-kHz sampling rate. This MLS signal was sent to a Tucker-Davis Technologies (TDT) D/A converter (PD1), which drove a Crown amplifier connected to a Bose minicube loudspeaker. Prior to each measurement, the Bose loudspeaker was hand-positioned by the experimenter to be at the correct location (e.g., at the desired direction and distance relative to KEMAR, oriented to face the manikin). Miniature microphones (Knowles FG-3329c) mounted in earplugs and

inserted into the entrance of KEMAR's ear canals measured the raw acoustic responses to the MLS. Microphone outputs drove a custom-built microphone amplifier connected to a TDT A/D converter (TDT PD1). For each BRIR measurement, the MLS was presented and the response measured ten times. The ten time-domain measurements were then averaged to obtain the final MLS response.

In order to achieve the best possible signal-to-noise ratio for each measurement, the maximum sound source level that did not cause clipping was found by trial and error for each source position and listener location. Signals were then presented roughly 5 dB below the clipping level.

Offline, the average measured response to the MLS was cross-correlated with the original sequence to obtain a raw estimate of a 743-ms-long BRIR (see also Kopco and Shinn-Cunningham, 2003). Each raw BRIR was digitally filtered to remove energy below 100 Hz and above 20 kHz. Visual inspection of the raw BRIRs verified that the first-arriving reflection in every one of the *center* conditions (off the floor) arrives between 9.75 and 11 ms from the start of the measurement (no earlier than 5 ms after the direct sound reaches the ears), as predicted from geometrical calculations. Ten-ms-long, *pseudo-anechoic* BRIRs were generated by multiplying the *center* BRIRs by a time window that was flat (equal to 1.0) up to 9 ms, with a 1-ms-long cosine-squared fall time from 9 to 10 ms, effectively removing all reverberant energy from the *center* BRIRs.

The measured BRIRs contain not only the transfer characteristics of the head and room, but also characteristics of the sound delivery and measurement system. Calibration measurements were taken of the impulse responses to the left and right microphones in the center of the room (without a listener present, orienting the Bose loudspeaker to point towards the microphones) using the same procedures described above (time-windowing out all reverberant energy). Results showed that the system magnitude response is smooth and equal in the two microphones, with a gentle low-pass characteristic (decreasing smoothly by approximately 10 dB from 200 Hz to 20 kHz). Ideally, BRIRs would be postcompensated to remove the transfer characteristics of the sound delivery and recording system by inverse filtering. However, because (1) the distortion due to the measurement system was small, (2) for the main comparisons of interest in the current study (i.e., comparisons across source position re: KEMAR and listener location in the room), any such filtering would have no effect, and (3) such compensation could introduce additional errors into the estimated BRIRs (e.g., if the error in the estimated compensation filter was of the same magnitude or greater than the compensation itself), no postcompensation was used in the reported measurements.

Calibration measurements established that the useful dynamic range in the BRIRs was at least 40 dB for all frequencies from 200 Hz to 20 kHz and that the exponentially decreasing energy in each of the BRIRs was in the electrical/acoustical noise floor by 700 ms. In order to reduce extraneous noise in the measurements, the 0.1–20-kHz bandpass-filtered BRIRs were multiplied by a 743-ms-long time window with a 50-ms-long cosine-squared fall time

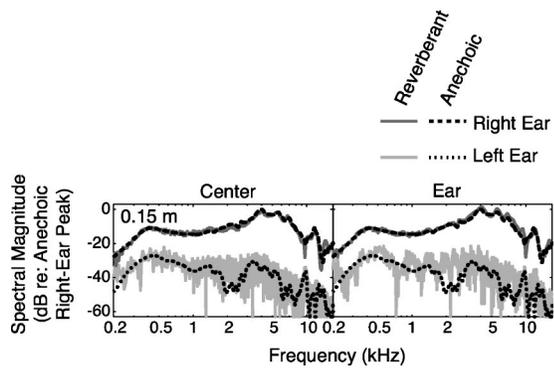


FIG. 2. Magnitude spectra of sample room transfer functions with the source at 90° azimuth, 1 m distance (*re*: KEMAR). Results for the right- and left-ear signals are shown in dark gray and light gray, respectively. Dashed lines show *pseudo-anechoic* results (derived from an independent set of *center* measurements). Results for the *center* and *ear* listener locations are shown in the left and right panels, respectively.

(from 693 to 743 ms) to produce the final BRIRs analyzed below.³

The current measurements quantify how source location relative to the listener and listener location in the room affect the spatial acoustic cues in BRIRs for an ordinary room. The study does not exhaustively explore what happens in arbitrary rooms, for arbitrary sound sources. For instance, because the Bose mini-cube speaker is not omnidirectional, it was always oriented to face the listener. Altering the orientation or model of the loudspeaker would alter both the direct sound level reaching the listener and the energy in and pattern of the reflections.⁴ Similarly, while there was nothing extraordinary about the classroom in which the measurements were taken, other rooms would yield different measurements. However, these results can be extrapolated to predict effects in other rooms by considering the statistics of reverberant energy in these other spaces. As described in Sec. I, because of the significant frequency-to-frequency variation in the BRTF phase and magnitude functions, the way in which reverberant energy distorts spatial cues for a particular source depends critically on the spectro-temporal content of the source. The current analyses explore how source location and listener location can alter different aspects of the BRIRs that relate to the magnitude of the effects of reverberant energy on the mean and variability in spatial cues likely to be observed for broadband sound sources. While in this sense the current results are specific (e.g., to the particular equipment, classroom, listener locations, etc.), the current results are similar to what would happen in these other settings. Finally, because the room measured here has been used in a number of behavioral studies, the specific details of this study may give insight into how reverberant energy influences perception in a range of tasks.

III. RESULTS

A. Example binaural room impulse responses

Figure 2 plots the magnitude spectra of BRIRs for a 0.15-m, 90° source for the *center* and *ear* conditions (left and right panels, respectively) to demonstrate the effects of diffuse reverberation and of an early, intense reflection. The two

panels show left- (light gray) and right-ear (dark gray) results from one of the repeated measurements as well as *pseudo-anechoic* results (black dashed lines) from a different *center* measurement. The spectral levels in each plot are normalized to the maximum peak in the right-ear *pseudo-anechoic* measurement.

Comparison of the right- and left-ear *pseudo-anechoic* results (black dashed lines) shows the large ILDs that arise for lateral sources very near the listener (see Brungart and Rabinowitz, 1999; Shinn-Cunningham *et al.*, 2000). The *pseudo-anechoic* results also show characteristic notches and peaks in the received spectral level, features important for signaling source location within a torus of confusion (e.g., see Butler and Belendiuk, 1977).

In the *center* conditions (left panel), the only obvious effects of the reverberant energy can be attributed to diffuse reverberant energy (adding frequency-to-frequency variability; filling in spectral notches; increasing the total energy in the shadowed, left ear). These effects increase as the direct-sound energy decreases, increasing in both ears with source distance, and increasing at the ear on the far side of the head as the source moves laterally.

In the *ear* results for the left-ear spectrum (gray plot in the right panel) there are systematic peaks and valleys consistent with frequency-dependent summation and cancellation of the initial direct sound and the prominent left-wall reflection. These notches are spaced by roughly 480 Hz, starting at 240 Hz, and the autocorrelation of the left-ear BRIR has a prominent peak at 2.1 ms (not shown), consistent with an intense early reflection off the left wall that reaches the left ear 2.1 ms after the direct sound. In all of the BRIRs, the exact timing of, intensity of, and interaural differences in any early reflection depend on source azimuth and distance, as predicted from geometrical considerations. For instance, as source azimuth increases, the delay between the direct sound and the left-ear reflection in the *ear* and *corner* conditions increases and the relative magnitude of the left-ear direct sound decreases. For all tested source locations, the left-ear spectra of the *ear* and *corner* conditions have pronounced comb filtering whose notch frequencies and notch depths vary with the relative timing and intensity of the early reflection. However, there is no pronounced comb filtering in any of the right-ear magnitude spectra for any of the source locations in the *ear* and *corner* conditions; similarly, there is no pronounced comb filtering in either ear in the *back* conditions, despite the proximity of the back wall.

B. Effects of reverberant energy on spectral magnitude

The FFT of each left- and right-ear impulse response was analyzed as a function of source and listener location. Summary statistics were computed to demonstrate how source distance, source direction, and listener location influenced the spectral magnitude of the signals reaching the listener.

1. Frequency-to-frequency fluctuations

Random, late-arriving reverberation tends to add frequency-to-frequency variability to the BRTF magnitude.

With such variability in the BRTF magnitude, the gain of the effective room filter acting at any point in time varies with the signal content at that time. Thus, whenever a signal has nonstationary spectro-temporal content, spatial cues in the signals at the ears will tend to vary, and the amount of fluctuation will be related to the variability in the BRTF from frequency to frequency. To characterize how such across-frequency variability depends on listener and source configuration in a room, the average size of the magnitude fluctuations in the BRTF (per Hz) were computed and then averaged across frequency. For each of the source locations (*re*: KEMAR) and listener locations in the room, this summary statistic was then averaged across the repeated measurements.

The BRTF can be thought of as a sum of direct sound, early echoes, and random late-arriving reverberation. Posed in this way, the random fluctuations in the BRTF magnitude as a function of frequency can be attributed primarily to reverberation. Statistical room acoustic analysis predicts that the mean spacing between adjacent maxima in the magnitude response of a room transfer function is approximately $3.91/T_{60}$ (e.g., see Schroeder and Kuttruff, 1962). Furthermore, the average dB change from a local maximum to an adjacent minimum should equal 10 dB (e.g., see Schroeder, 1987). For the current room with $T_{60}=565$ ms, statistical room acoustics analysis thus predicts that the late-arriving reverberation will have fluctuations of approximately 2.9 dB/Hz (i.e., 20 dB of fluctuation from peak to adjacent peak, which is, on average, 6.9 Hz away). Our measurements contain both direct sound and early reflections that vary relatively smoothly with frequency. As a result, the overall size of the fluctuations in the BRTF magnitude should vary between roughly 2.9 dB/Hz (if the reverberant energy dominates the total energy in the BRTF) down to almost no variation (when the reverberant energy is negligible compared to the sum of the energies in the direct sound and early reflections).

Variability in the right ear (not shown) was relatively small (at most less than 0.5 dB/frequency sample for sources straight ahead of the listener at 1 m), decreased slightly with source laterality, increased slightly with source distance, and was essentially identical for all listener locations in the room. The effect on the left ear, which generally received less direct sound energy than the right ear for all of the tested source locations, was greater overall and varied with source laterality, source distance, and listener location in the room.

Figure 3(a) plots the across-measurement mean of the absolute spectral magnitude change per Hz in the left-ear magnitude spectra and, where defined, the across-measurement standard deviation. Within each panel, results for each of the four listener locations in the room (dashed lines) and the *pseudo-anechoic* (solid line) results are shown as a function of source azimuth. Each row gives results for a different source distance.

Fluctuations in the *pseudo-anechoic* results [solid black lines in Fig. 3(a)] are essentially zero and independent of source direction and distance, as predicted. Variations in left-ear spectral energy increase with source azimuth for all listener locations in the room [in Fig. 3(a), values increase from

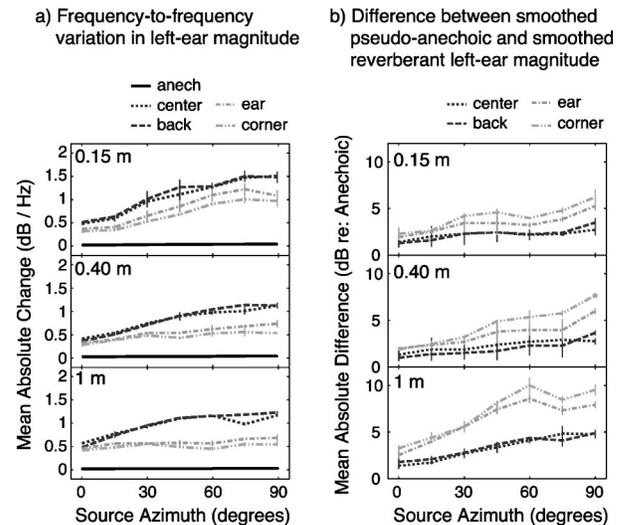


FIG. 3. Effect of reverberant energy on left-ear magnitude spectra. (a) Mean change in spectral magnitude per Hz as a function of source azimuth, calculated by taking the mean of the absolute values of the derivative of each measurement in dB per Hz, then calculating the mean across repeated measures. (b) Mean absolute difference between energy in smoothed third-octave bands of reverberant and *pseudo-anechoic* transfer functions as a function of source azimuth, calculated by computing the absolute value of the difference of the energy falling within each third-octave band of the reverberant and *pseudo-anechoic* transfer functions, calculating the mean across frequency bands, then calculating the mean across the three repeated measures. In both (a) and (b), error bars show the across-repetition standard deviation (where defined) and each row gives results for a different distance from source to listener (0.15, 0.40, and 1 m from top to bottom). Results within each panel correspond to different listener locations in the room.

left to right] as expected, as the direct-sound energy at the left ear decreases with source azimuth. Looking within each panel in Fig. 3(a), the average fluctuation is greatest for the *center* and *back* conditions (dark dashed lines) and much smaller for the *ear* and *corner* conditions (light-gray dashed lines).

The late portions of the left-ear results for the 0.15-m sources contain relatively more measurement noise than the other measurements (see footnote 3). Therefore, the average spectral fluctuation is slightly greater for the nearest source than the more distant sources [compare results in the top panel of Fig. 3(a) to those in the middle and bottom panels]. When the effect of the measurement noise is factored out,⁵ the magnitude of the frequency-to-frequency variations is roughly constant with increasing source distance. The maximum mean fluctuation in the BRTFs is on the order of 2 dB.

Because the magnitude of the BRTF fluctuations is directly related to the proportion of late-arriving energy making up each impulse response, the values analyzed here should be inversely correlated with standard room-acoustics metrics such as the clarity index, which computes the ratio of early-to-late energy in a room impulse response. C_{80} (the ratio of the energy in the first 80 ms of the impulse response over the energy in the remaining portion of the impulse response) was computed for each BRIR. The mean of this value was then computed across repeated measurements for each ear and each spatial configuration.

Table III gives the mean values of C_{80} for the various measurements. As expected, C_{80} decreases with source dis-

TABLE III. Mean clarity index C_{80} (in dB) averaged across the repeated measurements for each of the BRTFs. (Note that the clarity index is infinite, by definition, for the *anechoic* conditions.)

Ear	Listener location	Source distance (m)	Source azimuth						
			0°	15°	30°	45°	60°	75°	90°
Left	Center	0.15	22.9	21.5	16.8	16.7	16.0	13.7	10.9
		0.40	23.2	20.3	17.3	16.5	15.3	14.5	11.5
		1	18.3	15.2	13.0	12.0	11.1	10.8	10.2
	Back	0.15	22.9	20.8	17.1	17.3	16.4	14.2	11.1
		0.40	24.1	21.0	17.8	17.0	15.9	15.7	12.7
		1	19.5	15.8	14.0	12.7	12.2	11.2	9.4
	Ear	0.15	23.4	21.8	18.1	18.0	16.7	12.3	11.9
		0.40	24.2	22.6	20.1	20.2	18.7	16.9	15.5
		1	20.0	19.0	18.0	18.3	18.0	14.7	15.7
	Corner	0.15	25.3	22.4	18.6	19.6	17.5	15.1	12.3
		0.40	25.7	22.4	21.3	21.2	20.1	19.3	18.5
		1	21.6	20.2	19.4	19.7	19.3	16.8	17.8
Right	Center	0.15	23.5	26.6	31.2	34.1	36.3	36.9	39.3
		0.4	25.3	27.2	29.9	31.3	32.2	33.1	33.7
		1	18.6	20.5	22.1	22.8	23.3	23.0	23.0
	Back	0.15	23.1	25.4	30.9	34.7	35.8	36.3	40.1
		0.40	24.6	27.2	30.0	32.3	33.1	33.9	33.3
		1	19.0	21.1	22.8	24.0	24.0	24.1	23.7
	Ear	0.15	25.0	26.8	31.5	34.8	37.3	37.2	39.4
		0.40	24.7	26.6	29.6	32.0	32.9	33.9	34.2
		1	18.7	20.1	22.2	22.9	23.6	23.4	23.6
	Corner	0.15	24.1	25.9	31.8	35.0	37.7	35.8	38.2
		0.40	25.9	29.4	32.5	33.1	33.8	34.6	34.9
		1	20.2	21.9	23.1	24.1	23.8	23.9	23.8

tance in both ears. For the right ear, C_{80} increases with azimuth, whereas C_{80} decreases with azimuth in the left ear. Room location has little effect on the right ear results; however, the left-ear C_{80} is smaller in the *center* and *back* conditions than in the *ear* and *corner* conditions (where early reflections are more prominent); this difference is most pronounced at the greatest distance.

The correlation between C_{80} and the frequency-to-frequency variability (modified to remove measurement noise as discussed above) was calculated for the 168 independent measurements (2 ears, 4 room conditions, 3 distances, and 7 azimuths). As predicted, these measures are inversely correlated. The correlation coefficient was -0.89 (i.e., roughly 80% of the variance in C_{80} can be accounted for by knowing the frequency-to-frequency variability in the BRTF magnitude and vice versa).

2. Distortion of frequency-smoothed magnitude

If reverberant energy alters the BRTF magnitude averaged over a critical band, the mean spectral content of the BRTF-filtered source will be altered. Whereas frequency-to-frequency fluctuations in the BRTF magnitude can cause temporal fluctuations in the neural excitation pattern caused by a broadband source, systematic distortions of the BRTF magnitude (averaged over a critical band) will alter the mean excitation pattern. To quantify such systematic spectral distortion arising from reverberant energy, the BRTF magnitude was first smoothed over frequency and then compared to *pseudo-anechoic* results. The energy falling within third-octave wide energy bands (center frequencies spanning

0.2–20 kHz) was computed for each measurement in order to estimate how the average rate of a corresponding auditory nerve fiber would be influenced by the reverberant energy for an ongoing broadband stimulus. For each reverberant measurement, the smoothed spectral level was subtracted from the level of a smoothed *pseudo-anechoic* measurement (independent of the *center* result used for a given calculation). The absolute value of the difference between smoothed reverberant and *pseudo-anechoic* results was computed for each third-octave band and then averaged over center frequency. This statistic was then averaged across the independent repeated measures for each source direction, source distance, and listener location.

The mean effect on the right-ear spectra (not shown) is very small (2.4 dB or less) and does not vary significantly with room location, source direction, or source distance. Figure 3(b) plots the across-measurement mean absolute difference between the smoothed reverberant and *pseudo-anechoic* spectra for the left ear (including, where defined, the across-measurement standard deviation). Within each panel, results for each of the listener locations are shown as a function of source azimuth (measured relative to KEMAR's median sagittal plane). Each row gives results for a different source distance.

The effect of the reverberant energy on the left-ear spectra increases for all listener locations in the room as the source moves to the right. For the conditions with an early left-wall reflection [the *ear* and *corner* conditions; light gray results in Fig. 3(b)], the difference between the reverberant and *pseudo-anechoic* smoothed spectra is larger and in-

increases more rapidly with azimuth than in the other conditions [i.e., light gray dashed lines are above and steeper than the dark gray dashed lines in Fig. 3(b)]. The effects of reverberation on the *corner* magnitude spectrum are consistently larger than on the *ear* magnitude spectrum [in Fig. 3(b), the double-dot-dashed light gray line is above or equals the single-dot-dashed light gray line]. With increasing source distance, differences across the listener locations increase [e.g., the separation between the light and dark gray results in the left column increases as one looks from the top to the bottom panels in Fig. 3(b)].

3. Discussion

Whereas local frequency-to-frequency fluctuations in the spectral magnitude are biggest in conditions where diffuse reverberant energy is relatively more important [i.e., in the *center* and *back* conditions; dark gray lines in Fig. 3(a)], systematic deviations between the frequency-smoothed *pseudo-anechoic* and frequency-smoothed reverberant magnitude spectra are largest in the cases where there is a sufficiently strong early reflection to cause comb-filtering of the spectrum [i.e., in the *ear* and *corner* conditions shown by light gray lines in Fig. 3(b)]. Similarly, while source azimuth has the greatest effect on frequency-to-frequency variability in conditions dominated by diffuse reverberant energy (*center* and *back*; dark gray lines in Fig. 3(a); conditions producing the smallest values of C_{80}), azimuth has the greatest influence on changes in the smoothed spectral shape for conditions with an early intense reflection [*ear* and *corner*; light gray lines in Fig. 3(b)]. Overall, these results show that frequency-to-frequency variability in the BRTF magnitude spectra [Fig. 3(a)] depends primarily on the strength of the diffuse reverberant energy relative to the sum of the direct-sound energy and the energy in any early reflection (such as those present in the left ear for the *ear* and *corner* measurements). In contrast, systematic distortion of spectral shape cues depends primarily on the energy in any early intense reflections relative to the direct-sound energy. The location of the listener in the room alters the intensity of any early reflections reaching the listener and, therefore, influences both frequency-to-frequency variability and systematic spectral distortion.

As discussed in the Introduction, frequency-to-frequency variability in the BRTF is related to the across-stimulus-token and across-time fluctuations likely to be observed in a BRTF-filtered sound source. In contrast, systematic deviations between frequency-smoothed BRTFs and their anechoic counterparts are related to how the expected value of the received spectral shape will be affected by the reverberant energy in a room. Because perception of spectral shape is important for judging the location of a source within a cone or torus of confusion, both spectral variability and mean spectral distortion may influence sound localization by listeners in rooms. Frequency-to-frequency variability may reduce the reliability with which the expected spectral shape can be extracted from a finite-length broadband stimulus, and thus might increase variability in judging source location (either across time or across stimulus tokens). In contrast, systematic distortion of spectral cues

may induce localization bias, as it will tend to alter the mean perceived spectral shape (again, averaged across time for a long-duration stimulus or across tokens for shorter stimuli). Thus, the way in which reverberant energy influences sound localization in the up/down dimension may depend on the listener location in a room, with token-to-token response variability largest in conditions where the random reverberation energy is largest, and response bias largest in conditions where a listener is near a reflecting surface.

The frequency-to-frequency variability discussed in this section is directly related to the relative energy in the late-arriving, reverberant energy compared to the sum of the earlier portions of the room impulse responses. As a result, this metric is inversely correlated with such common room-acoustics metrics as C_{80} (which is commonly used to summarize the subjective influence of room acoustics on listening to music) and D_{50} (which predicts the effect of reverberant energy on speech quality in rooms). One advantage of examining the current metric rather than either of these metrics is that it does not depend on making any one choice about what time window should constitute “early” and “late” portions of the room transfer function.

C. Interaural level differences

1. ILDs in binaural room transfer functions

As with monaural spectral cues, reverberant energy will alter the across-time mean of and variability in ILD cues. Reverberant energy will cause frequency-to-frequency fluctuations in the ILD that are directly related to the monaural frequency-to-frequency fluctuations discussed in the previous section. Such fluctuations, in turn, will cause temporal fluctuations in the short-term ILD observed through any critical band filter for broadband stimuli. Because the influence of reverberant energy on monaural spectral level is larger in the ear far from the source than the near ear, the overall effect of reverberant energy on frequency-to-frequency variability in the ILD increases with source laterality and with the relative level of diffuse reverberant energy (i.e., ILD variability over frequency is inversely correlated with C_{80}).

To quantify how reverberant energy alters the expected ILD for broadband sources, the energy falling within third-octave wide energy bands from 200 Hz to 20 kHz was computed for the left- and right-ear BRTFs for each measurement. These values were subtracted to estimate the ILD in the room response filters as a function of frequency and then averaged across the independent repeated measures for each center frequency, source direction, source distance, and listener location.

Figure 4 plots the across-measurement mean ILD in the smoothed BRIR spectra and, where defined, the across-measurement standard deviation as a function of the center frequency of the third-octave bands. Within each panel, the ILDs are shown for each of the four listener locations in the room (dashed lines) and for an independent *pseudo-anechoic* BRIR (solid line). Each row gives results for a different

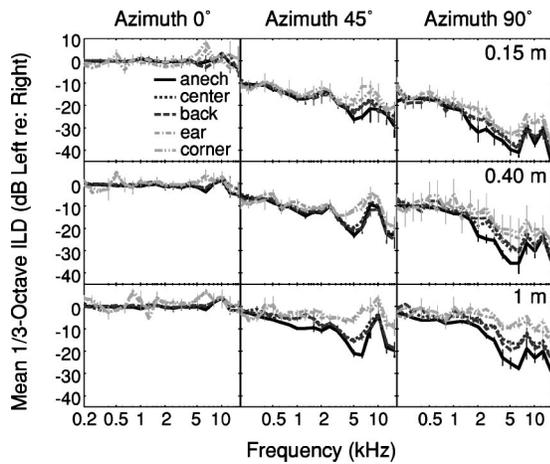


FIG. 4. Mean interaural level differences in the binaural room transfer functions as a function of frequency. Each point was calculated by computing the energy in each third-octave band of the transfer functions, taking the absolute value of the difference between the left- and right-ear values, then averaging these values across the independent repeated measures for each condition. Error bars show the across-repetition standard deviation (where defined). Results within each panel correspond to different listener locations in the room. Results are shown for sources at 0°, 45°, and 90° azimuth (left, middle, and right columns, respectively). Each row gives results for a different source distance (0.15, 0.40, and 1 m in the top, middle, and bottom row, respectively).

source distance. The left, middle, and right columns show results for sources at azimuth angles of 0°, 45°, and 90°, respectively.

For the source directly in front of the listener (left-hand column of Fig. 4), the ILD is very small at all frequencies. However, for the conditions in which the left ear receives an early intense reflection, there are small but consistent ILDs at some low to mid frequencies (a consequence of comb-filtering effects in the left-ear magnitude spectra for the *ear* and *corner* conditions; at higher frequencies, third-octave smoothing hides any comb-filtering and the only consistent effect is a boost in the left-ear energy from the early reflection in the left ear, which causes a slightly positive ILD). Around 6 kHz, there is a modest notch in the direct-sound spectrum that is symmetrical at the two ears for the 0° source location shown in Fig. 4. However, for conditions with a strong left-wall reflection, the left-ear notch is filled in more than in the right ear, producing a small but consistent positive ILD near 6 kHz.

Reverberant energy decreases the magnitude of ILDs and the size of this effect depends on the listener location in the room. In general, the ILD magnitude is smallest for the conditions with early intense reflections (light gray, dashed lines corresponding to the *ear* and *corner* conditions in Fig. 4), intermediate for the conditions with relatively diffuse reverberant energy (dark gray, dashed lines corresponding to the *center* and *back* conditions), and largest for the *pseudo-anechoic* results (solid dark lines). For instance, for the 1-m, 45° source the *pseudo-anechoic* ILD is nearly -20 dB at 5 kHz (solid line in the bottom middle panel of Fig. 4), roughly -12 dB for the *center* and *back* conditions (dark gray dashed lines), and only around -5 dB for the *ear* and *corner* conditions (light gray dashed lines). The effect of reverberant energy on the ILD increases with source distance (e.g., de-

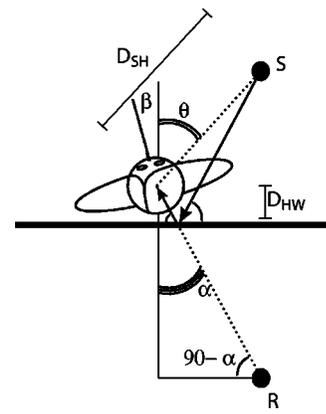


FIG. 5. Schematic diagram showing the source (S) and effective location of the back-wall reflection (R) for the *back* condition. The direct-sound path is shown by the black dashed line, arriving from angle θ relative to a line perpendicular to the wall. The reflection path is shown by solid black lines. The point R shows the effective location of the back-wall reflection (see Allen and Berkley, 1979). The angle of the reflection is given by α . The distance from source to the center of the head is given by D_{SH} and the distance from the center of the head to the wall is given by D_{HW} . When the listener's interaural axis is parallel to the wall (β is zero), the lateral angles (relative to the listener's median sagittal plane) for the arrival direction of S and R equal θ and α , respectively.

creasing the ILD magnitude by as much as 20 dB for a 1-m, 90° source in the bottom right panel of Fig. 4, but at most by 10 dB for the 0.15-m, 90° source in the top right panel).

2. Discussion

The impact of reverberant energy on ILDs is dominated by the effect of reverberation on the signals at the ear that is farther from the source. With increasing source distance, reverberant energy tends to increase the energy in the (left) shadowed ear, thereby decreasing the ILD. As the source moves to the right side of the head, reverberant energy causes a relatively larger change in the total energy reaching the left ear and the ILD magnitude decreases by a larger amount. For the *ear* and *corner* conditions, the early intense reflection reaching the left ear off of the nearby left wall adds significantly more energy to the left ear than to the right ear, which receives a head-shadowed version of this early wall reflection. Thus, for these conditions, the ILD magnitude decreases even more due to reverberant energy (particularly at the largest source distances) than for the reverberant conditions where diffuse energy is more dominant.

The similarity between *center* and *back* results can be partially explained by considering the geometry when the listener is near a wall. Figure 5 shows a schematic diagram of this situation. A source S reflects off the back wall (see black lines with arrows), producing a reflection that is roughly equivalent to a phantom source at location R [see Allen and Berkley (1979) for a description of this sort of geometric approximation]. For this geometry, the angle between the arrival direction of the sound source and a line perpendicular to the wall (θ) and the angle between the arrival direction of the reflection and the same reference (α) are related by

$$\alpha = \tan^{-1} \left[\frac{D_{SH} \sin \theta}{2D_{HW} + D_{SH} \cos \theta} \right]. \quad (1)$$

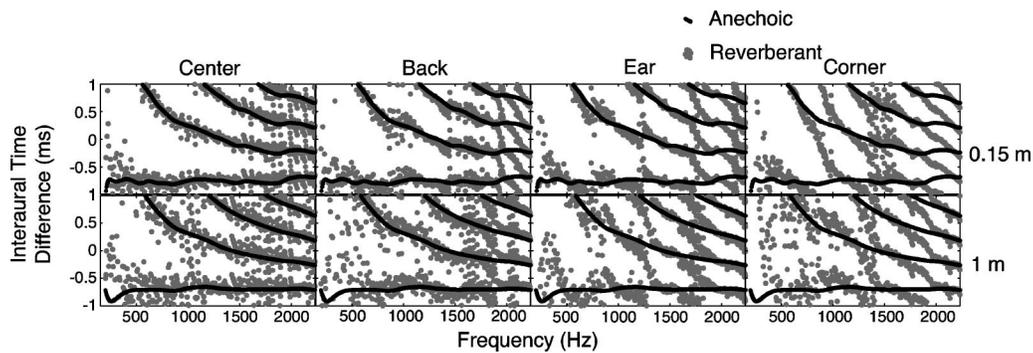


FIG. 6. Interaural time differences in sample room transfer functions with the source at 90° azimuth (*re*: KEMAR) as a function of frequency (shown up to 2.2 kHz). Values were calculated by subtracting the phase of the left- and right-ear transfer functions at each frequency, then dividing by frequency. For each frequency, multiple values are plotted, corresponding to all possible shifts of the ITD by integer multiples of the period (a pure, frequency-independent ITD would give rise to a horizontal line at the true ITD value). Gray points show results for the reverberant conditions. For comparison, black points show *pseudo-anechoic* measurements (derived from an independent set of *center* measurements). Each column shows results for one of the four listener locations in the room. The top and bottom rows give results for distances of 0.15 and 1 m, respectively.

The smaller the distance between listener and wall (D_{HW}) and the greater the distance from source to listener (D_{SH}), the more similar θ and α become. In addition, the difference between θ and α approaches 0 as θ approaches 0. When the listener's interaural axis is oriented parallel to the wall as in the *back* condition (i.e., when β in the figure is zero), θ equals the lateral angle of the direct sound and α equals the lateral angle of the reflection. Thus, when the listener is oriented with his interaural axis parallel to a nearby wall, both the direct sound and the early reflection produce nearly identical interaural differences. As the listener rotates (as the magnitude of β increases from zero), the lateral angle of the direct and phantom sources become increasingly disparate, reaching an extreme when the listener has one ear towards the wall (e.g., β is -90° , as in the *ear* and *corner* conditions) when the source is directly to one side of the listener (θ is 0°) and the phantom source is on the exact opposite side of the listener. As a result, early reflections in the *back* condition produce significantly less interaural distortion than the early reflections in the *ear* and *corner* conditions.

To the extent that the magnitude of the ILD influences perception of source laterality, sound sources may be systematically perceived as closer to the median sagittal plane in a room than in anechoic space, as the expected ILD magnitude is generally smaller in a room. Furthermore, any such systematic localization bias would be greatest in conditions where the listener is oriented with one ear facing a wall (such as the *ear* and *corner* conditions).

In addition to reducing the average ILD magnitude, diffuse reverberant energy increases the frequency-to-frequency variation in ILDs by introducing frequency-to-frequency variability in monaural spectral levels [e.g., see Fig. 3(a)]. Thus, ILD cues will tend to vary with the short-term spectral content of a source in the presence of diffuse reverberant energy.

Although low-frequency ILD cues [below 3 kHz; see Brungart (1999)] convey source distance for nearby sources, low-frequency ILD cues are less influenced by ordinary room reverberation than are ILDs in the mid to high frequencies (compare results above and below 3 kHz in Fig. 4).

Thus, distance perception of nearby sources may be robust in a room, especially given that reverberant energy provides additional distance information to listeners (Mershon *et al.*, 1989).

D. Interaural time differences

1. ITDs in BRTFs as a function of frequency

Frequency-to-frequency variation in ITD cues in BRTFs will cause temporal variability in the ITD cues reaching a listener for a broadband sound presented in a room. Figure 6 compares the ITDs in the reverberant and *pseudo-anechoic* conditions for a source at 90° , plotted as a function of frequency for the low to mid frequencies (200 Hz to 2.2 kHz, frequencies for which the ITD is thought to be most salient). Each plot in Fig. 6 was generated from one individual BRTF measurement. Rather than trying to extract the "true" ITD at each frequency, only the phase-wrapped IPD, denoted by $\phi_{\text{mod}(2\pi)}$, was computed: the phases of the left and right BRTFs at each frequency were subtracted and set to an equivalent value modulo 2π . In order to account for the inherent interaural phase ambiguity in $\phi_{\text{mod}(2\pi)}$, at each frequency f a vector of possible IPDs was generated by calculating $\phi_{\text{mod}(2\pi)} + 2\pi k$ for all integer values of k . The resulting values were divided by $2\pi f$ to generate multiple ITD values consistent with the observed IPD at each frequency. Figure 6 plots these ITD values as a function of frequency when the sound is 90° to the right of the listener's head. In each panel, gray symbols correspond to the reverberant results and black symbols correspond to *pseudo-anechoic* results. The *pseudo-anechoic* results in each row are identical and derived from an independent *center* measurement to make across-listener-location comparisons "fair." The top and bottom rows give results for 0.15- and 1-m sources, respectively. Each column shows results for one listener location.

All *pseudo-anechoic* plots (black symbols, repeated in each column) contain a roughly horizontal line, consistent with an ITD that is nearly the same at all frequencies. This ITD value is approximately $-750 \mu\text{s}$ for the 90° locations shown, independent of source distance. The other ITD values

consistent with the IPD at each frequency form stereotypically curved lines in the ITD-frequency plots (see, for example, Stern and Trahiotis, 1995).

Looking first at the results for the *center* condition (left column), the relatively diffuse reverberant energy causes small distortions of the ITD that appear to be random and independent from frequency to frequency (i.e., in the left-hand column results, the gray symbols fall near, but randomly scattered around, the black *pseudo-anechoic* results). The amount of distortion caused by the reverberant energy increases with source distance (deviations between the gray symbols and *pseudo-anechoic* results are larger in the bottom left panel than in the top left panel of Fig. 6).

Results for the *back* condition (second column from left) are similar, except that deviations from *pseudo-anechoic* results are greater than for the *center* condition. For a sound source at 90°, the back-wall reflection arrives from a different lateral angle than the direct sound [e.g., roughly 70° for the 1-m source; see Eq. (1)], thus altering the ITD cues reaching the ears. Furthermore, the distortion of the ITD cues varies systematically with frequency (see the upper panel in the second column of Fig. 6), a sign that part of the distortion of ITD cues is from a distinct reflection, rather than diffuse reverberant energy.

Results for the *ear* and *corner* conditions (third and fourth columns in Fig. 6) show prominent, systematic deviations from *pseudo-anechoic* ITD results. For example, for the 0.15-m source in the *ear* condition (top panel of third column), the reverberant results (gray symbols) nearly match the *pseudo-anechoic* results (black symbols) at 500 Hz, but are consistently above the *pseudo-anechoic* results just below 500 Hz and below the *pseudo-anechoic* results just above 500 Hz. With increasing source distance, the distortion caused by reverberant energy grows (e.g., compare the top and bottom panels in the third column). Indeed, the panel showing the 1-m source in the *ear* condition (gray symbols, bottom panel in the third column) shows a striking pattern of negatively sloped lines, without any obvious horizontal line; in other words, there is no single ITD in the plotted range that is consistent across frequency.

2. Cross-correlation analysis

Frequency-to-frequency variability in the ITDs contained in BRTFs influences how short-term ITD cues in a broadband sound source will vary. However, analysis of such variability does not address whether reverberant energy causes a change in the mean ITD. In this section, we compute a broadband cross-correlation of the left- and right-ear BRIRs to investigate whether reverberation causes systematic distortions of ITD information. This analysis is closely related to analysis of the interaural cross-correlation commonly employed in room acoustics studies to summarize the interaural decorrelation caused by room reflections (e.g., see de Vries *et al.*, 2001). Although neural computation of ITD information is performed by computing the narrow-band cross correlation within each critical band as a continuous time function, the broadband cross-correlation function is essentially a weighted average of narrow-band cross-correlation functions when the weighting of each constituent

frequency band is proportional to the energy in that band [see Shinn-Cunningham and Kawakyu (2003) for analysis of short-term, narrow-band cross-correlation results and how across-time integration affects ITD estimation from reverberant signals]. As such, the broadband analysis shown here provides a simple summary of the extent to which ITD cues in a room can provide useful information about source laterality.⁶

To quantify the effects of reverberant energy on ITD cues, the normalized cross-correlation of the low-to-mid-frequency portion of the left- and right-ear impulse responses was calculated. For each measurement, the left- and right-ear impulse responses were low-pass filtered with a cutoff frequency of 3 kHz (note that the upper-frequency cutoff has little influence on the results we will present; 3 kHz was chosen in order to focus analysis on the low-to-mid frequency region in which ITD cues are thought most salient). The normalized cross-correlation function $x(\tau)$ was then computed as

$$x(\tau) = \frac{\sum_n l[n]r[n-\tau]}{\sqrt{\sum_n l^2[n]\sum_m r^2[m]}}, \quad (2)$$

where $l[n]$ and $r[n]$ are the low-pass filtered left- and right-ear impulse responses, respectively. This normalized cross-correlation function takes on its largest value at the time delay τ that best aligns the left and right ear signals. For each measurement, the largest peak of $x(\tau)$, p_{\max} , was found along with the corresponding time delay, τ_{\max} . Because the largest such peak might fall outside the plausible range of ITD values for a sound in anechoic space, the largest peak within the limited range of $[-1, 1]$ ms was also found (p_{lim}) along with the corresponding interaural time delay (τ_{lim}).

Figure 7 plots the values of τ_{lim} and τ_{\max} [Figs. 7(a) and (b), respectively] and the normalized peak amplitudes p_{lim} and p_{\max} [Figs. 7(c) and (d), respectively] as functions of the source azimuth. Results for the three source distances are shown in individual panels; results within each panel show different listener locations in the room. Lines show the across-repeated-measurement mean and symbols show individual measurement results.

In all the panels of Fig. 7(a), the peak ITD magnitude increases from roughly 0 to roughly 750 μs as the azimuth ranges from 0° to 90°. Despite the strong distortion of ITD information by reverberation (see Fig. 6), τ_{lim} is essentially unaffected by reverberant energy: in all three panels of Fig. 7(a), results for all listener locations fall on top of *pseudo-anechoic* results. In other words, despite the frequency-to-frequency fluctuation in the ITD, integrating information across frequency (e.g., by calculating the broadband cross-correlation from 200 Hz to 3 kHz) gives a reliable ITD estimate when the range of candidate ITD values is limited.

Results in Fig. 7(b) show that τ_{\max} equals τ_{lim} for all cases in which there are no prominent early reflections (i.e., for the *pseudo-anechoic*, *center*, and *back* conditions; see solid and dashed dark lines). However, for the conditions with an early reflection (the *ear* and *corner* conditions, shown by light results), the largest overall peak in the normalized cross-correlation function does not always fall

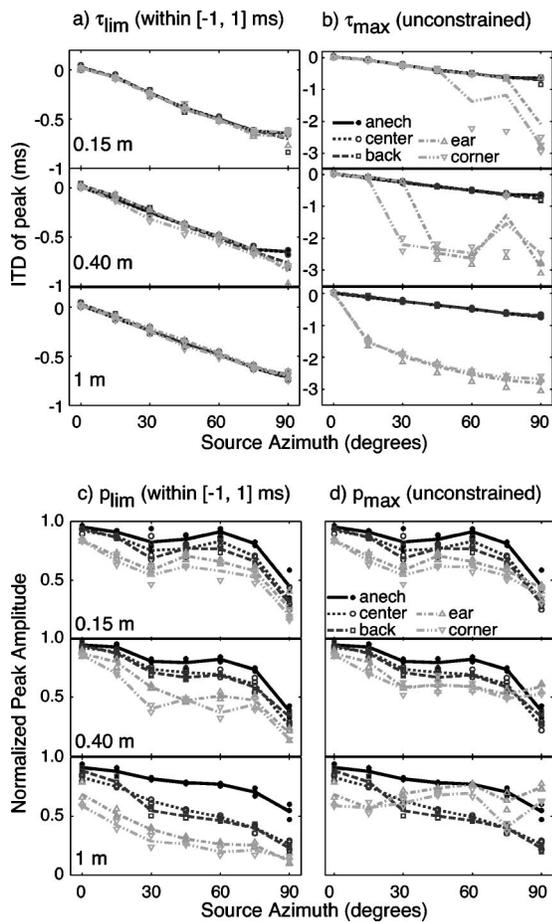


FIG. 7. ITD and normalized height of the peak in the broadband cross-correlation function of left- and right-ear impulse responses as a function of source azimuth. (a) ITD of largest peak within the physiologically plausible range of interaural time differences (-1 to $+1$ ms). (b) ITD of the largest overall peak (without limiting the ITD range). (c) Peak height within the physiologically plausible range of interaural time differences (-1 to $+1$ ms). (d) Height of the largest overall peak (without limiting the ITD range). Lines show the across repetition average for each listener location in the room. Results for each independent repeated measurement are plotted as symbols. Each row gives results for a different distance from source to listener (0.15, 0.40, and 1 m in the top, middle, and bottom rows, respectively).

within the plausible range. The likelihood that τ_{\max} differs from τ_{\lim} increases with increasing source azimuth and with increasing source distance. For instance, the number of *ear* and *corner* measurements (light gray results) with implausible peaks is largest in the bottom panel of Fig. 7(b) and increases from left to right in the panel.

Figure 7(c) shows that for all source distances, p_{\lim} decreases with increasing source laterality (in all panels, p_{\lim} decreases from left to right) and with increasing source distance (p_{\lim} decreases from top panel to bottom panel). Whereas τ_{\lim} does not vary with listener location in the room, p_{\lim} varies dramatically: in general, p_{\lim} is largest for the *pseudo-anechoic* results (solid black line), intermediate for the *center* and *back* conditions (dark gray results), and smallest for the *ear* and *corner* conditions (light gray results). Differences between the different listener locations also increase with source distance. In fact, for the most-distant sources tested [bottom panel of Fig. 7(c)], even

though τ_{\lim} is essentially unaffected by reverberation [see Fig. 7(a)], p_{\lim} for the *ear* and *corner* conditions is very small (less than 0.4 for sources beyond 15° to the right of the listener).

Figure 7(d) plots p_{\max} for the various source positions and listener locations. For listener locations in which there is no prominent early reflection, p_{\max} always equals p_{\lim} , independent of source azimuth and distance [i.e., the solid and dashed dark line results are identical in Figs. 7(c) and (d)]. For the nearest sources [top panel in Fig. 7(d)], the peak amplitude is essentially unchanged when the peak ITD value is unrestricted; only for the sources at 90° are there differences between p_{\max} and p_{\lim} , and only then for the *ear* and *corner* conditions (compare the rightmost light gray points in the top panels of Figs. 7(c) and (d)). With increasing source distance, however, p_{\max} in the *ear* and *corner* conditions changes dramatically. For the intermediate source distance, p_{\max} for the *ear* and *corner* conditions decreases as the source azimuth moves from 0° to 30° , but then changes very little as the source angle changes from 30° to 90° [light gray results in the middle panel of Fig. 7(d)]. For the most distant source in the listener locations for which there is an early left-wall reflection, p_{\max} actually increases as the source azimuth changes from 0° to 90° [light gray results in the bottom panel of Fig. 7(d)].

3. Discussion

Because monaural phase distortion is larger in the ear that receives less direct sound energy, ITD distortion depends most strongly on the phase distortion in the far ear, increasing with source laterality as well as source distance. For listener locations in which there is an early reflection from the left wall, the pattern of ITD as a function of frequency is grossly distorted, even when the source is very near the listener's head.

The small but consistent difference in ITD distortion in the *center* and *back* conditions can be partially explained by the geometry of these listening conditions. The early back-wall reflection causes ITD cues that are identical to those in the direct sound when the source is directly in front of the listener [see Fig. 5 and Eq. (1)]. Thus, for the 0° source angle, the distortion of ITD by reverberation is smaller in the *back* condition than the *center* condition. As the source azimuth increases, the back-wall reflection's lateral angle no longer exactly matches the direct-sound lateral angle [particularly for the smallest source distance; see Eq. (1)], and the ITD deviation in the *back* condition grows rapidly with source azimuth.

Whereas the ITD information in any specific frequency may be dramatically distorted, in the broadband cross-correlation function there is a local ITD peak corresponding to the direct-sound path. These results show that even when there is an early intense reflection, knowledge of the range of "true" ITDs and integration of information across frequency enables accurate estimation of the sound source lateral angle. Behaviorally, the distortion of ITD cues by reverberant energy may increase variability in judgments of source laterality across different stimulus tokens. However, because the ITD of the direct sound can be recovered with sufficient

across-frequency integration, laterality judgments based on ITD cues may not show any large bias, regardless of the location of the listener in the room.

IV. SUMMARY AND GENERAL DISCUSSION

This study examines the effects of source distance, source angle, and listener location on the distortion of spatial cues in BRIRs by reverberant energy when sources are nearby a listener in an ordinary room. All of these factors have a significant impact on the form and size of the distortion caused by reverberant energy. This kind of analysis is important for understanding both how reverberant energy alters the statistics of acoustic attributes in the signals reaching a listener (e.g., cues for source direction and for source content) and how reverberant energy can provide information to the listener (e.g., about source distance and the acoustic environment).

Room reflections alter the magnitude spectra of BRIRs by causing random frequency-to-frequency variation; filling in spectral notches; increasing the overall energy, particularly in the ear receiving less direct-sound energy; and, in some cases, generating comb-filter distortion. As source laterality increases, all of these effects become more pronounced at the far ear and smaller at the ear on the near side of the head. While the random variations in spectral content are greatest when most of the reverberant energy is diffuse (*center* and *back* conditions), the systematic distortions of spectral content are greatest when there is a prominent early reflection (*ear* and *corner* conditions).

Spectral magnitude distortion degrades the cues important for resolving source location within a particular torus of confusion. Such spectral distortion is much greater in the far ear than the ear near the source. This observation may explain why listeners give more perceptual weight to spectral shape cues in the near ear when the left- and right-ear spectra are pitted against one another (Morimoto, 2001): in many ordinary (reverberant) environments, the near-ear spectral cues are relatively more reliable than the far-ear cues.

In all cases, reverberation tends to reduce ILDs, especially when the source is relatively far from the listener. For the source locations tested in this study, the reduction of ILDs is especially pronounced when there is an early reflection from the side (i.e., in the *ear* and *corner* conditions). Furthermore, diffuse reverberant energy adds frequency-to-frequency variability in ILD cues. To the extent that listeners use ILD cues to judge source laterality, reverberant energy may therefore induce systematic localization bias, with listeners perceiving sources nearer the median sagittal plane than the true source location, as well as increase variability in judgments of source laterality.

ITD information becomes more distorted by reverberant energy as the source moves away from the median sagittal plane and as source distance increases. Distortion from reverberant energy can cause quite severe interaural decorrelation (e.g., significantly decreasing the amplitude of the “true” peak in the normalized cross correlation). Nonetheless, the ITD of the direct sound can still be extracted from the BRIR by combining ITD information across frequency [accomplished in the current analysis by computing a broad-

band cross correlation; however, see also Shinn-Cunningham and Kawakyu (2003) for a more physiologically motivated analysis of what ITD information can be extracted in a reverberant setting when considering short-term ITD cues arising within critical bands] and restricting the range of ITD values to those that are plausible. In fact, psychophysical tests and theoretical studies suggest that although ITD is computed independently within a critical frequency band, listeners integrate ITD information across frequency when judging source laterality (e.g., see Trahiotis and Stern, 1989; Buell and Hafter, 1991). Perhaps as a result of such across-frequency integration, listeners are able to make reasonable judgments of source laterality of broadband sounds in some reverberant settings (e.g., see Hartmann, 1983; Shinn-Cunningham, 2000b).

The current results suggest that for many aspects of BRIRs, the most important differences that arise from changing listener location in a room can be ascribed to the presence or absence of an early reflection arriving from the side of the head that is opposite the direct sound (i.e., such as arises in the current results in *ear* and *corner* conditions). The strength of any early reflections depends strongly on the listener location in the room and more modestly on the source location relative to the listener. Only when the listener is positioned with one ear facing a nearby wall is the initial reflection sufficiently strong (and from a sufficiently different direction than the direct sound) to cause any systematic distortion of spatial acoustic cues. In the current measurements, reflections from a wall behind the listener (i.e., in the *back* condition) do not cause obvious comb-filtering spectral distortions. Furthermore, the interaural differences in the early reflection from a back wall tend to reinforce rather than distort the interaural cues in the direct sound. As a result, current results from the *back* condition are comparable to those of the *center* condition and unlike those from the *ear* and *corner* conditions for nearly all of the statistics examined.

The acoustic similarity of many spatial cues in the *center* and *back* conditions is consistent with results of a recent behavioral study that examined the ability of trained listeners to identify the four listener locations in the room when listening to stimuli simulating different source and listener configurations using the BRIRs analyzed here (Shinn-Cunningham and Ram, 2003). In this study, listeners were given blocks of trials in which the simulated source location of a random token of noise was fixed relative to the listener, but the simulated listener location in the room varied from trial to trial. The listener’s task was to *identify* which of the listener locations was simulated on each trial (i.e., listeners had to categorize the room location while ignoring trial-to-trial variability caused by variations in the noise tokens presented). Listeners generally did not confuse listener locations in which there was an early reflection from the left (the *ear* and *corner* conditions) with stimuli in which there was no such reflection (the *center* and *back* conditions), but were very likely to confuse the *ear* and *corner* stimuli (mislabeling an *ear* trial as a *corner* trial and vice versa) and to confuse *center* and *back* stimuli (although listeners might have been able to *discriminate* between these stimuli, especially if

the noise token had been held constant across trials).

The current analyses may underestimate the effects of reverberation in an ordinary classroom, as the source distance was 1 m or less in all of the conditions considered here. In fact, results show that even for these conditions, where the source is very close to the listener, the effects of reverberation can be quite prominent. These results shed light on the ways in which source location relative to the listener (including source distance) and listener location in the room interact with the effects of reverberation to influence the signals reaching a listener. Of course, many other factors can alter the pattern of direct and/or reverberant energy reaching the listener, including, among other things, room properties (volume, dimensions, and surface properties), source directivity, and source orientation. However, many of the results from this study are applicable to other environments and source-listener configurations that give rise to similar relative levels of direct-sound, early reflection, and reverberant energy. Similarly, the current analyses focus on steady-state properties in the BRIRs and do not directly address how the spectro-temporal content of a sound source interacts with properties of the room or how reverberant energy builds up and decays over time. Instead, these results examine different properties of BRIRs that can influence the mean and variability in spatial cues received by a listener in a room and show how these properties vary with source and listener position.

V. CONCLUSIONS

Measurements of binaural room impulse responses in the classroom studied here demonstrate a number of important principles governing how reverberant energy will distort the spatial acoustic cues reaching the ears of the listener. Some of these principles, which will also apply in other acoustic settings, are listed here:

- (1) Especially for sources near the listener, the listener location in a room and the source position relative to the listener influence how reverberant energy distorts and degrades spatial acoustic cues. Systematic distortions are most prominent when a listener is oriented with one ear toward a wall. Reverberant energy influences signals more at the ear farther from the source than at the near ear, particularly when a source is nearby the listener. Spectral cues are more reliable in the near ear than the ear farther from the source, and interaural cues become less reliable with source laterality.
- (2) Interaural level differences are systematically reduced by reverberant energy, such that the mapping between the expected ILD (as a function of frequency) and the laterality of a sound source presented in reverberant space depends on acoustic properties of the environment as well as on the location of the listener in the environment.
- (3) For moderate levels of reverberation (such as observed here), the ITD in the direct sound can be reliably extracted in reverberant signals by integrating interaural cues across frequency and restricting the estimated ITD to a plausible range of values.
- (4) Reverberant energy may have different effects on sound localization in the left/right, up/down, and distance di-

mensions. Judgments of source laterality may be biased in a room because reverberant energy reduces ILD magnitude; additionally, because reverberation causes temporal fluctuations in both short-term ITD and ILD cues, laterality judgments may show greater token-to-token variability in a room than in anechoic space. Similarly, judgment of the up/down direction of a source may be biased because reverberant energy tends to alter the mean spectral shape of the signals reaching the listener. Finally, distance judgments may be more accurate in a room than in anechoic space, as low-frequency ILD cues that arise for nearby lateral sources are robust to the effects of reverberation and reverberant energy may provide additional distance cues.

ACKNOWLEDGMENTS

This work was supported in part by AFOSR Grant No. F49620-98-1-0108 and the Alfred P. Sloan Foundation. Steve Colburn, Armin Kohlrausch, and two anonymous reviewers provided very helpful and careful comments that vastly improved the manuscript.

¹Sound localization of human listeners has been studied in the same room, using the same listener locations and a similar range of source locations relative to the listener (e.g., see Santarelli, 2000; Shinn-Cunningham, 2000b; Kopco and Shinn-Cunningham, 2002). The BRIRs analyzed here also have been used in headphone-based experiments investigating how reverberation influences various aspects of perception (e.g., Shinn-Cunningham *et al.*, 2002; Devore and Shinn-Cunningham, 2003; Shinn-Cunningham and Ram, 2003; Shinn-Cunningham, 2004). Researchers interested in obtaining copies of these BRIRs for their own use can contact the first author at shinn@cns.bu.edu

²The combination of the source distances relative to KEMAR and the height of KEMAR relative to the floor were chosen to ensure that the initial floor reflection in the *center* position was temporally separated from the direct sound impulse response.

³Because the overall level of the MLS measurement signals varied (in order to present signals at the most intense level possible without distortion), the time at which the BRIRs reached the noise floor depended on the location of the source relative to the listener. For the nearest sources, the measurements were taken at a lower overall gain in order to avoid clipping of the direct sound impulse response. As a result, the BRIRs for the 0.15-m distant sources reached the noise floor as much as 150 ms earlier than the BRIRs for sources at 0.40 and 1 m. Because the BRIRs for the 0.15-m sources contain not only reverberation but also additional electro-acoustic noise, the distorting effects of reverberation (compared to the anechoic HRIRs) are, if anything, overestimated in the results presented here. However, the effect of this additional noise was found to be negligible. To validate results, all BRIRs were also processed with a more sophisticated time-windowing scheme. In this scheme, the time at which each measurement reached the noise floor was estimated by determining when the energy in the late portion of the BRIR no longer decayed with time. To approximate the correct exponential decay in the late portions of each BRIR, an appropriate exponentially decaying time envelope (whose decay time constant was determined from estimates of T_{60} for the room) multiplied the late portion of each BRIR from the time at which the measurement hit the noise floor to the end of the BRIR (note that the imposed envelope was longest for the nearest sources). In almost all cases, the results for the original and exponentially windowed BRIRs were indistinguishable. The only analysis for which the exponential window had a noticeable impact was on the summary statistic reported in Fig. 4; see footnote 5.

⁴Comparisons between *pseudo-anechoic* results and theoretical computations (see Kopco and Shinn-Cunningham, 2003) suggest that the radiation pattern of the speaker has little effect on the direct-sound impulse responses. Published specifications (<http://www.bose.com/pro/dd/product/main.html>) detail the frequency-specific polar radiation patterns for the Bose speaker used. At 8 kHz, the radiation gain is -6 dB relative to the on-axis gain for a direction 40° off axis; at 16 kHz, the -6 dB point occurs

at 20°. Because of this speaker directivity, the direct-to-reverberant energy ratio in our measurements is greater than would be obtained with an omnidirectional point source. However, most natural sources have some directivity; the goal of the current analyses is to explore how source and listener location influence results, not to quantify how reverberation impacts the signals reaching a listener in a room when the source is an ideal point source.

⁵The inclusion of electro-acoustical noise in the late portions of the BRIRs theoretically increased the estimated mean frequency-to-frequency fluctuations in all of the magnitude spectra, but only had a measurable effect on results for left-ear BRIRs for sources at 0.15 m. When an exponential decay was imposed on the late portions of the BRIRs, the mean frequency-to-frequency fluctuations in the left-ear, 0.15-m BRIRs decreased, particularly for sources to the right of the listener (see footnote 3). The maximum spectral fluctuations observed in the left-ear BRIRs for the nearest sources was essentially equal to that observed for the 0.40-m sources, with a maximum between-sample fluctuation for sources at azimuth 90° of approximately 1.1 dB/Hz for *center* and *back* conditions and 0.7 dB/Hz for the *ear* and *corner* conditions. For all other source locations, imposing an exponential decay in the late portion of the BRIR had no noticeable effect.

⁶The broadband cross-correlation analysis shown here essentially predicts the ITD that would be extracted by energy-weighting the narrow-band cross-correlation functions assumed to be available in the brainstem and reporting the peak of this energy-weighted average. However, the peak in the broadband cross-correlation does not always equal the average of the peaks of the narrow-band cross-correlation functions. In other words, this analysis can produce results different than if one first computed the best ITD estimate within a band and then combined these ITD estimates across frequency.

Algazi, V. R., Avendano, C., and Duda, R. O. (1999). "Low-frequency ILD elevation cues," *J. Acoust. Soc. Am.* **106**, 2237.

Allen, J. B., and Berkley, D. A. (1979). "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.* **65**, 943–950.

Asano, F., Suzuki, Y., and Sone, T. (1990). "Role of spectral cues in median plane localization," *J. Acoust. Soc. Am.* **88**, 159–168.

Barron, M. (2001). "Late lateral energy fractions and the envelopment question in concert halls," *Appl. Acoust.* **62**(2), 185–202.

Bech, S. (1998a). "Spatial aspects of reproduced sound in small rooms," *J. Acoust. Soc. Am.* **103**, 434–445.

Bech, S. (1998b). "Timbral aspects of reproduced sound in small rooms. I," *J. Acoust. Soc. Am.* **97**, 1717–1726.

Begault, D. R., Wenzel, E. M., Lee, A. S., and Anderson, M. R. (2001). "Direct comparison of the impact of head-tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *J. Audio Eng. Soc.* **49**(10), 904–916.

Berkley, D. A. (1980). "Normal listeners in typical rooms: Reverberation perception, simulation, and reduction," in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G. A. Studebaker and I. Hochberg (University Park, Baltimore, MD), pp. 3–24.

Bistafa, S. R., and Bradley, J. S. (2000). "Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics," *J. Acoust. Soc. Am.* **107**, 861–875.

Bradley, J. S. (1986). "Speech intelligibility studies in classrooms," *J. Acoust. Soc. Am.* **80**, 846–854.

Bradley, J. S. (1996). "Some effects of orchestra shells," *J. Acoust. Soc. Am.* **100**, 889–898.

Bradley, J. S., Reich, R. D., and Norcross, S. G. (1999). "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility," *J. Acoust. Soc. Am.* **106**, 1820–1828.

Bradley, J. S., Sato, H., and Picard, M. (2003). "On the importance of early reflections for speech in rooms," *J. Acoust. Soc. Am.* **113**, 3233–3244.

Bradley, J. S., and Soudoude, G. A. (1995). "The influence of late arriving energy on spatial impression," *J. Acoust. Soc. Am.* **97**, 2263–2271.

Bronkhorst, A. W. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acustica* **86**, 117–128.

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

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

Brown, T. J. (2000). "Characterization of Acoustic Head-Related Transfer Functions for Nearby Sources," in *Electrical Engineering and Computer Science* (MIT, Cambridge, MA).

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

Brungart, D. S., and Durlach, N. I. (1999). "Auditory localization of nearby sources. II. Localization of a broadband source in the near field," *J. Acoust. Soc. Am.* **106**, 1956–1968.

Brungart, D. S., and Rabinowitz, W. M. (1999). "Auditory localization of nearby sources I: Head-related transfer functions," *J. Acoust. Soc. Am.* **106**, 1465–1479.

Buell, T. N., and Hafter, E. R. (1991). "Combination of binaural information across frequency bands," *J. Acoust. Soc. Am.* **90**, 1894–1900.

Butler, R. A., and Belendiuk, K. (1977). "Spectral cues utilized in the localization of sound in the median saggital plane," *J. Acoust. Soc. Am.* **61**, 1264–1269.

Butler, R. A., and Humanski, R. A. (1992). "Localization of sound in the vertical plane with and without high-frequency spectral cues," *Percept. Psychophys.* **51**, 182–186.

Culling, J. F., Hodder, K. I., and Toh, C. Y. (2003). "Effects of reverberation on perceptual segregation of competing voices," *J. Acoust. Soc. Am.* **114**, 2871–2876.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (1994). "Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels," *Speech Commun.* **14**, 71–95.

Darwin, C. J., and Hukin, R. W. (2000). "Effects of reverberation on spatial, prosodic, and vocaltract size cues to selective attention," *J. Acoust. Soc. Am.* **108**, 335–342.

de Vries, D., Hulsebos, E. M., and Baan, J. (2001). "Spatial fluctuations in measures for spaciousness," *J. Acoust. Soc. Am.* **110**, 947–954.

Devore, S., and Shinn-Cunningham, B. G. (2003). "Perceptual consequences of including reverberation in spatial auditory displays," *International Conference on Auditory Displays*, Boston, MA.

Duda, R. O., and Martens, W. L. (1998). "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.* **104**, 3048–3058.

Durlach, N. I., Rigapulos, A., Pang, X. D., Woods, W. S., Kulkarni, A., Colburn, H. S., and Wenzel, E. M. (1992). "On the externalization of auditory images," *Presence* **1**, 251–257.

Ebata, M. (2003). "Spatial unmasking and attention related to the cocktail party problem," *Acoust. Sci. Technol.* **24**(3), 208–219.

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

Hartmann, W. M., and Rakerd, B. (1999). "Localization of sound in reverberant spaces," *J. Acoust. Soc. Am.* **105**, 1149.

Hartmann, W. M., Rakerd, B., Gaalaas, J. B., Vander Velde, T., Thorpe, W. R., and Oh, M. M. (1999). "Localization of sound in rooms: Broadband noise," Michigan State University.

Hartmann, W. M., Rakerd, B., and Koller, A. (2005). "Binaural coherence in rooms," *Acust. Acta Acust.* (to be published).

Hidaka, T., and Beranek, L. L. (2000). "Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas," *J. Acoust. Soc. Am.* **107**, 368–383.

Hodgson, M. (1999). "Experimental investigation of the acoustical characteristics of university classrooms," *J. Acoust. Soc. Am.* **106**, 1810–1819.

Kidd, Jr., G., Mason, C. R., Brughera, A., and Hartmann, W. M. (2005). "The role of reverberation in release from masking due to spatial separation of sources for speech identification," *Acust. Acta Acust.* in press.

Kistler, D. J., and Wightman, F. L. (1991). "A model of head-related transfer functions based on principle components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.* **91**, 1637–1647.

Kleiner, M., Dalenback, B.-I., and Svensson, P. (1993). "Auralization—An overview," *J. Audio Eng. Soc.* **41**, 861–875.

Kopco, N., and Shinn-Cunningham, B. G. (2002). *Auditory localization in rooms: Acoustic analysis and behavior*, 32nd International Acoustics Conference—EAA Symposium, Zvolen, Slovakia.

Kopco, N., and Shinn-Cunningham, B. G. (2003). "Spatial unmasking of nearby pure-tone targets in a simulated anechoic environment," *J. Acoust. Soc. Am.* **114**, 2856–2870.

Kulkarni, A., and Colburn, H. S. (2004). "Infinite-impulse-response models of the head-related transfer function," *J. Acoust. Soc. Am.* **115**, 1714–1728.

Macpherson, E. A., and Middlebrooks, J. C. (2002). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* **111**, 2219–2236.

Mershon, D. H., Ballenger, W. L., Little, A. D., McMurtry, P. L., and Bucha-

- nan, J. L. (1989). "Effects of room reflectance and background noise on perceived auditory distance," *Perception* **18**, 403–416.
- Middlebrooks, J. C. (1999). "Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency," *J. Acoust. Soc. Am.* **106**, 1493–1510.
- Middlebrooks, J. C., and Green, D. M. (1991). "Sound localization by human listeners," *Annu. Rev. Psychol.* **42**, 135–159.
- Morimoto, M. (2001). "The contribution of two ears to the perception of vertical angle in sagittal planes," *J. Acoust. Soc. Am.* **109**, 1596–1603.
- Nabelek, A. K., Letowski, T. R., and Tucker, F. M. (1989). "Reverberant overlap- and self-masking in consonant identification," *J. Acoust. Soc. Am.* **86**, 1259–1265.
- Naguib, M. (1995). "Auditory distance assessment of singing conspecifics in Carolina wrens: The role of reverberation and frequency-dependent attenuation," *Anim. Behav.* **50**(5), 1297–1307.
- Nishihara, N., Hidaka, T., and Beranek, L. L. (2001). "Mechanism of sound absorption by seated audience in halls," *J. Acoust. Soc. Am.* **110**, 2398–2411.
- Okano, T. (2002). "Judgments of noticeable differences in sound fields of concert halls caused by intensity variations in early reflections," *J. Acoust. Soc. Am.* **111**, 217–229.
- Plomp, R. (1976). "Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise)," *Acustica* **34**, 200–211.
- Rife, D. D., and Vanderkooy, J. (1989). "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc.* **6**, 419–444.
- Santarelli, S. (2000). "Auditory Localization of Nearby Sources in Anechoic and Reverberant Environments," in *Cognitive and Neural Systems* (Boston Univ., Boston, MA).
- Schroeder, M. R. (1965). "New method of measuring reverberation time," *J. Acoust. Soc. Am.* **37**, 409–412.
- Schroeder, M. R. (1987). "Statistical parameters of the frequency response curves of large rooms," *J. Audio Eng. Soc.* **35**(5), 299–305.
- Schroeder, M. R., and Kuttruff, K. H. (1962). "On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima," *J. Acoust. Soc. Am.* **34**, 76–80.
- Shaw, E. A. G. (1997). "Acoustical features of the human external ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. Anderson (Erlbaum, New York), pp. 25–48.
- Shinn-Cunningham, B. (2004). "The perceptual consequences of creating a realistic, reverberant 3-D audio display," International Congress on Acoustics, Kyoto, Japan.
- Shinn-Cunningham, B., and Kawakyu, K. (2003). "Neural representation of source direction in reverberant space," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY.
- Shinn-Cunningham, B. G. (2000a). "Distance cues for virtual auditory space," IEEE-PCM 2000, Sydney, Australia.
- Shinn-Cunningham, B. G. (2000b). "Learning reverberation: Implications for spatial auditory displays," International Conference on Auditory Displays, Atlanta, GA.
- Shinn-Cunningham, B. G. (2002). "Speech intelligibility, spatial unmasking, and realism in reverberant spatial auditory displays," International Conference on Auditory Displays, Kyoto, Japan.
- Shinn-Cunningham, B. G., and Ram, S. (2003). "Identifying where you are in a room: Sensitivity to room acoustics," International Conference on Auditory Display, Boston, MA.
- Shinn-Cunningham, B. G., Constant, S., and Kopco, N. (2002). "Spatial unmasking of speech in simulated anechoic and reverberant rooms," 25th mid-Winter meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL.
- Shinn-Cunningham, B. G., Santarelli, S., and Kopco, N. (2000). "Tori of confusion: binaural localization cues for sources within reach of a listener," *J. Acoust. Soc. Am.* **107**, 1627–1636.
- Shinn-Cunningham, B. G., Schickler, J., Kopco, N., and Litovsky, R. (2001). "Spatial unmasking of nearby speech sources in a simulated anechoic environment," *J. Acoust. Soc. Am.* **110**, 1118–1129.
- Stern, R. M., and Trahiotis, C. (1995). "Models of binaural interaction," in *Hearing*, edited by B. C. J. Moore (Academic, San Diego), pp. 347–386.
- Torres, R., Svensson, U. P., and Kleiner, M. (2001). "Computation of edge diffraction for more accurate room auralization," *J. Acoust. Soc. Am.* **109**, 600–610.
- Trahiotis, C., and Stern, R. M. (1989). "Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase," *J. Acoust. Soc. Am.* **86**, 1285–1293.
- Vanderkooy, J. (1994). "Aspects of MLS measuring systems," *J. Audio Eng. Soc.* **42**(4), 219–231.
- Vliegen, J., and van Opstal, A. J. (2004). "The influence of duration and level on human sound localization," *J. Acoust. Soc. Am.* **115**, 1705–1713.
- Wightman, F. L., and Kistler, D. J. (1989). "Headphone simulation of free-field listening. I. Stimulus synthesis," *J. Acoust. Soc. Am.* **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1997). "Monaural sound localization revisited," *J. Acoust. Soc. Am.* **101**, 1050–1063.
- 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.
- Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G. Studebaker and I. Hochberg (College-Hill, Boston, MA).
- Zurek, P. M., Freyman, R. L., and Balakrishnan, U. (2004). "Auditory target detection in reverberation," *J. Acoust. Soc. Am.* **115**, 1609–1620.