Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited[a]

Ewan A. Macpherson[b] and John C. Middlebrooks
Kresge Hearing Research Institute, University of Michigan, 1301 East Ann Street, Ann Arbor, Michigan 48109-0506

(Received 24 September 2001; accepted for publication 26 February 2002)

The virtual auditory space technique was used to quantify the relative strengths of interaural time difference (ITD), interaural level difference (ILD), and spectral cues in determining the perceived lateral angle of wideband, low-pass, and high-pass noise bursts. Listeners reported the apparent locations of virtual targets that were presented over headphones and filtered with listeners' own directional transfer functions. The stimuli were manipulated by delaying or attenuating the signal to one ear (by up to 600 μs or 20 dB) or by altering the spectral cues at one or both ears. Listener weighting of the manipulated cues was determined by examining the resulting localization response biases. In accordance with the Duplex Theory defined for pure-tones, listeners gave high weight to ITD and low weight to ILD for low-pass stimuli, and high weight to ILD for high-pass stimuli. Most (but not all) listeners gave low weight to ITD for high-pass stimuli. This weight could be increased by amplitude-modulating the stimuli or reduced by lengthening stimulus onsets. For wideband stimuli, the ITD weight was greater than or equal to that given to ILD. Manipulations of monaural spectral cues and the interaural level spectrum had little influence on lateral angle judgements. [DOI: 10.1121/1.1471898]

PACS numbers: 43.66.Qp [LRB]

I. INTRODUCTION

The locations of sound sources are not mapped directly at the sensory periphery. Instead, locations must be derived by combining acoustical cues that result from the interaction of incident sound waves with the external ears, head, and upper body. The acoustical cues for sound localization were explored as early as the end of the 19th century by the British physicist Lord Rayleigh (Strutt, 1907), among others. Rayleigh worked primarily with pure-tone stimuli produced by vibrating tuning forks. He determined that the primary cue to the lateral positions of sources with frequencies >500 Hz was the interaural difference in sound pressure levels (ILDs) resulting from acoustic shadowing by the head. At lower frequencies, however, the wavelength of sound is much larger than the diameter of the head, and ILDs are negligible. Through the ingenious use of a pair of mistuned low-frequency tuning forks, Rayleigh demonstrated compellingly that human listeners are sensitive to interaural differences in the ongoing phase of low-frequency sounds and, thus, that interaural time differences (ITDs) could provide cues to the lateral positions of low-frequency sources. Rayleigh's understanding of the localization of tones in the lateral dimension has come to be known as the "Duplex Theory" of sound localization and has been substantiated in numerous psychophysical and physiological studies. Rayleigh also appreciated that subjects could not discriminate the front-versus-back locations of pure-tone stimuli, but that such front/back discrimination was possible for "sounds of other character" (i.e., sounds with broader bandwidths). Research in the past few decades has filled in some understanding of the cues for front/back and vertical localization, revealing the importance of spectral-shape cues provided by the direction-dependent filtering of broadband sounds by the external ears.

The Rayleigh Duplex Theory is quite satisfactory to explain left/right localization of tonal stimuli. Nevertheless, in the real world most sounds have bandwidths of several octaves, and a listener rarely is exposed to a pure tone. Three sets of observations led us to revisit the Duplex Theory in the context of complex (i.e., broadband) sounds.

First, lateralization studies have shown than listeners are sensitive to ITDs in high-frequency complex sounds. Sensitivity to ongoing time differences in simple sounds such as pure tones is limited to low frequencies by the loss of phase-locking in the auditory nerve at high frequencies and likely also by a lower-frequency cutoff in the binaural system. Indeed, psychophysical studies show that sensitivity to ongoing ITD is limited to frequencies below ~1.3 kHz for pure tones (Zwislocki and Feldman, 1956). Nevertheless, the auditory system can extract timing information from the envelopes of higher-frequency sounds that contain multiple frequency components, and listeners can detect ITDs in high-frequency complex sounds presented through headphones (e.g., Henning, 1974; Leakey *et al.*, 1958; McFadden and Pasanen, 1976).

Second, Wightman and Kistler (1972, 1997b) have investigated localization cues using virtual auditory space (VAS) techniques, which permit more-or-less independent manipulation of ITD, ILD, and spectral-shape cues. They demonstrated that ITDs dominate listeners' judgments of the location of broadband sound sources that contain low-

[a]Portions of this work were previously presented in poster form at the 24th Annual Midwinter Research Meeting of the Association for Research in Otolaryngology in February 2001.

[b]Electronic mail: emacpher@umich.edu

frequency components (Wightman and Kistler, 1992). In some cases, the influence of ITDs persisted even when stimulus spectra were limited to high frequencies, although the effect of high-pass filtering varied widely among listeners. Also, imposition of an interaural imbalance (i.e., an ILD of 10–20 dB) had surprisingly little impact on lateral location judgments by some listeners (Wightman and Kistler, 1997b).

Third, spectral-shape cues, which provide essential cues to the vertical and front/back location of broadband sounds, also vary with source azimuth and might contribute to the judgment of lateral localization. For instance, some congenitally monaural listeners, whose only cues to sound-source location derive from direction-dependent filtering by one external ear, show reasonably accurate localization in the horizontal dimension (Slattery and Middlebrooks, 1994). Whether normal binaural listeners similarly use spectral cues in determining lateral angle is unknown.

These three sets of observations raise questions about the applicability of the Rayleigh duplex model to localization of naturally occurring broadband sounds. Might ITDs contribute to or even dominate localization of high-frequency complex sounds? How salient are spectral-shape cues compared to interaural difference (i.e., ITD and ILD) cues? Although ILDs have an obvious impact on lateralization of high-frequency tones, do ILDs have any influence on localization of broadband sounds?

We addressed these questions by measuring localization of broadband, low-pass, and high-pass sounds presented using VAS techniques. We quantified the degrees to which cues addressed by the Duplex Theory (such as low-frequency ITD and high-frequency ILD) and other cues (such as spectral cues and high-frequency, envelope-based ITD) contributed to the lateral component of listeners' localization judgments. Apart from some influence of envelope-based ITDs in high-pass signals, we found that the localization of complex sounds in the lateral dimension accorded nicely with a more-than-century-old theory based on the localization of vibrating tuning forks.

II. METHODS

A. Subjects

Thirteen paid listeners (five female and eight male, ages 18–35 years, including the first author) were recruited from the student body of the University of Michigan and the staff of the Kresge Hearing Research Institute. All had extensive experience in free-field localization experiments. Only one (the first author, S18) had previous experience in localizing virtual auditory stimuli presented over headphones, but the others received several hours of practice in the virtual free-field localization task described in Sec. II D prior to the beginning of the present study. These practice sessions involved the localization of virtual broadband noise-burst targets. No feedback was provided. All listeners had normal hearing as defined by standard audiometric testing. Ten of the listeners (not including S18) participated in Experiments I and II-A. Two of these ten plus S18 and two additional listeners participated in Experiment II-B. Three of the ten plus S18 participated in Experiment III, and three from Experiment II-B participated in Experiment IV.

B. Directional transfer function measurements

In order to compute the filters necessary to synthesize the virtual free-field stimuli, measurements were made of each listener's directional transfer functions (DTFs). The details of this procedure are given by Middlebrooks (1999a). Briefly, 512-point, 50-kHz Golay codes (Zhou *et al.*, 1992) were presented from a loudspeaker positioned 1.2 m from the listener's head at 400 locations approximately evenly distributed in space around the listener's head. The responses to these excitation signals were recorded simultaneously by two miniature electret microphones (Knowles, model 1934) inserted approximately 5 mm into the listeners' ear canals. Head-related transfer functions (HRTFs) were extracted by cross-correlation of excitation and response, transformation to the frequency domain, and division by the previously measured loudspeaker transfer function. For each ear, DTFs were computed from the HRTFs by dividing by a complex common component computed for the set of HRTFs for that ear (Middlebrooks, 1999a). This nondirectional component was a combination of the ear canal resonance and the diffuse-field average response at the ear canal entrance. The DTFs were transformed into the time domain, yielding a set of directional impulse responses (DIRs) for each listener. The DIRs were used in the synthesis of the virtual free-field targets as described below.

C. Stimulus synthesis

Stimulus noise waveforms were computed on an Intel-based desktop personal computer using custom MATLAB scripts (The Mathworks). An inverse-Fourier-transform method was used to produce flat-spectrum, random-phase noise waveforms with the desired passband and duration sampled at 50 kHz. Raised-cosine (i.e., $\cos^2$) ramps of 1, 20, or 50 ms duration (depending on the stimulus condition) were applied to the onsets and offsets. The resulting waveform was convolved with the right- and left-ear DIRs corresponding to the desired target location. The stimuli were presented over ''diffuse-field equalized'' circumaural headphones (Sennheiser HD 265) at a sampling rate of 50 kHz using digital-to-analog converters, attenuators, and headphone amplifiers from Tucker-Davis Technologies (models DD1, PA4, and HB6, respectively). The stimulus level on each trial was equivalent to that of a free-field source at a sound pressure level of approximately 65 dB.

We did not attempt a rigorous equalization of the headphone response. Rather, the headphone response itself restored an approximation of the diffuse-field component removed in the computation of the DIRs, and the listener's own ear canal restored the ear canal resonance. We have discussed this approach previously and have shown that listeners can localize accurately in the VAS generated by this method (Middlebrooks, 1999b).

In Experiments I and II, the interaural time difference (ITD) and interaural level difference (ILD) cues naturally

TABLE I. Summary of cue manipulations for Experiments I–IV. The cues referred to are interaural time difference or interaural phase spectrum (ITD/IPS), interaural level difference or interaural level spectrum (ILD/ILS), and the directional transfer functions for the ears nearer to and farther from the source (DTF_{near} and DTF_{far}). Symbols indicate whether the cue corresponded to the original target location (□), to a biased location (■), or (in the case of the artificial DTFs used in Experiments III and IV) to no actual location (⋆).

Cue	Experiments I and II		Experiment III ILS bias	Experiment IV DTF_{near}, bias
ITD/IPS	■	□	□	□
ILD/ILS	□	■	■	□
DTF_{near}	□	□	□	■
DTF_{far}	□	□	⋆	⋆

present in the stimulus were manipulated by imposing a whole-waveform delay or attenuation on the signal at one of the ears. We refer to this procedure as *biasing* the stimulus and to the amount of ITD or ILD offset as the *imposed bias*. In Experiments III and IV, the stimuli were manipulated by modifying the DTF spectrum at one or both ears in order to bias the interaural level spectrum (ILS) or the DTF of the ear nearest the source. The cue manipulations for Experiments I–IV are summarized in Table I.

D. Localization procedure

The localization procedure was similar to that described by Middlebrooks (1999b). Listeners stood in the center of a darkened anechoic chamber, and at the beginning of each trial oriented towards a light-emitting diode (the centering LED) positioned at eye level 2 m directly in front of the listener. A trial was initiated by pressing a hand-held button. The centering LED was extinguished, and the listener's initial head position was measured by a head-mounted electromagnetic tracking device (Polhemus FASTRAK). Following a delay of 500 ms, the stimulus was presented over headphones. After hearing the virtual free-field stimulus, the listener oriented towards its perceived location, at which time a second button press triggered a second measurement of head orientation. If the initial head orientation deviated by more than 10 degrees in either azimuth or elevation from the centering LED, the trial data were discarded. Depending upon the condition, runs consisted of 79, 99, or 111 trials, and the longest were typically completed in 10–12 min. Listeners rested after every two to three runs, and typically completed six to eight runs in one session.

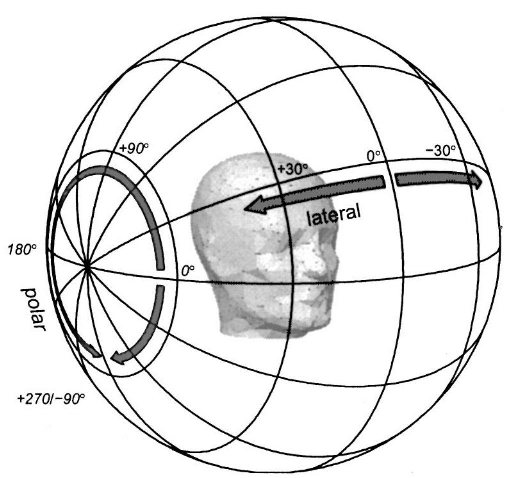

FIG. 1. Horizontal polar coordinate system. Lateral angle is the angle between the location and the median sagittal plane; positive values are to the listener's right. Polar angle combines elevation and front/back position. −90 degrees: below; 0 degrees: front; +90 degrees: overhead; +180 degrees: rear.

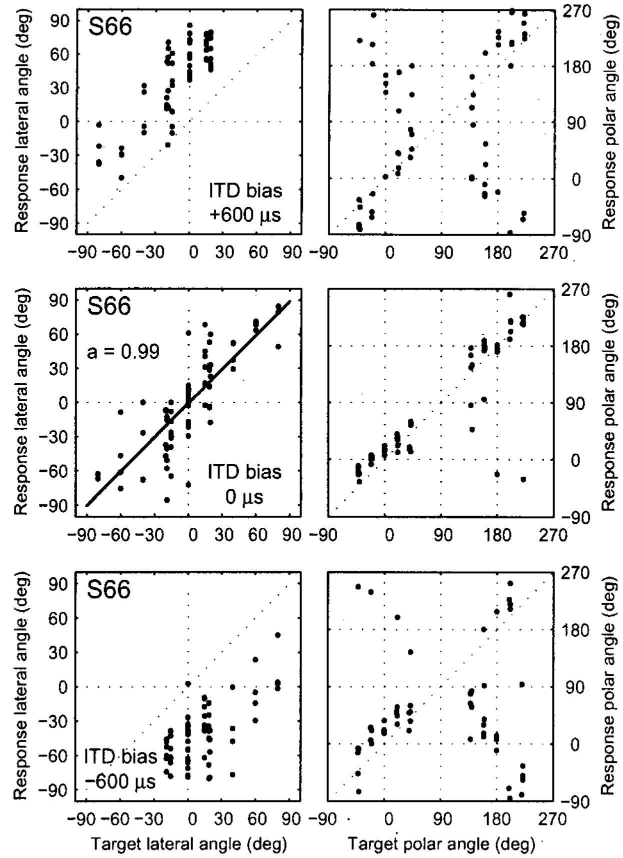

FIG. 2. Sample lateral- (left column) and polar-angle (right column) response data for ITD-biased wideband targets for listener S66. ITD bias: +600 μs (top), 0 μs (middle), −600 μs (bottom). Unbiased lateral angle gain, a, was computed as the slope of a linear fit to the unbiased stimulus lateral-angle responses (solid line). In the polar angle plots, accurate responses fell near the positive diagonal, and front/back reversed responses fell near the negative diagonal.

E. Data analysis

1. Lateral-polar coordinate system

Target and response location data were converted from the vertical polar (azimuth and elevation) coordinate system to the horizontal polar (lateral- and polar-angle) coordinate system (Fig. 1). Only the lateral angle data were analyzed in detail. Sample lateral and polar angle response data for wideband, ITD-biased stimuli are shown in Fig. 2. In this example, the ITD was biased by +600 μs (top row), 0 μs (middle row), or −600 μs (bottom row). Note that for these stimuli, positive (right-leading) and negative (left-leading) ITD biases shifted the listener's responses consistently towards the right and left sides, respectively. For this listener, applying ITD bias increased the rate of front/back confusions (more responses fell in the upper-left or lower-right quadrants), but the elevation component of the listener's polar-angle judgments remained reasonably accurate (i.e., responses fell near the positive or negative diagonals). The effect of bias on polar-angle responses is discussed further in Sec. III B 2.

2. Unbiased lateral angle gain

In order to reveal any detrimental effect on lateral angle localization caused by the restriction of stimulus bandwidth, an *unbiased lateral angle gain* (Hofman and Van Opstal, 1998; Macpherson and Middlebrooks, 2000) was computed for responses to the filtered, but unmanipulated (unbiased), targets in each stimulus set. The unbiased lateral angle gain was simply the slope of a linear fit to the target and response lateral angle data (Fig. 2).

3. ITD and ILD bias weights

We wished to assess the weighting or salience of the manipulated cue (ITD or ILD) in a manner which would permit meaningful comparisons between weights. Neither the classical time-intensity trading ratio (in units of μs/dB) nor ratios between angular displacements of judgments and magnitudes of imposed cue bias (yielding values in degrees/μs and degrees/dB) are useful in isolation in describing the relative effectiveness of ITD and ILD in localization. To use such values to make comparisons between the weighting of different cues, reference must also be made both to the naturally occurring relations between the physical cues and to the spatial disposition of the cues.

Our strategy was to derive *dimensionless* weights relating bias in responses to imposed bias in the underlying cue by (a) measuring the correspondence between the physical cues and lateral angle, and (b) using this relation to convert any shift in the angular response to a quantity expressed in the units of the manipulated cue itself.

As an example, consider a target location for which the natural ILD is 5 dB. If an ILD bias of 10 dB is imposed, the ILD in the presented stimulus will be 15 dB. If the listener responds at a location for which the natural ILD were also 5 dB (even if that were at a different lateral angle), we would conclude that the imposed bias had no effect, the observed bias would be 0 dB, and we would estimate the perceptual weighting of ILD to be close to 0. Conversely, if the listener responds at a location for which the natural ILD were 15 dB, we would conclude that the imposed bias was fully effective in shifting the perceived lateral angle, the observed bias would be 10 dB (equal to the imposed bias), and we would estimate the ILD bias weight to be close to 1. Responses at locations with intermediate ILDs would yield intermediate weights.

This conversion of bias in response location to bias in an underlying physical cue and the resulting derivation of a dimensionless weight permits the computation and comparison of perceptual weights on the ITD and ILD cues without concern about the physical correspondence of, or the auditory system's differential sensitivity to, the ITD and ILD cues. In addition, by incorporating the relation between cue and lateral angle into the procedure, the exact (and possibly nonmonotonic) relation between the interaural cues and lateral angle was rendered unimportant.

The following procedure was used to compute the bias weights. *First*, the value of the physical interaural cue at each of the 400 measured DTF locations was computed for each listener using procedures similar to those of Gaik (1993). For ILD, the energies in the right and left ear directional transfer functions were integrated over the stimulus passband (*wideband*, 0.5–16 kHz; *low-pass*, 0.5–2 kHz; *high-pass*, 4–16 kHz; see Secs. II C and III A) and their ratio represented in dB such that positive ILDs corresponded to higher intensity at the right ear.

To compute ITD, the DIR for each ear was passed through a gammatone filter bank (Slaney, 1994) with low-frequency channels at 600, 700, and 800 Hz, and high-frequency channels at 4, 4.5, and 5 kHz. These center frequencies were chosen because they produced the smoothest ITD-versus-azimuth functions across listeners. For the low-frequency channels, the ITD was taken from the lag of the peak in the cross-correlation of the left and right ear signals. For the high-frequency channels, the envelopes of the filter outputs were extracted using a Hilbert transform prior to cross-correlation in order to extract the group delay. This paralleled the loss of phase-locking and the onset of envelope following in the auditory nerve at high frequencies. If multiple peaks appeared in the cross-correlation, the one closest to the predicted ITD based on a spherical head model (Kuhn, 1987) was chosen. If no peak was found within 250 μs of the predicted value, the mean of the computed ITDs for neighboring locations was used. The median of the ITDs in the low-frequency channels was used as the ITD cue for low-pass and wideband stimuli, and the median of the high-frequency channel ITDs was used as cue for high-pass stimuli. Note that this process was intended as a means of measuring the physical ITD, not an attempt to model the extraction of ITD information by the auditory system.

Next, having associated a frequency-dependent ITD and ILD with each DTF location, an *observed cue bias* (ITD or ILD) was computed for each localization response in the manner illustrated in Fig. 3. The *natural cue*, c_{nat} (in μs or dB), was that present in the unmodified DTFs used in synthesizing the spatialized stimulus. The *observed cue*, c_{obs}, was that associated with the DTF location closest to the listener's response location. The observed cue bias was the difference between these values; bias$=c_{\mathrm{obs}}-c_{\mathrm{nat}}$. Trials in which the manipulated cue exceeded the range of the listener's measured ITD or ILD values were eliminated from the analysis. Overall, 10%–15% of trials were discarded for this reason.

Finally, the listener's weighting of the manipulated interaural cue was computed as the slope of the linear regression between the observed cue bias and the *imposed cue bias* (the magnitude of the added ITD, in μs, or ILD, in dB), as illustrated in Fig. 4. We refer to these dimensionless values as the $\mu s/\mu s$ and dB/dB weights. The standard error of the regression coefficient was taken as a measure of the uncertainty in the computed weight. If the manipulated cue had little influence on the listener's response, the response loca-

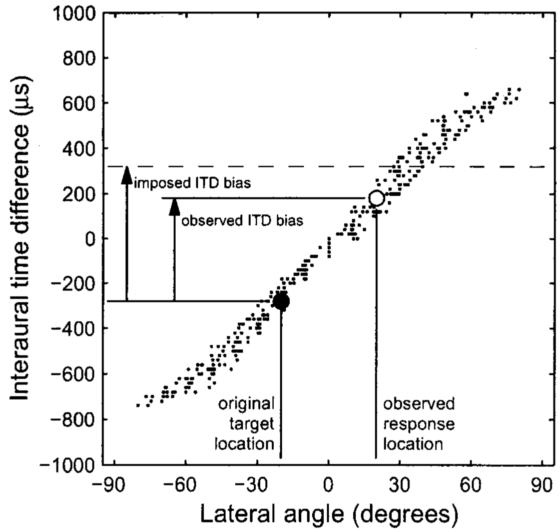

FIG. 3. Computation of observed cue bias. Small symbols show the measured low-frequency ITD for listener S92 as a function of lateral angle. Each small symbol represents a distinct DTF measurement location (unique lateral angle and polar angle combination). The large filled circle indicates the lateral angle and natural ITD of an example original target location. An ITD bias of $+600$ μs was applied. The large open circle indicates the lateral angle and natural ITD of the DTF measurement location closest to the listener's response (i.e., the observed ITD). The observed cue bias is the signed difference between the observed ITD and the natural target ITD.

tion was close to the original target location, the observed cue bias was close to zero in all trials, and the cue weight, W_{ITD} or W_{ILD}, was also close to zero (Fig. 4, upper-right and lower-left panels). Conversely, if the listener derived the judgment of lateral angle primarily from the manipulated cue, then the response was expected to lie at a location for which the natural cue was similar to the stimulus cue value. In such cases, the cue weight was close to 1 (Fig. 4, upper-left and lower-right panels).

III. EXPERIMENT I: WEIGHTING OF ITD AND ILD CUES IN LOW-PASS, HIGH-PASS, AND WIDEBAND NOISE

A. Stimuli and locations

In Experiment I, we measured listeners' weighting of the ITD and ILD cues to lateral angle under three passband conditions: *wideband*, 0.5–16 kHz; *low-pass* 0.5–2 kHz; and *high-pass*, 4–16 kHz. The target stimuli were 100-ms noise bursts with 1-ms raised-cosine onsets and offsets. Each stimulus set contained interleaved stimuli from four classes: (1) unfiltered (i.e., wideband) noise bursts, (2) filtered, unmanipulated (i.e., no imposed ITD or ILD bias) noise bursts, (3) filtered noise bursts with *medium* imposed cue bias (± 300 μs ITD or ± 10 dB ILD), and (4) filtered noise bursts with *large* imposed cue bias (± 600 μs ITD or ± 20 dB ILD). Some listeners also completed an additional set of ILD-bias conditions in which the medium and large biases of 10 and 20 dB were replaced by 4- and 8-dB biases, respectively. Analysis of these data showed that the computed weights were insensitive to which range of ILD biases was used, and the 4- and 8-dB ILD-bias data were not included in the following analysis.

Unfiltered targets, 36 in all, were placed at 10-degree increments in azimuth from -170 to $+180$ degrees with elevations of -30 or $+30$ degrees. Of the filtered, unbiased targets, 24 were placed at azimuths 0, ± 20, ± 160, and 180 degrees with elevations of ± 20 and ± 40 degrees (i.e., on or near the median plane, both in the front and rear, and above and below the horizontal plane). An additional 18 were

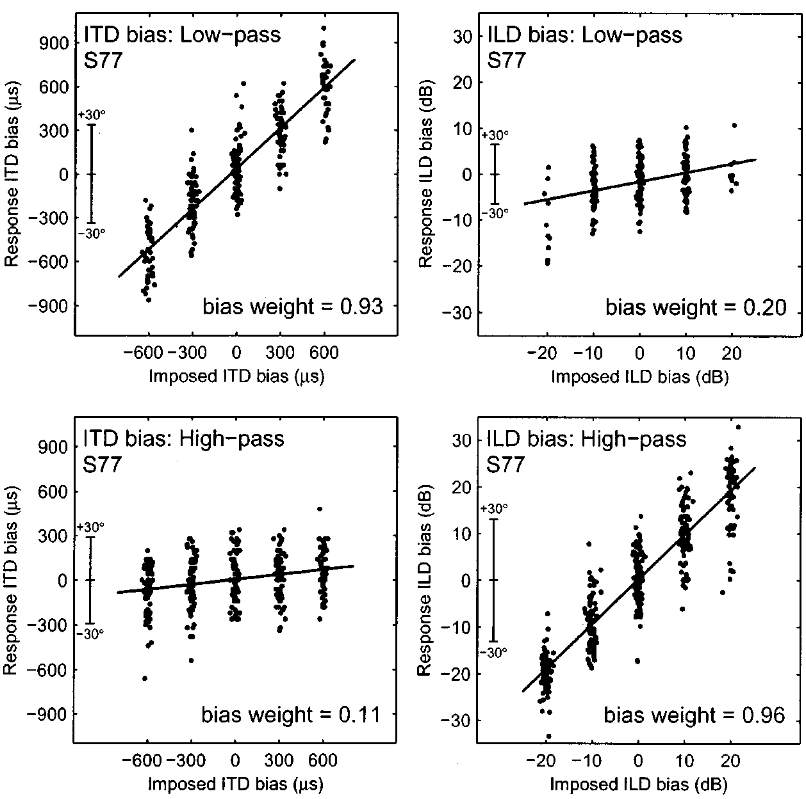

FIG. 4. Illustration of cue bias weight computation. Observed ITD and ILD cue biases are plotted against imposed cue bias for listener S77 in the low-pass (0.5–2 kHz) and high-pass (4–16 kHz) passband conditions. Upper-left: low-pass ITD-bias condition (bias=0, ± 300, ± 600 μs); upper-right: low-pass ILD bias condition (bias=0, ± 10, ± 20 dB); lower-left: high-pass ITD-bias condition; lower-right: high-pass ILD bias condition. The cue bias weight was the slope of a linear fit to these data. The ± 30 degrees scale near the vertical axis indicates the natural variation of the manipulated cue within 30 degrees of the midline.

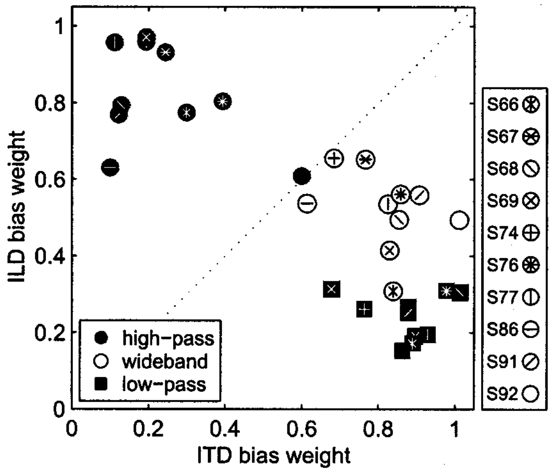

FIG. 5. Measured ITD- and ILD-bias weights for Experiment I. ILD-bias weight (vertical axis) is plotted against ITD-bias weight (horizontal axis) for each listener in the wideband, low-pass, and high-pass passband conditions.

placed in 20-degree increments of azimuth around the horizontal plane, for a total of 42 locations. To minimize the presentation of stimuli with ITD or ILD cues outside of the physiological range, the medium bias was not applied to targets more than 40 degrees away from the median plane on the side toward which the bias was applied, leaving 38 locations. Similarly, the large bias was not applied to targets more than 20 degrees from the median plane on the biased side, leaving 34 locations. In total there were six stimulus sets (3 passbands$\times$2 cue-bias types) of 222 targets each. The targets in each set were presented twice in a shuffled order over the course of four blocks of 111 trials each. Blocks from different stimulus sets were intermixed within an experimental session but trials from different sets were not intermixed within blocks.

B. Results

1. Lateral angle responses

For each listener, the unbiased lateral angle gain was computed for the unmanipulated stimuli of each passband, and the cue weight was computed for each bias and passband condition. The lateral angle gain was unaffected by stimulus passband, and was close to unity for all listeners except for listener S69, who showed a mild decrement in lateral gain in the low-pass condition. Similar, near-unity, unbiased lateral angle gains were obtained for all other listeners in all conditions of Experiments I–IV. This result indicated that low-pass or high-pass filtering did not impair listeners' ability to judge accurately the lateral angle of the unbiased virtual free-field targets, and it increased our confidence that, in the biased conditions, the listeners also had access to usable cues to lateral angle in all passband conditions.

The cue bias weights are shown in the scatter plot of Fig. 5, in which ILD-bias weight is plotted against ITD-bias weight for each of the ten listeners in each passband condition.

The computed weights for the ITD and ILD cues were always greater than 0, and only in two cases exceeded 1 (for S68, $W_{\mathrm{ITD,LP}}=1.01$; for S92, $W_{\mathrm{ITD,WB}}=1.01$). Weights less than 1 were expected because the manipulated cue was placed in opposition to all the unmanipulated cues, which corresponded to a zero-bias target location. The unmanipulated cues thus opposed or diluted the effect of the biased cue. For clarity, the computed standard errors of the weights are not shown in Fig. 5, but they ranged from 0.01 to 0.04 for both ITD and ILD. The median standard error was $\sim$0.025.

The relation between the ITD and ILD weights depended on the stimulus passband condition. In the wideband condition (open circles), ITD was weighted more heavily than ILD for all listeners (mean $W_{\mathrm{ITD,WB}}=0.82$, mean $W_{\mathrm{ILD,WB}}=0.52$), although the difference between the weights was small for listeners S67 and S86, and close to zero for S74. In the low-pass condition (filled squares), ITD weights were >0.8 for eight of the ten listeners, and were substantially higher than the ILD weights for all listeners (mean $W_{\mathrm{ITD,LP}}=0.88$, mean $W_{\mathrm{ILD,LP}}=0.24$). In the high-pass condition (filled circles), the relation between ITD and ILD weights was reversed from that observed in the low-pass case: ITD was weighted less heavily than ILD for all listeners except S92 (mean $W_{\mathrm{ITD,HP}}=0.24$, mean $W_{\mathrm{ILD,WB}}=0.82$). For all listeners, ITD weights were much lower than those observed in the wideband and low-pass conditions.

There were, however, marked individual differences in the weight given to the ITD cue in the high-pass passband condition. Six of the ten listeners (S68, S69, S74, S77, S86, S91) had small (<0.20) high-pass ITD weights, whereas the other four (S66, S67, S76, S92) had larger weights (0.24 to 0.60). Our high-pass stimuli were limited to frequencies above 4 kHz, and because the auditory system has virtually no representation of waveform fine structure at frequencies this high (Palmer and Russell, 1986), we presume that any ITD information utilized by these listeners must have been derived from the envelopes of the signals. We investigated the relative influence of the ITD cues provided by the onset and ongoing portions of the envelope in Experiment II.

2. Polar angle responses

We inspected the polar angle components of the listeners' responses under the conditions of restricted stimulus passband and imposed ITD or ILD bias. With no imposed bias, polar angle responses were similar to those observed in other free-field and virtual localization studies. In the wideband, unbiased condition, most listeners responded accurately in polar angle, but with rates of front/back confusion that varied among listeners. Errors of elevation and front/back position were most common for targets high above and behind the listener. The rate of front/back confusions increased for some listeners in the high-pass condition. Listeners' ability to localize the virtual stimuli in the vertical plane demonstrated that our stimulus synthesis was of sufficient quality to deliver accurate spectral cue information. We therefore presume that any influence of these spectral cues on apparent lateral angle (see Secs. V and VI) was similar to that occurring in real-world auditory environments. In the low-pass condition, no listeners were able to judge accurately the polar angle of the targets, and all made responses near the horizontal plane in either the front or rear hemisphere for such filtered targets. This is a response pattern typically observed for low-pass stimuli (for example, Carlile *et al.*, 1999; Morimoto and Aokata, 1984).

The change in polar angle responses to ITD- or ILD-biased targets varied among listeners, but in general, the medium-magnitude biases (± 300 μs ITD or ± 10 dB ILD) produced little change from the unbiased condition. For approximately half of the listeners, large-magnitude biases (± 600 μs ITD or ± 20 dB ILD) substantially increased the rate of front/back confusions (as seen, for example, in Fig. 2) and caused mild compression of polar angle responses towards the horizontal plane. Wightman and Kistler (1992) have reported similar effects on vertical-plane localization for ITD-biased virtual stimuli. For the other listeners, even the large biases had little effect on polar responses. Sensitivity of polar response patterns to imposed ITD or ILD bias did not appear to be correlated with the weight given by individual listeners to the biased cue.

C. Discussion

1. Previous lateralization studies

Much of our knowledge about the auditory system's processing of interaural difference comes from so-called lateralization studies, in which stimuli are delivered over headphones without DTF filtering and their apparent positions lie inside the listener's head. In many respects, our derived weights for the ITD and ILD cues are consistent with the results of lateralization experiments employing both pure-tone and noise stimuli. We observed moderate to large weights on both ITD and ILD for wideband targets, although ITD was usually weighted more strongly; large ITD and small ILD weights for low-pass targets; and small ITD and large ILD weights for high-pass targets.

The intracranial position of diotic wideband noise can be displaced from the midline by the application of either ITD or ILD, and can be fully shifted to one side by either cue; similar sensitivities to ITD and ILD exist for low-frequency tones and low-pass noise (Blauert, 1997; Pinheiro and Tobin, 1969). ILD can displace the images of high-frequency tones (Fedderson *et al.*, 1957) and bands of noise (Simon and Aleksandrovsky, 1997), but high-frequency tones above ~1.3 kHz cannot be lateralized on the basis of ongoing ITD (Zwislocki and Feldman, 1956). It is not clear whether the decline in interaural phase sensitivity above this frequency is caused by loss of phase-locking in the auditory nerve or to a lower-frequency cutoff particular to the binaural system. Phase-locking at frequencies up to ~4 kHz has been observed in the squirrel monkey auditory nerve (Rose *et al.*, 1967), but whether this is representative of the human auditory system is unknown.

In principle, the auditory system should be able to exploit envelope fluctuations in such signals to extract ITD information, but thresholds for detecting ITD changes in amplitude-modulated high-frequency tones are typically much larger than those for low-frequency tones (e.g., McFadden and Pasanen, 1976). Also, several studies have shown that the extent of lateralization possible based on high-frequency envelope ITD is in general small and listener-dependent (Blauert, 1982; Henning, 1974; Trahiotis and Bernstein, 1986). This parallels the typically small, but occasionally substantial, high-frequency ITD weights derived in our experiments I and II. (See Sec. IV D to follow.)

The one aspect in which our results do not correspond to those obtained in lateralization studies is the low weight given to ILD in the low-pass condition of our Experiment I (mean weight, 0.24; see Fig. 5). When presented over headphones without DTF filtering, the intracranial position of low-pass noise is sensitive to ILD. Pinheiro and Tobin (1969) found that the image of broadband noise and noise low-pass filtered at 1.2 kHz could be shifted fully to one side by an ILD of 9 dB. Although naturally occurring ILDs are smallest in the low-frequency regime, this 9-dB value does not exceed the physiological range. For all of our listeners, ILD computed over the 0.5–2-kHz band approached or exceeded 10 dB at lateral angles of 50–60 degrees. Similar ILDs were observed in a 0.8–1-kHz band by Wightman and Kistler (1997a, Fig. 6).

The difference between the low-pass lateralization result and our own might be related to the externalization of our stimuli. Our unbiased lateral angle gain measure indicated that listeners were able to judge accurately the lateral angle of these low-pass targets, and, anecdotally, none reported in-the-head-localization for biased or unbiased low-pass targets. Presumably the 2-oct bandwidth and DTF filtering of these targets was sufficient to create externalized auditory images and perhaps to engage a different mode of cue processing, in which ILD plays a much reduced role in the determination of perceived lateral angle. The role of low-frequency ILDs in near-field distance perception is discussed in Sec. III C 3.

The relative potency of ITD and ILD has been explored in many lateralization studies in which a time-intensity trading ratio was measured. The trading ratio measured using a "centering" method describes the amount of time difference favoring one ear that is required to center the image of a signal presented with a level difference favoring the opposite ear. In the "pointer" method, the listener adjusts the ITD of a broadband noise to match the lateral location of the experimental stimulus presented with some combination of ITD and ILD (e.g., Moushegian and Jeffress, 1959). Although prone to intersubject differences and level dependencies, a typical result is that the trading ratio for low-frequency stimuli is lower than that for high-frequency stimuli, which indicates that ITD is more potent relative to ILD at low frequencies. For example, Harris (1960) reported trading ratios of ~25 μs/dB and ~60 μs/dB for low-passed and high-passed clicks, respectively.

Stimuli with images centered using conflicting ITD and ILD cues can be readily discriminated from diotic stimuli (Hafter and Carrier, 1969) perhaps because narrow-band stimuli presented over headphones with conflicting ITD and ILD cues can generate multiple intracranial images (e.g., Whitworth and Jeffress, 1961). The location of the so-called "time" image is determined primarily by ITD (and thus exhibits a very low trading ratio), while the "intensity" image location is controlled by both ILD and ITD (higher trading ratio). Hafter and Jeffress (1968) found that trading ratios for both types of images were higher for high-passed clicks than for 500-Hz tone pips, again indicating the reduced potency of ITD at high frequencies. Gaik (1993) examined the conditions under which narrow bands of noise were most likely to produce multiple or noncompact intracranial images. Natural combinations of ITD and ILD in each frequency band were identified using measured directional impulse responses from a human subject. Results showed that the lateralized images were most likely to be unitary and compact when the imposed ITD and ILD were close to a natural pairing of these cues. It is not known whether similar image splitting occurs for wideband noise signals with conflicting binaural difference cues. Our data (see Sec. III C 2) and those of Wightman and Kistler (1992, 1997b) do not provide strong evidence for such an effect in localization experiments.

2. Localization or lateralization?

Although our results are substantially in agreement with many lateralization studies describing the relative strengths of ITD and ILD cues and the perception of binaural stimuli with conflicting time and level cues, we believe that ours was not a lateralization experiment. First, our stimuli consisted of multi-octave, DTF-filtered bands of noise, whereas a majority of trading ratio, binaural discrimination, and extent-of-lateralization studies have used stimuli of restricted bandwidth. Our stimuli thus provided listeners with the opportunity to integrate binaural difference information across frequency, to use monaural spectral cues, and to experience externalized auditory images.

Second, several pieces of evidence make us confident that our listeners perceived externalized (rather than intracranial) images and were localizing them as they would real free-field targets. It was not practical to combine distance estimation with our localization response method, but in a VAS study using ILD-biased stimuli very similar to ours (Wightman and Kistler, 1997b), listeners reported externalized images even for extreme ILD biases. None of our listeners reported trouble making orienting responses to the biased stimuli, which might have been expected were the images not externalized. S18 (the first author), S04 and S93 (laboratory colleagues) all described the images as well externalized and noticed no obvious differences between biased and unbiased stimuli.

Third, our data provide no indication that listeners were responding to multiple images produced by the biased stimuli. Narrow-band signals presented in a lateralization paradigm with conflicting time and amplitude cues often produce multiple intracranial images (i.e., ‘‘time’’ and ‘‘intensity’’ images). Had this occurred frequently with our VAS stimuli, we would expect to observe bimodality in lateral angle responses and in plots of observed versus imposed bias. This would have been particularly evident in the ILD-biased conditions, for which the location of the ‘‘time’’ image would be highly insensitive to the manipulation. Bimodal response patterns were not observed (for example, Fig. 4). Even if multiple images were generated, we are content that our derived weights describe the contributions of ITD and ILD to the image that dominated the percept and drove the listeners’ orienting responses. Scatter in listeners’ lateral angle responses did not increase markedly as ITD or ILD bias was imposed.

Finally, the strongest evidence that the biased stimuli were localized rather than lateralized comes from the polar angle response data. Imposition of large ITD or ILD biases did produce an increase in front/back confusions for some listeners, but, as in the example shown in Fig. 2, listeners continued to respond accurately to the original elevation of the target. This accuracy would seem unlikely if the orienting responses were derived from a nonexternalized image.

3. Previous localization studies

Our work is not the first to use a localization task in addressing the relative roles of ITD and ILD in spatial hearing. Sandel *et al.* (1955) used a loudspeaker array to produce natural and unnatural combinations of interaural phase and intensity for pure tones, and concluded that ITD was the dominant lateral angle cue for frequencies below 1.5 kHz.

Other researchers have used VAS techniques and noise stimuli to explore the binaural cue weighting. Wightman and Kistler (1992) presented virtual free-field targets in which the interaural phase spectrum was manipulated to correspond to a lateral angle of 90 degrees, rather than to the natural lateral angle of the target. For wideband noise stimuli, listeners’ lateral angle judgments agreed with the manipulated interaural phase cue, but with frequencies below 2.5 kHz removed by high-pass filtering, the influence of the fixed ITD cue was almost eliminated for most, but not all, listeners. This result demonstrated the dominance of low-frequency ITD information over other interaural cues and the lack of influence of high-frequency ITD cues. In order to prevent listeners from learning the individual spectral characteristics of their loudspeakers, Wightman and Kistler scrambled the spectra of their noise targets from trial-to-trial in $\frac{1}{3}$-oct bands. This might have obscured or weakened the salience of the veridical spectral cues competing with the manipulated ITD cue. For some of the listeners in our Experiment I, the wideband ITD weight was lower than the low-pass ITD weight, which might indicate increased influence of high-frequency spectral or ILD cues in our unscrambled, wideband noise targets. Wightman and Kistler did not seek to differentiate between the roles of DTF spectra and ILD as the salient cues to lateral angle for high-pass stimuli. The results of our Experiment IV (Sec. VI), however, suggest that DTF spectra have little influence on lateral angle judgments.

Wightman and Kistler (1997b) also obtained results in agreement with ours in a virtual free-field localization condition almost identical to the wideband-ILD condition of our Experiment I. Broadband (0.2–14 kHz) noise targets were presented with an attenuation of 0-40 dB applied to the left-ear signal. For level imbalances (equivalent to our ILD bias) of 10 or even 20 dB, the listeners’ responses were generally unperturbed, and corresponded to the position indicated by the unmanipulated ITD and spectral cues. This result was in agreement with our finding that ILD is a weak cue for lateral angle in wideband noise stimuli.

Our results are in agreement with those of Wightman and Kistler (1992, 1997b) that ITD is the dominant lateral angle cue for stimuli containing low-frequency components. We note, however, that this dominance has been detected only in experiments conducted in anechoic or virtual free-field environments, and might not apply in reverberant environments. Hartmann and Constan (1998) found that the binaural coherence of low-frequency signals presented in rooms is lower than that required to support lateralization on the basis of ITD. Virtual auditory space (VAS) studies incorporating synthesized reverberation would be useful in determining the weighting of cues under nonanechoic conditions.

Low-frequency ILDs for lateral sources in the near-field (<1 m) are much larger than those observed for distant sources, but low-frequency ITDs are not strongly distance-dependent (Brungart and Rabinowitz, 1999). Lateral angle localization judgments remain accurate in the near-field even as low-frequency ILDs diverge from their far-field values (Brungart *et al.*, 1999). These results suggest that, in accord with our finding from the low-pass passband condition of Experiment I, apparent lateral angle for low-pass sources is determined on the basis of ITD (distance-independent) and is independent of ILD (distance-dependent). Low-frequency ILDs have been shown to be important cues for near-field apparent distance (Brungart, 1999). We hypothesize that low-frequency ILDs have little effect on apparent source *direction* because they are *reserved* as cues for near-field *distance* perception.

IV. EXPERIMENT II: WEIGHTING OF ONSET AND ONGOING ENVELOPE-BASED ITD CUES IN HIGH-PASS NOISE

A. Motivation

Experiment I revealed marked individual differences in the perceptual weighting of ITD in high-pass noise stimuli. Because the peripheral auditory system cannot transduce the fine structure of these stimulus waveforms, listeners who placed substantial weight on the high-pass ITD cue must have derived interaural timing information from the envelopes of the signals. Information might be extracted by interaural processing of stimulus onsets or offsets. Moreover, information might be extracted from ongoing fluctuations in the envelopes of the noise waveform, although Middlebrooks and Green (1990) have hypothesized that modulation depth in the outputs of high-frequency auditory filter channels is too small for robust extraction of ongoing ITD cues from noise stimuli.

In experiment II, we explored the relative influence of ITD cues from the onset and ongoing portions of the envelope in high-pass noise stimuli. We attempted to *weaken* the transient envelope ITD cues by *lengthening* the duration of the stimulus onset and offset ramps. In a separate condition, we attempted to *strengthen* the ongoing cues by *amplitude-modulating* the target noise bursts.

B. Stimuli and locations

In Experiment II-A, the high-pass ITD and ILD bias manipulations of Experiment I were repeated with identical sets of target location and bias combinations, but the temporal characteristics of the target noise bursts were altered. The onset and offset ramps of the noise bursts were lengthened from 1 to 20 ms in order to weaken the onset/offset ITD cue. To preserve the 98-ms plateau duration present in the 1-ms-ramped stimuli, the length of the noise bursts was increased to 138 ms. All ten listeners from Experiment I participated in Experiment II-A.

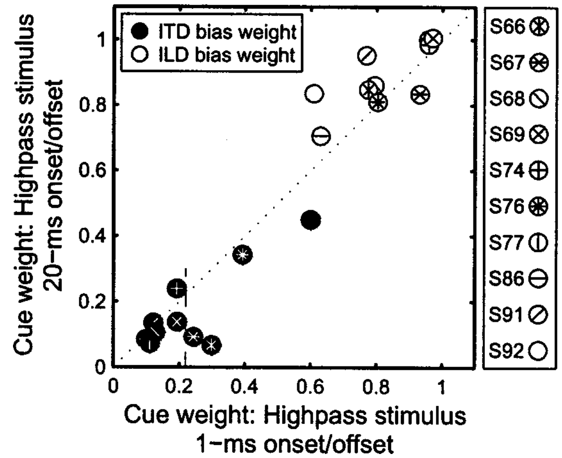

FIG. 6. Effect of onset ramp duration on high-pass ITD and ILD bias weights. The scatter-plot shows high-pass ITD (filled symbols) and ILD (open symbols) bias weights for each of the ten listeners in Experiment II-A. Horizontal axis: cue bias weights for 1-ms onset stimuli; vertical axis: cue bias weights for 20-ms onset stimuli.

In experiment II-B, alterations of the onset/offset and ongoing portions of the envelope were combined factorially. Onset/offset ramps were either short (1 ms) or long (50 ms, 198 ms total duration), and the envelope of the ongoing portion was either *flat* or *modulated*. In the modulated case, the ongoing portion of the signal consisted of 4-ms segments of silence alternating with 6-ms noise bursts with 1-ms onset and offset ramps.[1] The long onsets were intended to reduce the salience of the onset ITD cue, and the amplitude modulation was intended to enhance the salience of the ongoing envelope ITD cue. Data were collected for the high-pass passband condition with either ITD or ILD biases and for all four combinations of the ramp and modulation parameters. Each of the eight stimulus sets (1 passband×2 cue bias types×2 onsets×2 modulations) was presented only once. All other details of the stimulus sets and their presentation were identical to those of Experiment I, including the block size of 111 trials. In the 1-ms-onset/flat-envelope condition for listeners S91 and S92, the data from the equivalent condition of Experiment I were used.

C. Results

1. Experiment II-A

The high-pass ITD and ILD weights for the 20-ms onset stimuli are compared to those for the 1-ms onset targets in Fig. 6. The four listeners whose 1-ms ITD weights were ⩾0.24 in Experiment I all exhibited reduction of the ITD weight in the 20-ms onset condition. This suggests that they had been relying on the onset ITD cue. The ITD weights for two of these four listeners remained fairly high, however, suggesting either that onset ITD information was still available with the 20-ms onset or that these listeners could exploit

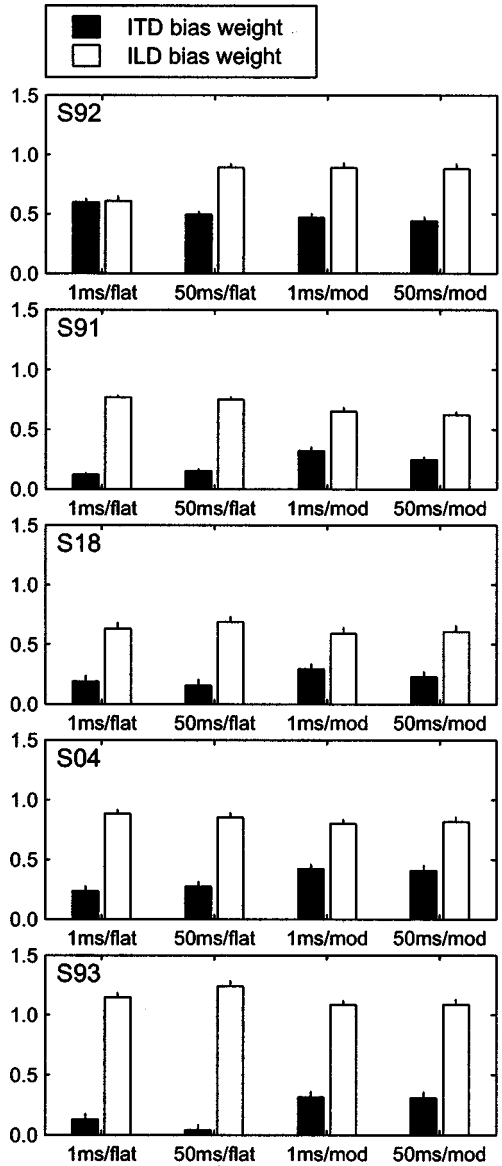

FIG. 7. Effect of onset ramp duration and envelope modulation on high-pass ITD and ILD bias weights. ITD bias weight (black bars) and ILD bias weight (open bars) are plotted for five listeners for each of the four combinations of onset duration (1 or 50 ms) and noise envelope modulation (flat or 100 Hz, 100% modulation depth). Error bars show the standard error of the regression coefficient (i.e., of the cue bias weight).

intrinsic envelope fluctuations to obtain ongoing ITD information. Further evidence for the reduced salience of the onset ITD in the 20-ms condition is provided by the increase in the ILD weights obtained for the longer-onset stimuli. In ILD bias conditions, the ITD cue was consistent with the unbiased target location. Thus weakening of the ITD cue should have reduced the influence of cues competing with the biased ILD.

2. Experiment II-B

ITD and ILD weights for the stimuli of Experiment II-B are shown for five listeners in Fig. 7. The unbiased lateral angle gain was consistent across envelope condition for all listeners, which indicates that the envelope manipulations did not affect the listeners' localization accuracy. In the 1-ms-onset/flat-envelope condition, the weights for listeners S18, S04, and S93 (0.19, 0.24, and 0.13, respectively) were consistent with the distribution of weights observed in the high-pass ITD bias condition of Experiment I.

We consider first the effect of the onset-time manipulation on ITD weight. In the flat-envelope conditions, the effect of lengthening the signal onset and offset from 1 to 50 ms was small (mean ITD weight reduction, 0.03) except for listener S92, whose larger ITD weight decreased from 0.60 to 0.50. Thus, as found in experiment II-A, weakening the onset ITD cue had little impact on the listeners whose ITD bias weights were already small in the 1-ms-onset condition (S91, S18, S04, and S93), likely because of a simple floor effect. For S91 and S04, the ITD bias weights actually increased slightly (by only 0.03 and 0.04, respectively). In the modulated-envelope conditions, lengthening the onset had little effect on the ITD weight for any listener (mean ITD weight reduction, 0.04).

In contrast, the envelope modulation manipulation had a pronounced effect in both the 1- and 50-ms onset conditions. Four of the five listeners exhibited an increase in the weighting of the ITD cue in both onset-length conditions when envelope modulation was added (for these four, mean ITD weight increase over both onset lengths, 0.16). For these four listeners, adding envelope modulation also produced modest decreases in the ILD weight (mean ILD weight reduction, 0.09). A similar inverse relation between ITD and ILD weights was observed in Experiment II-A. Listener S92, for whom the high-pass ITD weight in Experiment I was already high (0.60), maintained a consistently high weight on the ITD cue but did not exhibit an increase in the weight in the modulated-envelope conditions.

Together, the results of Experiments II-A and II-B suggest that both onset and ongoing envelope ITD cues play a role in the sensitivity of some listeners to high-frequency ITD. It appears, however, that the highest high-frequency ITD bias weights are obtained when the listener is able to process ongoing envelope ITD cues, and that under such circumstances, the onset cue is of reduced importance.

D. Discussion: Role of high-frequency envelope ITD cues

As noted previously, complex, high-frequency stimuli such as amplitude-modulated tones, tone complexes, and bands of noise provide the auditory system with ongoing envelope ITD information that is absent from high-frequency pure-tone stimuli. ITD discrimination thresholds measured for such sounds are typically found to be two to ten times larger than those for low-frequency tones—as large as a few hundred μs (e.g., Bernstein and Trahiotis, 1994; Blauert, 1982; Henning, 1974; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1981). Although easily detectable, ITDs of physiologically plausible magnitude have been found to be remarkably ineffective in displacing the intracranial images of high-frequency, complex stimuli away from the midline in lateralization experiments, and both detection thresholds and lateral displacement sensitivity display marked individual differences (Trahiotis and Bernstein, 1986).

Our results from Experiment I paralleled these findings (in a localization rather than lateralization task) both with respect to the generally low weights accorded to high-frequency ITD and to the variation of those weights among listeners.

The results of Experiments II-A and II-B show that (a) some listeners with substantial weights on high-frequency ITD derive this information primarily from the onset ITD cue, (b) others are able to extract ongoing ITD from intrinsic noise envelope fluctuations, and (c) if robust envelope modulation is present, weights on high-frequency ITD increase and the strength of the onset cue has little effect. These results are consistent with the findings that listeners vary in their ability to extract envelope-based, ongoing ITD information, and that onset cues are salient only when the normally potent ongoing cues are ambiguous. The psychophysical literature contains a diversity of reports on the relative strength and discriminability of onset and ongoing ITD cues (for example, Buell *et al.*, 1991; Hafter and Dye, 1983; Tobias and Schubert, 1959). In a synthesis of these and other results, Freyman and colleagues (1997) concluded that results differ because the contribution of the onset cue is strongly dependent upon the spectral, temporal, and binaural characteristics of the ongoing stimulus. In particular, ''lateralization of a spectrally dense signal with an unambiguous ongoing delay is not subject to dominance by onsets even if the onset cue itself is strong'' (Freyman *et al.*, 1997). In our Experiment II, the high-pass noise signals satisfied the condition of spectral density, and ambiguity in ongoing envelope ITDs was reduced by applying envelope modulation.

The results of a simulation of high-frequency, envelope-based ITD discrimination (see Appendix) suggest that intrinsic envelope fluctuations in unmodulated high-frequency noise bands are not sufficiently large for useful discrimination of ongoing, envelope-based ITDs. This result supports the hypothesis of Middlebrooks and Green (1990). Modulation depth was found to be independent of differences in listeners' DTFs, and therefore individual differences in high-pass ITD bias weights must be related to individual differences in envelope extraction processes or in information processing strategies. Constan and Hartmann (2001) have shown that in reverberant environments such as those in typical rooms, the coherence of the two ear signals is often insufficient to permit lateralization on the basis of high-frequency envelope ITD. This may be another reason that most listeners discount these cues.

For noise stimuli, intrinsic envelope fluctuations in different auditory filter channels are not correlated. That is, the channel signals are not *comodulated*. Comodulation is known to be an important factor in across-frequency integration of information (e.g., Hall and Grose, 1990), and it is possible that, for noise stimuli, the lack of comodulation among envelopes in different auditory filter bands reduces the salience of the high-frequency envelope cue for most listeners. The increased ITD bias weight we observed when amplitude modulation was imposed on the noise targets in Experiment II-B might have been a function of increased interchannel comodulation as well as increased modulation depth. Saberi (1995) has shown that ITD discrimination thresholds for stimuli composed of two spectrally separated, high-frequency, narrow bands of noise are lower when the temporal envelopes of the bands are identical than when they are different. Similarly (but in a low-frequency regime), Trahiotis and Stern (1994) have found that complexes of spectrally separated SAM tones with consistent carrier ITDs produce compact unitary binaural images only when identical modulators are applied to each carrier.

Eberle *et al.* (2000) investigated the salience of envelope-based ITD cues in a free-field localization task. Listeners reported the apparent location of a high-frequency octave band of noise (7–14 kHz), with and without an applied 20- 80- or 320-Hz, 100%, sinusoidal amplitude modulation. This signal was similar to the modulated-envelope targets used in Experiment II-B of the present study. The authors found that the introduction of amplitude modulation produced no reduction in the mean magnitude of errors in listeners' lateral angle judgements. They inferred that the amplitude modulation did not facilitate the extraction of envelope-based, high-frequency ITD cues. The results of our Experiment II-B suggest that ITD salience is enhanced by amplitude modulation, and thus are not in agreement with this conclusion.

The discrepancy between our conclusions and those of Eberle *et al.* (2000) likely resulted from the fact that our VAS technique permitted the elimination of a confound between ITD and ILD cues inevitably present in the free-field situation. In our Experiments I and II, listeners were able to judge accurately the lateral angle of the unmanipulated high-pass targets despite their generally low sensitivity to the ITD cue. The veridical judgements were, therefore, likely to have been based on the ILD or on spectral cues (but see Sec. VI D to follow). Using free-field target presentation, it is not possible to present stimuli in which the various localization cues are in conflict. If, similarly to our subjects, the listeners in the Eberle *et al.* (2000) study were able to judge lateral angle accurately on the basis of the sufficient and veridical ILD cues, then adding ITD information consistent with the ILD via amplitude modulation would not have changed the apparent location of the source. The consistently low lateral localization error across conditions might reflect the resolution of the motor response method rather than evidence of complete discounting of the high-frequency ITD cue.

V. EXPERIMENT III: WEIGHTING OF THE INTERAURAL LEVEL SPECTRUM CUE

A. Motivation

The frequency-independent ILD bias manipulation used in Experiments I and II resulted in stimuli with unnatural *interaural level spectra* (ILS; i.e., patterns of ILD across frequency). As a wideband sound source is moved away from the median plane, ILD grows more rapidly at high frequencies than at low ones, and thus the frequency-independent ILD created by applying ILD bias to a midline source in Experiments I and II would never be observed under natural conditions. In principle, the ILS could serve as a robust cue to source elevation and front-back location for wideband sources (Duda, 1997), because it is the difference between the spectra at the left and right ears, and therefore largely independent of irregularities in the source spectrum. There is, however, little psychophysical evidence to support the salience of ILS as a vertical-plane localization cue (see Discussion, Sec. V E).

In Experiment III, we attempted to bias the perceived location of virtual free-field targets without introducing unnatural patterns of ILD across the frequency spectrum.

B. Stimuli and locations

The targets were 100-ms noise bursts with 1-ms raised-cosine onsets and offsets. Stimuli with biased interaural level spectra were generated as follows. Starting with an original target location, **A**, we identified a second location, **B**, with the same polar angle, but displaced in lateral angle from **A** by ±30 or ±60 degrees.

For the ear on the same side as the target (the *near* ear), we retained the DTF of location **A**, $\mathrm{DTF}_{\mathrm{near}}(\mathbf{A},f)$. For the opposite ear (the *far* ear), we generated a new transfer function, $X(f)$, such that the resulting ILS matched that of location **B**. That is,

$$\left|\frac{\mathrm{DTF}_{\mathrm{near}}(\mathbf{A},f)}{X(f)}\right| = \mathrm{ILS}(\mathbf{B},f) = \left|\frac{\mathrm{DTF}_{\mathrm{near}}(\mathbf{B},f)}{\mathrm{DTF}_{\mathrm{far}}(\mathbf{B},f)}\right|.$$

It follows that

$$|X(f)| = \left|\mathrm{DTF}_{\mathrm{near}}(\mathbf{A},f) \times \frac{\mathrm{DTF}_{\mathrm{far}}(\mathbf{B},f)}{\mathrm{DTF}_{\mathrm{near}}(\mathbf{B},f)}\right|.$$

In practice, we computed $X(f) = \mathrm{DTF}_{\mathrm{far}}(\mathbf{A},f) \times H(f)$, where $H(f)$ was the *zero-phase* filter with transfer function

$$H(f) = \left|\frac{\mathrm{DTF}_{\mathrm{near}}(\mathbf{A},f)}{\mathrm{DTF}_{\mathrm{far}}(\mathbf{A},f)}\right| \Bigg/ \left|\frac{\mathrm{DTF}_{\mathrm{near}}(\mathbf{B},f)}{\mathrm{DTF}_{\mathrm{far}}(\mathbf{B},f)}\right| = \frac{\mathrm{ILS}(\mathbf{A},f)}{\mathrm{ILS}(\mathbf{B},f)}.$$

This operation preserved the natural near-ear DTF and the interaural phase spectrum (and hence the ITD) measured at location **A**, but altered the ILS to correspond to that of location **B** (see Table I). Positive (rightward) ILS bias was applied only to targets originally on or to the right of the median plane. Similarly, negative ILS bias was applied only to midline or left-hemisphere targets. Thus in the above expression for $H(f)$, the ear designated "near" remained the same for location **A** and location **B**, as did the ear designated "far."

The wideband, low-pass, and high-pass passband conditions used in Experiment I were employed. The groups of target locations for Experiment III were similar to those for Experiments I and II, but because ILS bias was applied only to shift apparent position laterally *away* from the median plane, the total number of biased targets was reduced from 144 to 80. In total there were three stimulus sets (3 passbands×1 cue bias type) of 158 targets each. The targets in each set were presented twice in a shuffled order over the course of four blocks of 79 trials each. Blocks from different passband conditions were intermixed within an experimental session but trials from different sets were not intermixed within blocks.

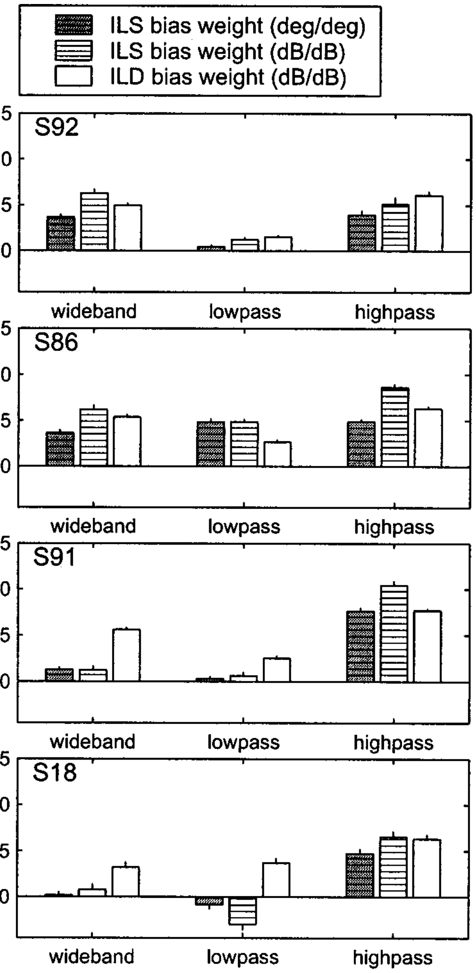

FIG. 8. Interaural level spectrum and interaural level difference bias weights for Experiment III. Each panel shows data for one listener. ILS-degree/degree bias weight (shaded/hatched bars), ILS-dB/dB bias weight (unshaded/hatched bars) and ILD bias weight (open bars) are plotted for the wideband, low-pass and high-pass passband conditions. Error bars show the standard error of the regression coefficient (i.e., of the cue bias weight).

C. Analysis

Because the change in the ILS caused by the manipulation was not a scalar quantity, as were the imposed ITD or ILD biases of Experiments I and II, we first computed a weight for the ILS cue by simply computing the slope of a linear fit to the observed lateral angle response bias versus imposed ILS bias data, both measured in degrees. This gave a weight in units of degree/degree. We also analyzed the responses as we did in Experiment I, treating *gross ILD*, the overall interaural difference in energy integrated across the stimulus passband, as the effective cue. In this case, the imposed bias was computed from the difference between the interaural energy ratios present in the original and modified DTFs, and the observed bias was computed as in Experiment I (Sec. II E). A linear regression between the observed and imposed biases yielded an ILS bias weight in units of dB/dB.

D. Results

The cue weights derived from the responses to ILS-biased stimuli are shown in Fig. 8, along with the corresponding ILD weights from Experiment I (or equivalent conditions for S18). In general the relation between ILS-degree/degree weights (shaded/hatched bars) and passband condition was similar to that found for the ILD weights (open bars) of Experiment I. That is, low weight was given to the ILS cue in the low-pass condition (three of four listeners), and a higher weight was given in the high-pass condition. Curiously, small negative ILS-degree/degree and ILS-dB/dB weights were observed for listener S18 in the low-pass condition.

Several features of the data suggest that the detailed shape of the interaural level spectrum itself is not a particularly salient cue for lateral angle, but rather that the effective interaural level cue is the overall energy difference between the ears. For three of the four listeners, the ILS-degree/degree weight was lower than the ILD weight from experiment I in both the wideband and high-pass conditions. Also, in the majority of cases in the wideband and high-pass conditions, the ILS-dB/dB (gross ILD) weight was higher than the ILS-degree/degree weight computed from the same responses.

In some cases, a particular gross ILD produced more response bias when it was produced by the ILS manipulation than when it was applied as a flat attenuation. Evidence for this can be found in the high-pass weights for S86 and S91. For these listeners, the high-pass ILS-dB/dB weight (Fig. 8, hatched bars) was substantially higher than the high-pass ILD weight from Experiment I (open bars).

E. Discussion: Interaural level spectrum and the nature of ILD processing

The shape of the interaural level spectrum has been proposed as a cue for vertical-plane localization for locations both on (Searle *et al.*, 1975) and off (Duda, 1997) the median plane. The results of several psychophysical studies contradict this proposal. First, vertical-plane localization of wideband sources can be disrupted by certain source-spectrum irregularities which preserve the ILS (e.g., Macpherson, 1996, 1998; Rakerd *et al.*, 1999; Wightman and Kistler, 1997a). Second, there is evidence that the shape of the spectrum at the ear contralateral to the target can be altered substantially without consequent degradation in vertical-plane localization accuracy (Humanski and Butler, 1988; Morimoto, 2001; Wightman and Kistler, 1999), and that the influence of the contralateral-ear DTF declines as the source location is moved away from the median plane (Morimoto, 2001). These results suggest that the details of the ILS are unimportant for human listeners' vertical-plane localization, and our results from Experiment III lead to a similar conclusion about the lack of efficacy of the ILS as a cue to lateral angle.

The effective cue for lateral angle in high-pass stimuli seems to be better modeled as the overall difference in proximal stimulus energy between the ears, or, perhaps more realistically, as an integration of the ILDs observed in discrete frequency channels. This would parallel the auditory system's use of low-frequency ITD as a cue to lateral angle and its apparent insensitivity to fine details in the interaural phase spectrum (Kulkarni *et al.*, 1999). Both the ITD and ILD findings are consistent with the suggestion that " . . . the system does not evaluate every detail of the complicated interaural dissimilarities, but rather derives what information is needed from definite, easily recognizable attributes" (Blauert, 1997, p. 138).

The results of Experiment III do, however, reveal some effect of the distribution of ILD across frequency. As discussed earlier, we observed that a net ILD bias generated by the ILS manipulation was more effective in biasing lateral angle judgments than was an equivalent flat interaural attenuation (Fig. 8). That is, the dB/dB weights for ILD obtained in Experiment III were higher than those obtained in Experiment I. A possible explanation for this is that listeners placed greater weight on the gross ILD when the ILS was more natural, or at least when it did not contain unnatural low-frequency ILDs. This would be consistent with the proposal of Wightman and Kistler (1997a) that the naturalness of observed localization cues plays a mediating role in their salience.

VI. EXPERIMENT IV: WEIGHTING OF THE NEAR-EAR SPECTRAL CUE

A. Motivation

Some have suggested that monaural spectral cues are important in determining perceived lateral angle as well as being the primary cues to sound source elevation and front/back location (for example Butler and Flannery, 1980). In Experiments I, II, and III of the present study, we could not directly assess the influence of spectral cues on perceived lateral angle because we did not manipulate these cues independently of both of the binaural difference cues (ITD and ILD) simultaneously. In all cases, the spectral cue in the ear on the side of the original target location corresponded with the unmanipulated binaural difference cue. The results of Experiments I and II suggest, however, that the spectral cue contribution is minimal, as we discuss below in Sec. VI D.

In Experiment IV, we biased the lateral angle corresponding to the spectral cue in one ear independently of both the ITD and ILD cues. The near-ear DTF (as defined in Sec. V B) was replaced by another corresponding to a new location displaced in lateral angle from the original target location. The opposite-ear impulse response was altered in order to maintain the original interaural level and phase spectra. We manipulated the near-ear DTF because it is thought to be the most potent cue for vertical plane localization and because, as discussed earlier, it appears that the far-ear spectrum can be manipulated severely without affecting vertical plane localization (Humanski and Butler, 1988; Morimoto, 2001; Wightman and Kistler, 1999).

B. Stimuli and locations

As in the previous experiments, the targets were 100-ms noise bursts with 1-ms raised-cosine onsets and offsets. Stimuli with biased near-ear DTFs were generated as follows. Starting with an original target location, **A**, we identified a second location, **B**, with the same polar angle, but displaced in lateral angle from **A** by ±30 or ±60 degrees.

The interaural level- and phase-difference spectra for location **A** were computed as

$$\text{ILS}(\mathbf{A},f)=\left|\frac{\text{DTF}_{\text{near}}(\mathbf{A},f)}{\text{DTF}_{\text{far}}(\mathbf{A},f)}\right|$$

and

$$\text{IPS}(\mathbf{A},f)=\phi\{\text{DTF}_{\text{near}}(\mathbf{A},f)\}-\phi\{\text{DTF}_{\text{far}}(\mathbf{A},f)\}.$$

For the ear on the target side, $\text{DTF}_{\text{near}}(\mathbf{B},f)$ was substituted for $\text{DTF}_{\text{near}}(\mathbf{A},f)$; this constituted the spectral cue bias. In the far-ear channel, a new transfer function, $X(f)$, was synthesized with magnitude

$$|X(f)|=\left|\frac{\text{DTF}_{\text{near}}(\mathbf{B},f)}{\text{ILS}(\mathbf{A},f)}\right|$$

and phase

$$\phi\{X(f)\}=\phi\{\text{DTF}_{\text{near}}(\mathbf{A},f)\}-IPS(\mathbf{A},f).$$

Thus the stimulus had the ITD and ILS cues of location **A**, but the near-ear DTF of location **B**. This differed from experiment III, in which the near-ear DTF corresponded to **A** and the ILS to **B** (see Table I). Prior to computing the far-ear DTF magnitude spectrum, $\text{ILS}(\mathbf{A},f)$ was smoothed by frequency-domain convolution with a 1-kHz-wide rectangular filter. This reduced the creation of sharp resonances in the synthesized DTF caused by notches in $\text{DTF}_{\text{near}}(\mathbf{A},f)$.

Biased targets were located at lateral angles of ±30 or ±60 degrees and at elevations of 0, ±20, and ±40 degrees in the front and rear hemispheres. At each location there were three possible biases that did not place locations **A** and **B** on opposite sides of the median plane. For example, a target at a lateral angle of −30 degrees could be biased by −60, −30, or +30 degrees without crossing the median plane.

The wideband and high-pass passband conditions used in Experiment I were employed. In total there were two stimulus sets (2 passbands×1 cue bias type) of 198 targets each. Of these targets, 120 were filtered and biased, 40 were filtered but unbiased, and the remaining 38 were wideband, unmanipulated targets. The targets in each set were presented once in a shuffled order over the course of two blocks of 99 trials each. Three listeners (S04, S18, S93) who had participated in Experiment II-B participated in Experiment IV.

C. Results

We computed a bias weight for the ispilateral DTF cue by simply computing the slope of a linear fit to the observed lateral angle response bias data and the imposed DTF bias data, both measured in degrees. This gave a weight in units of degree/degree, as in our first analysis of the ILS bias data in Experiment III (Sec. V C). Although our manipulation preserved the ILS of the original target location, it did not necessarily preserve the overall ILD because changing the near-ear spectrum changed the spectral distribution of energy. Because we did not wish our results to be confounded by an unintended ILD bias, trials for which the overall stimulus ILD differed from that of the original target location by more than 5 dB were excluded from the analysis. The proportion of excluded trials varied between 4% and 10% among listeners.

The obtained weights are shown in Table II. For all listeners and in both passband conditions, the magnitudes of all weights were <0.1, and in a majority of cases were not significantly different from 0. These results suggest that monaural spectral cues play a negligible role in determining apparent lateral angle even in the high-pass condition, in which the strong conflicting low-frequency ITD cue was removed.

TABLE II. Measured DTF-bias weights for the wideband and high-pass targets of Experiment IV.

Listener	Wideband	High-pass
S04	−0.01±0.03	0.09±0.04
S18	−0.05±0.05	0.00±0.06
S93	0.08±0.04	0.05±0.05

D. Discussion: Influence of spectral cues on lateral localization

In Experiments I and II of the present study, we did not manipulate monaural spectral cues independently of ITD and ILD to assess their influence on perceived lateral angle, but those results also suggest that the spectral cue contribution is minimal. In each stimulus condition of these experiments, the spectral cues at both ears corresponded to the original target location prior to the ITD or ILD manipulation and were in agreement with the unmanipulated interaural cue. Thus, any reliance on the spectral cues should have resulted in a reduction of the weight placed on the manipulated cue, although, considering each condition in isolation, the effects of spectral cues and the unmanipulated cue cannot be disambiguated. Our finding, however, that weights approaching unity were found for one or other of the interaural cues in each passband condition is evidence that spectral cues do not contribute substantially to lateral angle localization even in a binaural listening situation. A high weight on a biased interaural cue indicates that it was possible to shift the apparent position of the target to a location inconsistent with the natural spectral cues. In Experiment IV, we directly tested the influence of monaural spectral cues by biasing the source-side DTF independently of the natural ITD and ILS cues. We found that the DTF bias had little or no influence on the perceived lateral angle of the target.

These results are in agreement with and extend the work of Slattery and Middlebrooks (1994) and of Wightman and Kistler (1997b), who found that normal-hearing listeners rendered monaural by plugging one ear were incapable of accurate lateral angle judgements. In such an acute monaural situation, listeners reported all localization targets originating from a position opposite the unoccluded ear. Slattery and Middlebrooks did find that approximately half of their chronically monaural listeners were able to make use of spectral information to determine source azimuth, but they considered this to be the result of a long-term adaptation to the listeners' lack of access to binaural cues, rather than being indicative of the functioning of the normal auditory system.

VII. SUMMARY AND CONCLUSIONS

Our results suggest that, in broad outline, the duplex theory does serve as a useful description of (if not a principled explanation for) the relative potency of ITD and ILD cues in low- and high-frequency regimes. In Experiment I, we found that listeners weighted the ITD cue strongly as a cue for lateral angle for low-pass stimuli and (with some exceptions) weighted ITD weakly for high-pass stimuli. The opposite pattern was observed for listeners' ILD weights. Even when substantial biases were introduced, ILDs were generally ignored at low frequencies but were given high weights for high-pass stimuli. For wideband targets, both cues were given substantial weight, but ITD dominated for most listeners. In Experiment II, we found that both onset and ongoing envelope ITD cues contributed to the relatively minor role of high-frequency ITD information, but that the greatest high-frequency ITD weights were observed for listeners who were able to make use of ongoing cues. In Experiment III, we examined the role of the detailed shape of the interaural level spectrum as a lateral angle cue. The results suggested that the precise shape of the ILS was not as effective a cue as the overall energy difference between the ears. In Experiment IV, we found that monaural spectral cues had little or no influence on perceived lateral angle.

ACKNOWLEDGMENTS

The authors are grateful to Leslie Bernstein, Fred Wightman, and an anonymous reviewer for their constructive comments on an earlier version of this paper. William Hartmann, Zachary Constan, Douglas Brungart, Brian Mickey, and Christopher Stecker also provided helpful suggestions. Zekiye Onsan provided invaluable technical assistance. This work was funded by NIH Grant Nos. R01DC00420 and T32DC00011.

APPENDIX: CAN INTRINSIC ENVELOPE FLUCTUATIONS IN WIDEBAND NOISE SUPPORT ITD DISCRIMINATION AT HIGH FREQUENCIES?

1. Motivation

For most of our listeners in experiment I, ITD had very little influence on judgments of lateral angle for high-pass noise stimuli. One possible explanation for this result is that the auditory system is unable to extract robust ITD information from such signals, as suggested by Middlebrooks and Green (1990). Alternatively, the system might in principle be able to obtain useful ITD information, but listeners discount that information for some other reason. Psychophysical studies of ITD sensitivity in high-frequency noise have tended to use narrowband stimuli (e.g., Bernstein and Trahiotis, 1994), which might encourage listeners to adopt an off-frequency listening strategy not possible with a broadband stimulus. Because psychophysical data are not available for the multi-octave high-pass stimuli we used in Experiments I and II, we conducted a series of simulations in order to determine whether high-frequency, envelope-based ITD is indeed a viable lateral angle cue for wideband noise stimuli.

To play a role in high-frequency ITD sensitivity, modulations in the right- and left-ear signal envelopes must exhibit two properties. *First*, the envelope fluctuations must be highly *coherent* (i.e., similar), or else ITD cannot be defined. For wideband noise sources off the midline, the strong asymmetry between source-side and far-side DTFs leads to decorrelation of the wideband right- and left-ear envelopes. This occurs because the envelopes of the narrowband components that compose the wideband noise are uncorrelated and the overall envelope depends on the relative levels (and phases) of these components. Within individual auditory filter bands, however, this decorrelation might not occur if the effect of interaural DTF asymmetry is primarily the imposition of a level difference, which would not alter the shapes of the envelopes.

Second, the envelope modulations must be of sufficient *depth*. In a study of interaural envelope delays, Middlebrooks and Green (1990) hypothesized that, because of the low-pass limits of high-frequency envelope following, the depth of intrinsic noise envelope modulations in high-frequency auditory filter bands might be too small to allow effective extraction of envelope-based ITD information from noise signals. Reduction of modulation depth increases ITD just-noticeable differences (jnd's; i.e., discrimination thresholds) in sinusoidally amplitude-modulated (SAM) tones and beating two-tone complexes (Henning, 1974; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1981).

A metric which simultaneously captures the effects of changes in modulation depth and interaural envelope coherence is the *normalized correlation* of the left- and right-ear signal envelopes. This is computed in the same manner as the Pearson product-moment correlation (or *normalized covariance*), but the mean (or d.c.) components of the envelopes are retained. Using this metric, Bernstein and Trahiotis (1996) have successfully accounted for the dependence of ITD jnd's on both SAM-tone modulation depth and two-tone modulation depth as measured by Nuetzel and Hafter (1981) and Pasanen (1976), respectively.

In our simulations, we passed wideband Gaussian noise signals through an auditory filter-bank model. We then estimated ITD discrimination thresholds (Δ_{ITD}) based on the outputs of individual filter channels by determining the minimum increment in ITD required to reduce the normalized correlation by a criterion amount. We derived estimates of minimum audible angle (MAAs) from these Δ_{ITD} estimates. We also examined the sensitivity of the jnd estimates to DTF filtering of the noise signals and to the low-pass cutoff frequency of the envelope-following process.

2. Methods

We passed 16 100-ms exemplars of wideband Gaussian noise through a gammatone auditory filter-bank model (Slaney, 1994) with binaural channels centered at 4, 6, 8, 10, 12, and 14 kHz. Envelopes for each ear in each frequency band were extracted by half-wave rectification followed by low-pass filtering at either 250 or 500 Hz (fourth-order Butterworth filter). All processing was done at a sampling rate of 50 kHz.

To estimate an ITD discrimination threshold for each filter channel and low-pass filter cutoff, we first computed the normalized correlation, ρ, of the right- and left-channel envelopes as we imposed increasing interaural delays in 20-μs (1-sample) steps. The resulting cross-correlation function had a peak value of 1 at 0-lag, and declined monotonically with increasing lag. The imposed interaural delay required to reduce the normalized correlation to 0.99 was then taken as the estimate of the ITD jnd. This criterion matched the mean threshold decrement in normalized correlation of ~0.01 found by Bernstein and Trahiotis (1996, Fig. 4) for their four most sensitive listeners. For each filter channel and low-pass filter cutoff, the final jnd estimate was taken as the mean value over the 16 samples of Gaussian noise.

As a descriptive statistic, we also estimated the envelope modulation depth in each filter channel for each low-pass filter cutoff. Envelope modulation depth is a measure of the magnitude of a signal's amplitude fluctuations relative to the mean level of the signal. For a sinusoidally amplitude-modulated tone, the modulation depth, m_{SAM}, is defined as the ratio of the modulation amplitude, a (half the peak-to-trough range) and the mean level:

$$m_{\mathrm{SAM}}=\frac{a}{\mathrm{mean}_{\mathrm{ENV}}}=\frac{\frac{1}{2}(\max_{\mathrm{ENV}}-\min_{\mathrm{ENV}})}{\frac{1}{2}(\max_{\mathrm{ENV}}+\min_{\mathrm{ENV}})}.$$

For sinusoidal modulation, $a=\sqrt{2}\sigma_{\mathrm{SAM}}$, where σ is the standard (or root-mean-square) deviation of the envelope about the mean level. Using this relation we derived the more general expression

$$m=\frac{\sqrt{2}\sigma_{\mathrm{ENV}}}{\mu_{\mathrm{ENV}}}$$

(in which σ_{ENV} and μ_{ENV} are respectively the estimated standard deviation and mean of the envelope) as an estimator for the modulation depth of a nonsinusoidal envelope.

To investigate how DTF filtering might affect ITD discrimination, for each of the participants in Experiments I–III, we also filtered samples of Gaussian noise with the measured right- and left-ear DTFs corresponding to azimuths of 0, 30, 60, and 90 degrees in the horizontal plane. For each listener, azimuth, filterbank channel and low-pass filter cutoff, we repeated the ITD jnd and modulation-depth estimation procedures using these DTF-filtered signals. Before computing Δ_{ITD}, we removed the DTF-related ITD by time-aligning the right- and left-ear envelopes on a channel-by-channel basis to maximize the normalized correlation. The peak value of ρ was recorded and the jnd was then computed relative to the adjusted ITD. Psychophysical data are not available for discrimination of changes in envelope correlation for pedestal correlations other than 1, but in the majority of cases the peak value of normalized correlation observed was >0.95, and we therefore retained the 0.01 reduction in ρ as a reasonable threshold criterion.

3. Results and discussion

We first consider the results obtained without DTF filtering. Estimates of Δ_{ITD} with the 500-Hz low-pass envelope filtering increased roughly linearly with increasing channel frequency from a mean of 256 μs at 4 kHz to 298 μs at 14 kHz. This range of thresholds is similar to those measured by Bernstein and Trahiotis (1994) for noises of various bandwidths centered at 4 and 8 kHz, for which the mean jnd's were ~150 and ~300 μs, respectively. Our jnd estimates were very sensitive to the cutoff frequency of the low-pass filter used to extract the envelope; using the 250-Hz cutoff, Δ_{ITD} increased to 511 μs at 4 kHz and to 740 μs at 14 kHz. These values are substantially higher than those found psychophysically with bands of noise, suggesting that the 250-Hz low-pass filter cutoff was too low.

In the absence of DTF filtering, the right- and left-ear envelopes were perfectly correlated, and therefore the reduction in estimated ITD acuity must have been caused by reduction in modulation depth, m. With 500-Hz low-pass filtering, m declined approximately linearly with increasing channel frequency from a mean of 0.65 (i.e., 65% modulation) in the 4-kHz channel to 0.43 in the 14-kHz channel. Reduction of the low-pass cutoff to 250 Hz lowered modulation depth to 0.53 at 4 kHz and to 0.34 at 14 kHz. These dependencies on frequency and low-pass cutoff were expected because of the increase of channel bandwidth with increasing center frequency. The high-frequency limit of channel envelope fluctuations is roughly proportional to channel bandwidth (Dau *et al.*, 1999; Lawson and Uhlenbeck, 1950), and therefore a fixed low-pass filter cutoff removes a greater proportion of the envelope fluctuation energy in the higher-frequency channels.

The addition of DTF filtering caused an increase in Δ_{ITD} of 50–100 μs for most listeners and most frequency channels. There was no consistent effect on modulation depth of listener, azimuth, or near/far ear, and therefore the loss of sensitivity must have been caused by decorrelation of the envelopes related to DTF asymmetry. This manifested itself in reduction of the peak value of the normalized correlation, which, however, remained >0.95 unless a pronounced spectral notch in either the left or right DTF was aligned with the center frequency of the filter channel. In such cases, the peak normalized correlation was further reduced, and typically fell in the range 0.90–0.95, presumably because of decorrelation of the right- and left-channel envelopes caused by the rapid changes in phase associated with the spectral notches. The consequence of this decorrelation was to increase Δ_{ITD} in the affected frequency bands by at least 200–300 μs.

An ITD jnd on the order of 250 μs corresponds to an MAA of approximately 25 degrees because ITD varies by ~10 μs per degree of lateral angle across the midline. Although one would expect that combining ITD information across frequency channels would lead to lower overall thresholds, such large MAAs do not appear small enough to permit accurate lateral localization. These simulation results suggest that, in agreement with the hypothesis of Middlebrooks and Green (1990), high-frequency Gaussian noise signal envelopes *cannot* convey robust ITD information.

[1]In effect, the ongoing portion of the noise burst was amplitude modulated at a depth of 100% by a 100-Hz, smoothed square-wave modulator. We computed the levels of the low-frequency sidebands generated by this modulator, and determined that any sidebands appearing below 1.5 kHz were at least 60 dB lower in level than components of the high-frequency passband. Bernstein and Trahiotis (1982) caution that low-frequency sidebands as low in level as 50 dB below the passband components of nominally high-pass signals can be informative for ITD discrimination. We believe, and the small ITD weights observed in the envelope-modulated conditions suggest, that the sidebands in our stimuli were low enough to avoid this problem.

Bernstein, L. R., and Trahiotis, C. (**1982**). ''Detection of interaural delay in high-frequency noise,'' J. Acoust. Soc. Am. **71**, 147–152.

Bernstein, L. R., and Trahiotis, C. (**1994**). ''Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise,'' J. Acoust. Soc. Am. **95**, 3561–3567.

Bernstein, L. R., and Trahiotis, C. (**1996**). ''On the use of the normalized correlation as an index of interaural envelope correlation,'' J. Acoust. Soc. Am. **100**, 1754–1763.

Blauert, J. (**1982**). ''Binaural localization: Multiple images and applications in room and electroacoustics,'' in *Localization of Sound: Theory and Applications*, edited by R. W. Gatehouse (Amphora, Groton, CT), pp. 3–25.

Blauert, J. (**1997**). *Spatial Hearing* (MIT, Cambridge, MA).

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

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

Brungart, D. S., Durlach, N. I., and Rabinowitz, W. M. (**1999**). ''Auditory localization of nearby sources. II. Localization of a broadband source,'' J. Acoust. Soc. Am. **106**, 1956–1968.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (**1991**). ''Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays,'' J. Acoust. Soc. Am. **90**, 3077–3085.

Butler, R. A., and Flannery, R. (**1980**). ''The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane,'' Percept. Psychophys. **28**, 449–457.

Carlile, S., Delaney, S., and Corderoy, A. (**1999**). ''The localisation of spectrally restricted sounds by human listeners,'' Hear. Res. **128**, 175–189.

Constan, Z. A., and Hartmann, W. M. (**2001**). ''Sound localization by interaural time differences at high frequencies,'' J. Acoust. Soc. Am. **109**, 2485.

Dau, T., Verhey, J., and Kohlrausch, A. (**1999**). ''Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers,'' J. Acoust. Soc. Am. **106**, 2752–2760.

Duda, R. O. (**1997**). ''Elevation dependence of the interaural transfer function,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. R. Anderson (Laurence Erlbaum, Mahwah, NJ), pp. 49–75.

Eberle, G., McAnally, K. I., Martin, R. L., and Flanagan, P. (**2000**). ''Localization of amplitude-modulated high-frequency noise,'' J. Acoust. Soc. Am. **107**, 3568–3571.

Fedderson, W. E., Sandel, T. T., Teas, D. C., and Jeffress, L. A. (**1957**). ''Localization of high-frequency tones,'' J. Acoust. Soc. Am. **29**, 988–991.

Freyman, R. L., Zurek, P. M., Balakrishnan, U., and Chiang, Y. C. (**1997**). ''Onset dominance in lateralization,'' J. Acoust. Soc. Am. **101**, 1649–1659.

Gaik, W. (**1993**). ''Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling,'' J. Acoust. Soc. Am. **94**, 98–110.

Hafter, E. R., and Carrier, S. C. (**1969**). ''Inability of listeners to trade completely interaural time for interaural intensity in a detection task,'' J. Acoust. Soc. Am. **49**, 125.

Hafter, E. R., and Dye, D. H. (**1983**). ''Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number,'' J. Acoust. Soc. Am. **73**, 644–651.

Hafter, E. R., and Jeffress, L. A. (**1968**). ''Two-image lateralization of tones and clicks,'' J. Acoust. Soc. Am. **44**, 563–569.

Hall, J. W., and Grose, J. H. (**1990**). ''Comodulation masking release and auditory grouping,'' J. Acoust. Soc. Am. **88**, 119–125.

Harris, G. G. (**1960**). ''Binaural interactions of impulsive stimuli and pure tones,'' J. Acoust. Soc. Am. **32**, 685–692.

Hartmann, W. M., and Constan, Z. A. (**1998**). ''Binaural coherence and the localization of sound in rooms,'' in *Proceedings of the 16th International Congress on Acoustics*, pp. 2889–2890.

Henning, G. B. (**1974**). ''Detectability of interaural delay in high-frequency complex waveforms,'' J. Acoust. Soc. Am. **55**, 84–90.

Hofman, P. M., and Van Opstal, A. J. (**1998**). ''Spectro-temporal factors in two-dimensional human sound localization,'' J. Acoust. Soc. Am. **103**, 2634–2648.

Humanski, R. A., and Butler, R. A. (**1988**). ''The contribution of the near and far ear towards localization of sound in the median plane,'' J. Acoust. Soc. Am. **83**, 2300–2310.

Kuhn, G. F. (**1987**). ''Physical acoustics and measurements pertaining to directional hearing,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 3–25.

Kulkarni, A., Isabelle, S. K., and Colburn, H. S. (**1999**). ''Sensitivity of human subjects to head-related transfer-function phase spectra,'' J. Acoust. Soc. Am. **105**, 2821–2840.

Lawson, J. L., and Uhlenbeck, G. E. (**1950**). *Threshold Signals, Vol. 24 of Radiation Laboratory Series* (McGraw–Hill, New York).

Leakey, D. M., Sayers, B. M., and Cherry, C. (**1958**). ''Binaural fusion of low- and high-frequency sounds,'' J. Acoust. Soc. Am. **30**, 222.

Macpherson, E. A. (**1996**). ''Effects of source spectrum irregularity and uncertainty on sound localization,'' J. Acoust. Soc. Am. **99**, 2515–2529.

Macpherson, E. A. (**1998**). ''Spectral cue processing in the auditory localization of sounds with wideband non-flat spectra,'' Ph.D. thesis, University of Wisconsin—Madison.

Macpherson, E. A., and Middlebrooks, J. C. (**2000**). ''Localization of brief sounds: Effects of level and background noise,'' J. Acoust. Soc. Am. **108**, 1834–1849.

McFadden, D., and Pasanen, E. G. (**1976**). ''Lateralization at high frequencies based on interaural time differences,'' J. Acoust. Soc. Am. **59**, 634–639.

Middlebrooks, J. C. (**1999a**). ''Individual differences in external-ear transfer functions reduced by scaling in frequency,'' J. Acoust. Soc. Am. **106**, 1480–1492.

Middlebrooks, J. C. (**1999b**). ''Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency,'' J. Acoust. Soc. Am. **106**, 1493–1510.

Middlebrooks, J. C., and Green, D. M. (**1990**). ''Directional dependence of interaural envelope delays,'' J. Acoust. Soc. Am. **87**, 2149–2162.

Morimoto, M. (**2001**). ''The contribution of two ears to the perception of vertical angle in sagittal planes,'' J. Acoust. Soc. Am. **109**, 1596–1603.

Morimoto, M., and Aokata, H. (**1984**). ''Localization cues of sound sources in the upper hemisphere,'' J. Acoust. Soc. Jpn. **5**, 165–173.

Moushegian, G., and Jeffress, L. A. (**1959**). ''Role of interaural time and intensity differences in lateralization of low-frequency tones,'' J. Acoust. Soc. Am. **31**, 1441–1445.

Nuetzel, J. M., and Hafter, E. R. (**1981**). ''Discrimination of interaural delays in complex waveforms: Spectral effects,'' J. Acoust. Soc. Am. **69**, 1112–1118.

Palmer, A. R., and Russell, I. J. (**1986**). ''Phase-locking in the cochlear nerve of the guiniea-pig and its relation to the receptor potential of inner hair cells,'' Hear. Res. **24**, 1–15.

Pinheiro, M. L., and Tobin, H. (**1969**). ''Interaural intensity difference for intercranial lateralization,'' J. Acoust. Soc. Am. **46**, 1482–1487.

Rakerd, B., Hartmann, W. M., and McCaskey, T. L. (**1999**). ''Identification and localization of sound sources in the median saggital plane,'' J. Acoust. Soc. Am. **106**, 2812–2820.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). ''Phase locked responses to low-frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 769–793.

Saberi, K. (**1995**). ''Lateralization of comodulated complex waveforms,'' J. Acoust. Soc. Am. **98**, 3146–3156.

Sandel, T. T., Teas, D. C., Fedderson, W. E., and Jeffress, L. A. (**1955**). ''Localization of sound from single and paired sources,'' J. Acoust. Soc. Am. **27**, 842–852.

Searle, C. L., Braida, L. D., Cuddy, D. R., and Davis, M. F. (**1975**). ''Binaural pinna disparity: Another auditory localization cue,'' J. Acoust. Soc. Am. **57**, 448–455.

Simon, H. J., and Aleksandrovsky, I. (**1997**). ''Perceived lateral position of narrow-band noise in hearing-impaired and normal-hearing listeners under conditions of equal sensation level and sound-pressure level,'' J. Acoust. Soc. Am. **102**, 1821–1826.

Slaney, M. (**1994**). ''Auditory toolbox: A MATLAB toolbox for auditory modeling work,'' Technical Report 45, Apple Computer.

Slattery, W. H. I., and Middlebrooks, J. C. (**1994**). ''Monaural sound localization: Acute versus chronic impairment,'' Hear. Res. **75**, 38–46.
Strutt, J. W. (**1907**). ''On our perception of sound direction,'' Philos. Mag. **13**, 214–232.
Tobias, J. V., and Schubert, E. D. (**1959**). ''Effective onset duration of auditory stimuli,'' J. Acoust. Soc. Am. **31**, 1595–1605.
Trahiotis, C., and Bernstein, L. R. (**1986**). ''Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth,'' J. Acoust. Soc. Am. **79**, 1950–1957.
Trahiotis, C., and Stern, R. M. (**1994**). ''Across-frequency interaction in lateralization of complex binaural stimuli,'' J. Acoust. Soc. Am. **96**, 3804–3806.
Whitworth, R. H., and Jeffress, L. A. (**1961**). ''Time vs intensity in the localization of tones,'' J. Acoust. Soc. Am. **33**, 925–929.
Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661.
Wightman, F. L., and Kistler, D. J. (**1997a**). ''Factors affecting the relative salience of sound localization cues,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Laurence Erlbaum, Mahwah, NJ), pp. 1–23.
Wightman, F. L., and Kistler, D. J. (**1997b**). ''Monaural localization revisited,'' J. Acoust. Soc. Am. **101**, 1050–1063.
Wightman, F. L., and Kistler, D. J. (**1999**). ''Sound localization with unilaterally degraded spectral cues,'' J. Acoust. Soc. Am. **105**, 1162.
Zhou, B., Green, D. M., and Middlebrooks, J. C. (**1992**). ''Characterization of external ear impulse responses using Golay codes,'' J. Acoust. Soc. Am. **92**, 1169–1171.
Zwislocki, J., and Feldman, R. S. (**1956**). ''Just noticeable differences in dichotic phase,'' J. Acoust. Soc. Am. **28**, 860–864.