The effects of attenuation of frequency segments on binaural localization of sound

JASON A. BURLINGAME and ROBERT A. BUTLER
University of Chicago, Chicago, Illinois

Perceived location of tonal stimuli d narrow noise bands presented in two-dimensional space varies in an orderly manner with changes in stimulus frequency. Hence, frequency has a referent in space that is most apparent during monaural listening. The assumption underlying the present study is that maximum sound pressure level measured at the ear canal entrance for the various frequencies serves as a prominent spectral cue for their spatial referents. Even in binaural localization, location judgments in the vertical plane are strongly influenced by spatial referents. We measured sound pressure levels at the left ear canal entrance for 1.0-kHz-wide noise bands, centered from 4.0 kHz through 10.0 kHz, presented at locations from 60° through −45° in the vertical plane; the horizontal plane coordinate was fixed at −90°. On the basis of these measurements, we fabricated three different band-stop stimuli in which differently centered 2.0-kHz-wide frequency segments were filtered from a broadband noise. Unfiltered broadband noise served as the remaining stimulus. Localization accuracy differed significantly among stimulus conditions ($p < .01$). Where in the vertical plane most errors were made depended on which frequency segment was filtered from the broadband noise.

A generally accepted statement by those studying localization of sound in space is that interaural temporal and level differences are the major cues for localizing sound in the horizontal plane (HP) and that spectral cues dominate for localizing sounds in the vertical plane (VP) (Middlebrooks & Green, 1991). The spectral cues are those generated primarily by the pinna's interaction with the incident sound wave. The pinna, because of its small size, interacts effectively only with sounds of short wave lengths. Unless the higher frequencies are contained in the sound's spectrum, the stimulus cannot be localized accurately. For example, normally hearing persons are unable to localize sounds accurately in the median sagittal plane when the higher frequencies (>3.0 kHz) are removed (Butler & Humanski, 1992); patients with high frequency hearing losses perform at chance level (Butler, 1970). The importance of the higher auditory frequencies in localization may be attributable to the fact that different segments within this frequency range have different referents in space. With source location fixed in the median sagittal plane, the apparent location of tonal stimuli or narrow noise bands changes with changes in frequency content (Blauert, 1969; Pratt, 1930; Roffler & Butler, 1968). We call the apparent location of a stimulus of a given frequency the spatial referent of that frequency, which in many respects is comparable to Blauert's "directional band." Usually, the spatial referent for a particular

frequency is distributed within a restricted region of space, not necessarily confined to a specific location. Hence, the *spatial referent* is defined as the mean of the independent location judgments for the stimulus frequency in question. If a given frequency band is notched from a broadband noise, monaural listeners are less likely to judge the sound as emanating from the spatial referent of the notched frequency segment (Butler & Flannery, 1980; Rogers & Butler, 1992). Butler and Musicant (1993) demonstrated that even when interaural temporal and level difference cues were available, notch-filtering a broadband noise resulted in a significant decrease in the number of location judgments directed toward the spatial referent of the filtered frequency band.

A linkage exists between a behavior variable (spatial referent) and a physical variable (maximum sound pressure levels). Consider first the spatial referent. We know, for example, that the spatial referents of noise bands (1.0 kHz wide) centered at 5.0 to about 9.5 kHz lie above 0° in the VP and that the spatial referents of noise bands centered at about 9.5 through 14.0 kHz lie below 0° in the VP. This phenomenon is illustrated in the upper panel of Figure 1 (see Butler, 1997). Here, the migration of the mean monaural location judgments (spatial referents) of narrow noise bands for Listener V.C. is charted when a band (1.0 kHz wide) was presented at a fixed position in space (−90° in the HP; 0° in the VP). As the center frequency of the noise band was increased, the apparent location of the stimulus moved from the front toward the side and upward. At a center frequency of 10.0 kHz, the apparent location dropped below 0° in the VP. The data shown in lower panel of Figure 1 are based on sound pressure levels measured at the entrance of the external ear canal. We derived from these measurements an entity that we call the "covert
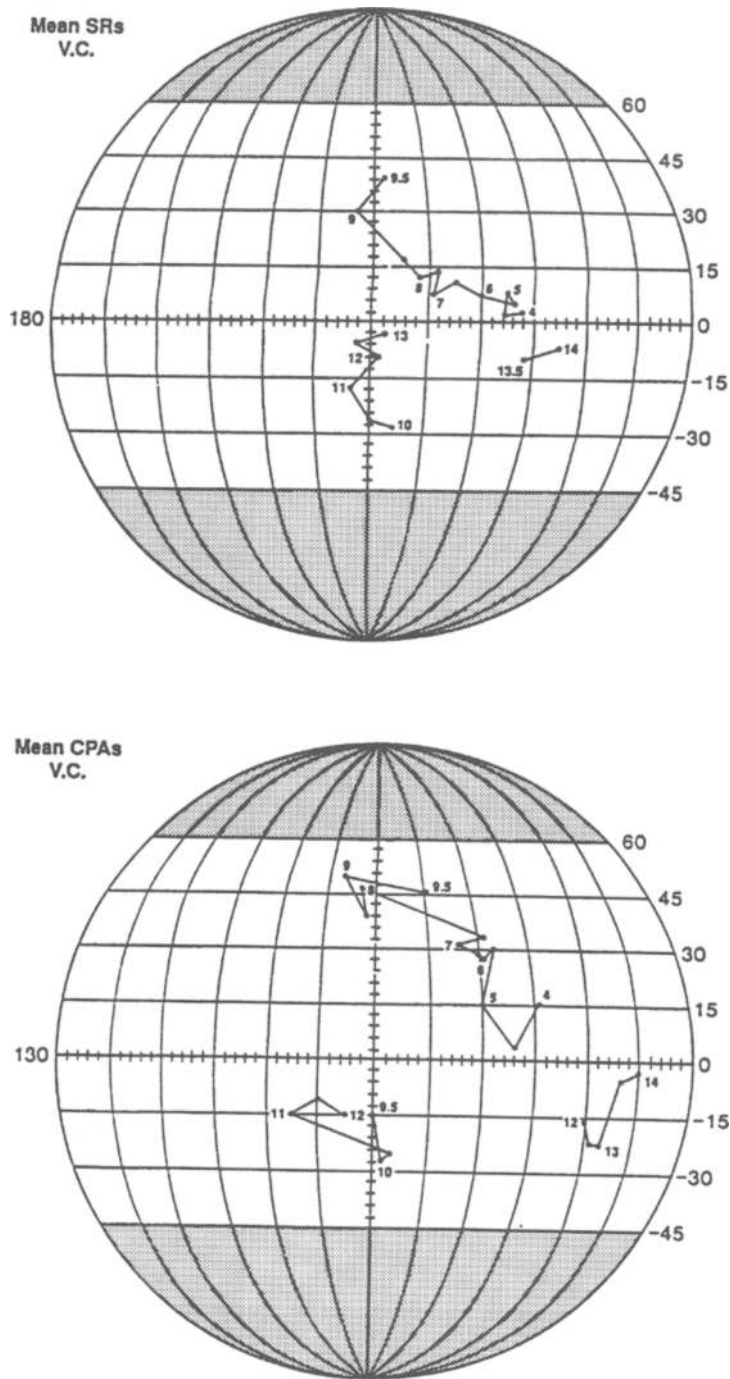
1374

Figure 1. Upper panel: The pattern of mean SRs as a function of the center frequency of a 1.0-kHz-wide noise band for Listener V.C. Lower panel: The pattern of mean covert peak areas (CPAs) as a function of the center frequency of a 1.0-kHz-wide noise band for Listener V.C. (from Butler, 1997).

peak areas" (CPAs). The CPA is defined as the mean of the spatial locations giving rise to outputs within 1 dB of the maximum output obtained from all locations.[1] As is evident in the lower panel, the spatial migration of the mean CPA follows a pattern similar to that of the spatial

referent as the center frequency of the narrow noise band is increased. Note in particular the abrupt drop in elevation of the mean CPA when the center frequency is increased from 9.0 to 10.0 kHz. We call this discontinuity in the pattern of spatial referents and the pattern of maximum

sound pressure levels the "breakpoint." This pattern of sound pressure levels is in accord with those iso-level contours published by Middlebrooks, Makous, and Green (1989). The assumption underlying the present study is that the place of maximum sound pressure for a given frequency segment is the spectral cue, which then serves as the spatial referent for the stimulus.

One further set of measurements warrants description since they are directly relevant to our experiment. Musicant, in establishing background material preceding the Butler–Musicant study (1993), measured head-related-transfer-functions on 11 subjects. Data derived from these transfer functions for Listener A.B., who was representative of the group, are illustrated in Figures 2, 3, and 4. Here, sound pressure levels are plotted as a joint function of the HP and VP coordinates of the sound source, with center frequency of the 1.0-kHz-wide noise band serving as the parameter. Notice in Figure 2 the changes in sound pressure levels that occurred when the HP component of the sound source was placed at −90° (see the heavy line). There is a distinct trough in the sound pressure level associated with a 1.0-kHz-wide noise band centered at 8.5 kHz when that stimulus was moved from 60° to −45° elevation.[2] Sound pressure levels at each side of the trough are comparable to each other. When the center frequency of the noise band was reduced by 2.0 kHz, sound pressure levels generated at higher elevations exceeded those generated at lower elevations (Figure 3). The converse was observed when the center frequency of the noise band was increased by 2.0 kHz (Figure 4). In accordance with our thesis that place of maximum sound pressure level serves as the spectral cue for spatial referents, we assert that spatial referents both "high" and "low" are represented in Figure 2; that spatial referents for "high" are predominantly represented in Figure 3; and that only spatial referents for "low" are predominantly represented in Figure 4. When the data are plotted as in Figure 1, the discontinuity (breakpoint) is clear-cut. When the data are plotted as in Figures 2, 3, and 4, the discontinuity is expressed as a trough in the pattern of sound pressure levels. These are simply different ways of illustrating the phenomenon.

The present experiment was specifically motivated by a post hoc finding of Butler and Musicant (1993). Their data suggested that a notch-filtered broadband noise was more difficult to localize (larger errors) when the center frequency of the notch was near the breakpoint rather than farther away. Here, we further probe that result.

METHOD

Subjects
Eleven normally hearing college students participated in each stage of the experiment. Their auditory thresholds were within 15 dB of audiologic zero (ISO; Davis & Krantz, 1964) for frequencies ranging from 0.25 through 8.0 kHz.

Test Room
The acoustic measurements and sound localization tests were conducted in a sound-treated room, 4.8 × 4.1 × 2.4 m. The floor was carpeted, a heavy curtain was positioned behind the loudspeaker array, and the walls were covered by loose felt, 10 cm in thickness contained within a fine wire mesh. Supported by a framework of small-diameter tubing bent to specifications were 104 midrange tweeters (Realistic, Model 40-1289) with a nominal frequency range of 0.3 to 20 kHz. They were selected from a pool of 220 on the basis of similarity in their frequency characteristics. An equalizer (Rane, Model PE-15) was employed to flatten the frequency response curves. Within the range of 0.3 to 15.0 kHz, the loudspeaker response to a broadband noise (Grason–Stadler noise generator Model 1285) was ±3 dB. The loudspeakers were arrayed in a 13 × 8 matrix, each loudspeaker separated from an adjacent one by 15° within each row and within each column. They extended from 0° (straight ahead) to 180° (directly behind) in the horizontal plane. All loudspeakers were positioned on the subjects' left side. In the vertical plane, the loudspeakers extended from 60° to −45°, with 0° approximating the interaural axis.

Acoustic Measurements
We first indentified the frequency band that was associated with a distinct trough in sound pressure level as that band was presented sequentially from 60° to −45° in the VP. This trough would mark the transition between maximum sound pressure levels generated when the sound was presented from above 0° elevation and those generated when the sound was presented from below 0° elevation (see Figure 2). The left ear canal was occluded with a malleable plastic material (Audi-Sil), and a miniature microphone (Knowles, Model EA-1934) was embedded in the plastic with its diaphragm facing outward. Since we knew from Musicant's measurements (Butler & Musicant, 1993) that a prominent trough occurs when the sound source is in the vicinity of 90° in the HP, we measured the microphone output for narrow noise bands originating only from this azimuth with the VP component ranging from 60° to −45°. As apparent from the plot shown in Figure 2, we could have chosen other HP coordinates to identify the frequency segment showing a clear-cut trough as the sound was moved from above to below in the VP, but the choice of −90° was satisfactory for our purposes.

Subjects were seated in a chair positioned 1.6 m from the column of loudspeakers delivering the sound. They were urged not to move their heads while sound pressure measurements were in progress. A headrest attached to the back of the chair promoted compliance. Chair height could also be adjusted vertically. A 1.0-kHz-wide noise band centered at 5.0 kHz was presented successively from each of the eight loudspeakers beginning with the one positioned at 60° in the VP and −90° in the HP and ending with the loudspeaker positioned at −45° in the VP—the HP coordinate remaining at −90°. Then with the center frequency fixed at 6.0 kHz, we repeated the procedure. With one exception, the center frequency of the noise band was increased in steps of 1.0 kHz until reaching 10.0 kHz. The output of the microphone was amplified and delivered to a voltmeter whose output was recorded in dB. A stimulus level was chosen to insure that the lowest registration on the voltmeter in response to the stimulus was at least 6 dB above the noise level. We were searching for the center frequency at which a distinct trough occurred in sound pressure level as the stimulus elevation was lowered from 60° to −45°. Knowing the actual sound pressure levels was not critical for identifying the trough in the sequence of sound pressure level measurements. For each center frequency delivered from each of the eight loudspeakers, we recorded the sound pressure level. When the center frequency was increased to around 9.0 kHz and the stimulus was presented sequentially from above to below 0° in the VP, the output of the miniature microphone declined to a minimum—usually at a VP position of −15°—and then increased when the narrow noise band was presented at the next two lower elevations. To more precisely identify the center frequency at which this trough would be most prominent, we changed the center frequency of the noise band in 0.5-kHz steps and re-
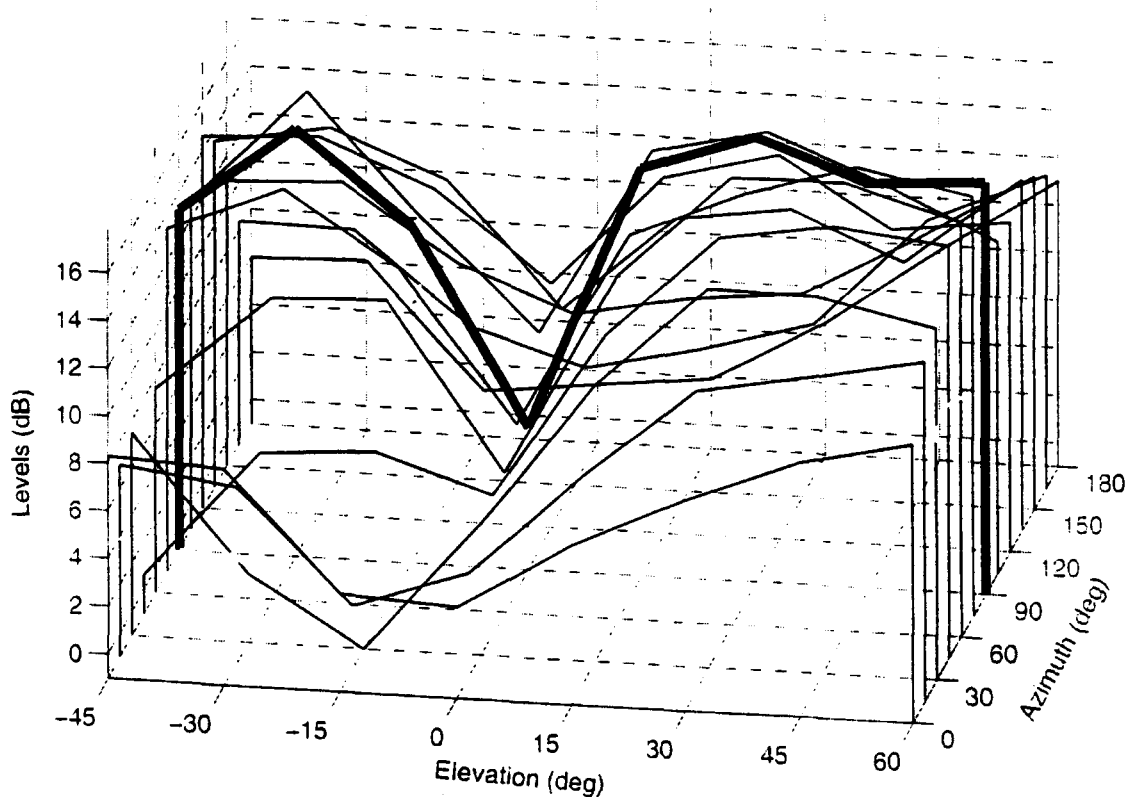
Figure 2. Total energy recorded as a function of sound location. These data were derived from head-related-transfer functions (HRTFs) measured at the left ear canal entrance of Listener A.B. Center frequency of a 1.0-kHz-wide noise band was set at 8.5 kHz. The heavy line represents the measurements when the HP coordinate was −90°.

peated the measurements at center frequencies surrounding that at which the trough first appeared.

Our experimental design called for a broadband noise and three different bandstop noises in which a 2.0-kHz-wide notch was embedded. For one bandstop stimulus, the center frequency of the notch was to be fixed at the center frequency of the narrow noise band associated with the distinct trough in the sound pressure levels as the sound was delivered at the various locations in the vertical plane. This sound is referred to as the target (T) stimulus.

Table 1 provides an example of measurements obtained on Listeners S.B., J.H., and T.J. The relative dB readings in bold print identify the segment of center frequencies that best encompassed the transition in sound pressure measurements from peak to trough to peak. Consider the data for Listener S.B. Scanning from 60° to −45°, the microphone output exhibited for the first time an unmistakable trough when the center frequency was increased to 9.0 kHz; the output at −30° and −45° abruptly increased instead of continuing to decrease, in contrast to the microphone's output at the lower center frequencies. We reasoned that 9.0 kHz would be our best choice as the notch's center frequency. Sound pressure levels above and below the trough, while not equal, were similar. Centering a 2.0-kHz-wide notch in a broadband noise at 9.0 kHz would filter frequency segments whose spatial referents were assumed to be both above and below 0° in the VP. The pattern of the microphone output data for Listener J.H. resembled closely that of Listener S.B. Nine kHz was again chosen as the center frequency for the 2.0-kHz-wide notch representing the T stimulus. The pattern of the micro-

phone output for Listener T.J. also resembled that of the others, but the center frequency at which the trough was most apparent differed. We chose 8.0 kHz as the center frequency for a 2.0-kHz-wide notch. The T stimulus for the remaining listeners had center frequencies for the 2.0-kHz-wide notches at 8.5, 8.5, 9.0, 9.0, 9.5, 9.5, 9.5, and 9.5 kHz. The decisions on center frequencies of a notch that most adequately encompassed the breakpoint are open to argument. But, if we erred by 0.5 kHz, the breakpoints we chose for each listener would still have been encompassed by a 2.0-kHz-wide notch. The other two bandstop stimuli were characterized by a 2.0-kHz-wide notch in a broadband noise in which the center frequency of the notch was either 2 kHz below that for the T stimulus (the T − 2k stimulus), or the notch was centered 2.0 kHz above that for the T stimulus (the T + 2k stimulus). Notching the broadband noise was accomplished by delivering the stimulus to two filters (A. P. Circuit, Model 403) connected in series. This resulted in a symmetrical notch, 2.0-kHz-wide at the 3-dB down point. The depth of the notch at its center frequency was 15 dB.[3]

Behavioral Measurements

Listeners were positioned so that the loudspeaker array, described above, occupied the left hemifield. They reported the sound's location by means of a simple code—the columns of loudspeakers arranged in the HP were numbered from 1 through 13, with 1 positioned straight ahead and 13 positioned directly behind; the rows of loudspeakers arranged in the VP were numbered from 1 through 8 with 1 positioned at −45° elevation and 8 positioned at
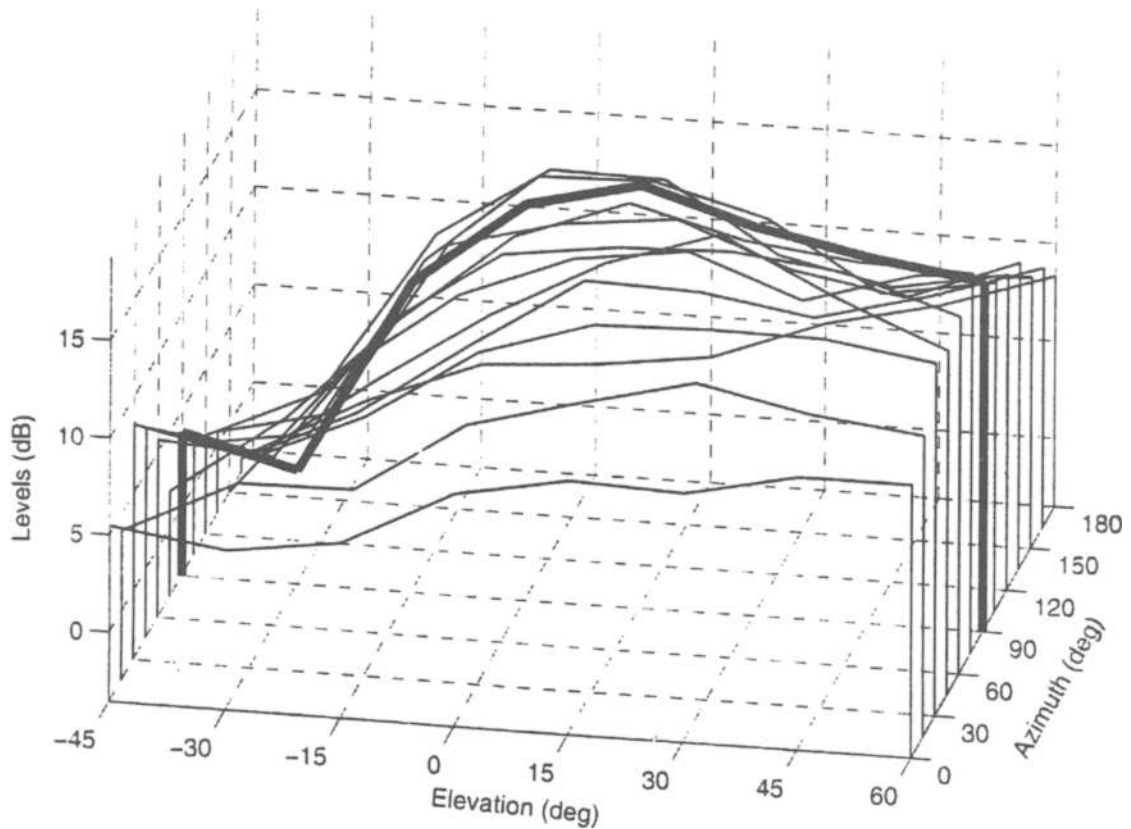
Figure 3. Total energy recorded as a function of sound location. These data were derived from head-related-transfer functions (HRTFs) measured at the left ear canal entrance of Listener A.B. Center frequency of a 6.5-kHz-wide noise band was set at 8.5 kHz. The heavy line represents the measurements when the HP coordinate was −90°.

60° elevation. Listeners were required to report the sound's coordinate in the HP first, then its coordinate in the VP. So for a sound originating at −105° in the HP (left hemifield) and 45° in the VP, its coded location would be 8:7; the coded location for a sound originating at −15° in the HP (left hemifield) and −15° in the VP would be 2:3. A brief practice period in which the experimenter pointed to different loudspeakers and the listener identified them in terms of the code sufficed to master response requirements.

First, thresholds were obtained (method of limits) for each of the four different stimuli whose duration was 100 msec in duration and whose rise–fall time was 10 msec. Interstimulus interval was 300 msec. Given that 104 loudspeakers were employed in the localization tests, taking thresholds for each loudspeaker was prohibitive. Hence we measured thresholds from only nine loudspeakers that were strategically positioned in the loudspeaker matrix—those whose coordinates in the HP were 0°, −90°, and 180° and whose coordinates in the VP were −45°, 0°, and 60°. Thresholds for the remaining 95 loudspeakers were estimated by means of interpolation.

All testing was conducted under binaural listening conditions. During the localization tests, stimulus level was 40 dB above measured or estimated threshold. Some of the estimated thresholds may have been in error by perhaps as much as 5 dB, thereby tempting listeners to use relative loudness as a cue for location notwithstanding its unreliability. Hence, ±3 dB re 40 dB SL changes were scattered throughout the presentation series according to a prearranged schedule. Before formal testing commenced, listeners received one

practice session. We chose for this purpose a stimulus that would not be used in the main study—a broadband noise from which a 2.0-kHz-frequency segment centered at 4.0 kHz was filtered out. For the practice and formal test sessions, the noise bursts continued until a listener made a location judgment. Listeners were admonished to keep their heads firmly against the headrest. Formal testing consisted of 16 test sessions; each session consisted of 104 trials, with every loudspeaker delivering the sound once in an irregular order. The sequence of presentations of the test sessions followed a 4 × 4 Latin square arrangement in which columns represented sessions, rows represented listeners, and Latin letters represented stimulus conditions. Four different schedules of presentation were employed. Knowledge of results of their performance was withheld.

Location judgments for the HP and VP coordinates of the stimuli were evaluated separately. We used error scores (ESs) as an index of accuracy. If a listener correctly identified both coordinates of the sound source, a 0 would be recorded for the HP and VP entries. If a listener chose a loudspeaker one, two, three (etc.) loudspeakers removed from the one generating the stimulus, an ES of 1, 2, 3 (etc.) would be recorded in the appropriate HP or VP ledger. Say, for example, that the sound originated from −60° in the HP (left hemifield) and 15° in the VP. For the response code that we followed, the correct response would have been 5-5. If the listener reported that the sound came from the loudspeaker positioned at −90° in the HP and 0° in the VP (i.e., 7-4), an error score of 2 would be recorded for the HP judgment and 1 for the VP judgment. Not
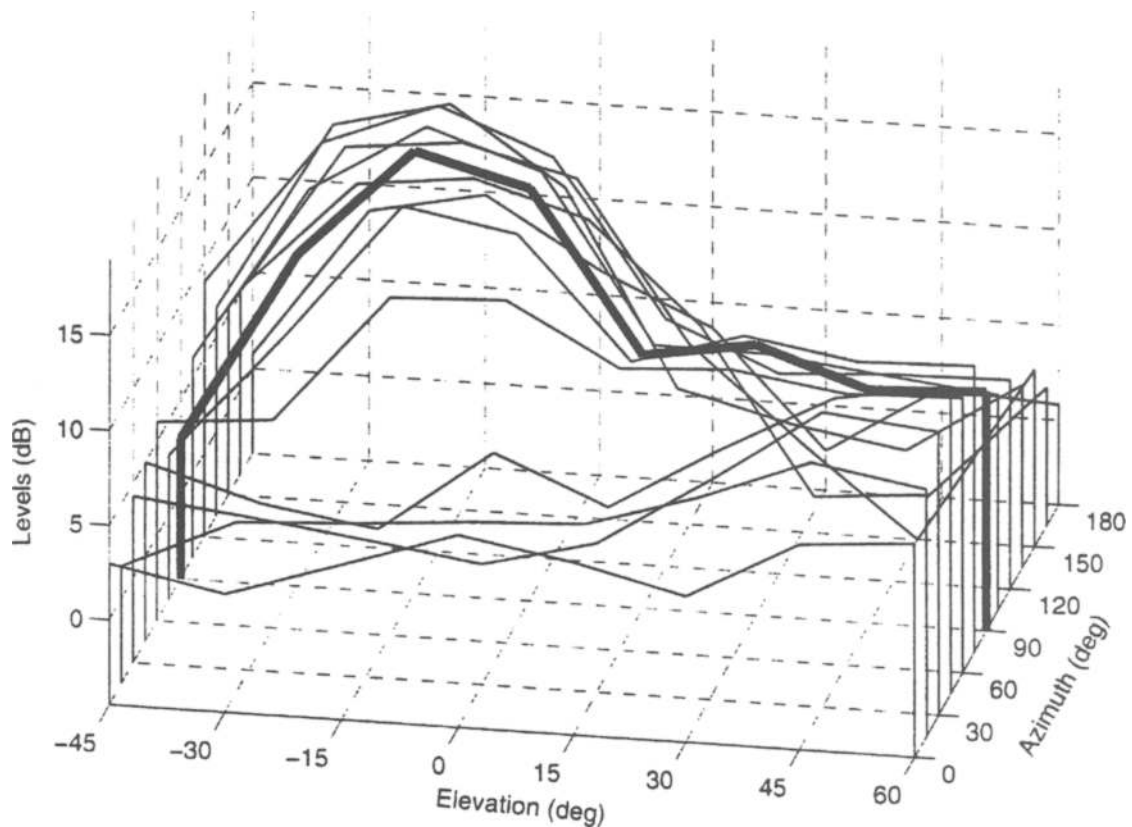
Figure 4. Total energy recorded as a function of sound location. These data were derived from head-related-transfer functions (HRTFs) measured at the left ear canal entrance of Listener A.B. Center frequency of a 10.5-kHz-wide noise band was set at 8.5 kHz. The heavy line represents the measurements when the HP coordinate was −90°.

infrequently, listeners report the HP coordinate of a sound as originating from the mirror image of the correct location; the so-called front–back reversal, and vice versa. We followed a common procedure of correcting for reversals. If, for example, the stimulus originated at −45° in the HP (left hemifield; 4 in the response code) and the listener reported the sound as coming from −135° in the HP (also left hemifield, but 10 in the response code), an error score of 0 would be recorded. Had the listener reported the HP coordinate as coming from 150° or 120° (11 or 9 in the response code), an ES of 1 would have been recorded. Rarely did our listeners reverse VP location (i.e., top–bottom reversals), so we did not need to apply a correction.

On the basis of the reasons put forth in the introduction, we predicted best performances when listeners were localizing the broadband stimulus and poorest performances when they were localizing the T stimulus.

RESULTS

Attenuating a 2.0-kHz frequency segment from the broadband stimulus did not result in a dramatic decline in localization accuracy. Nonetheless, location judgments were influenced in specific ways depending on the band-stop stimulus being employed. Figure 5 shows in histographic form mean errors in degree for localizing sounds

in the VP (upper panel) and HP (lower panel). Clearly, notch-filtering various segments of the spectrum exerted a much greater influence when subjects were localizing the VP coordinates. ESs were subjected to a repeated measures analysis of variance (ANOVA). Those recorded for the VP coordinates were analyzed separately from those recorded for the HP coordinates. With respect to the former, ESs among stimulus conditions differed significantly [$F(3,30) = 13.98$, $p < .01$]. Tukey's honestly significant difference (HSD) test was used to detect where, among stimulus conditions, a statistically significant difference occurred. Table 2 lists the outcome for the VP performances. A difference of 2.4° was deemed significant at the .05 confidence level. As can be seen, the significant differences took place between (1) the broadband stimulus and the T − 2k stimulus, (2) the broadband stimulus and the T stimulus, and (3) between the T stimulus and the T + 2k stimulus. The differences in rank between T − 2k and T + 2k were not statistically significant.

The location judgments for HP coordinates were first corrected for front–back reversals and then analyzed statistically. The results indicated that ESs among stimulus conditions also differed significantly [$F(3,30) = 6.31$,

Table 1
Sound Pressure Measurements at the External Ear Canal to 1.0-kHz-Wide Noise Bands Centered at Different Frequencies (CF$_{kHz}$) and Presented at Various Locations in the VP With the Location in the HP Fixed at −90°

VP Degrees	CF$_{kHz}$							
Listener S.B.								
	5.0	6.0	7.0	8.0	**8.5**	**9.0**	**9.5**	10.0
60	17.0	15.5	16.5	15.0	**14.5**	**13.5**	**15.0**	11.0
45	17.0	16.0	17.0	16.0	**16.0**	**16.5**	**15.0**	12.5
30	18.5	16.5	19.5	17.5	**16.5**	**17.5**	**16.5**	12.0
15	18.0	17.5	19.5	18.5	**16.5**	**15.0**	**13.5**	12.5
0	18.0	17.0	19.5	17.5	**16.5**	**16.0**	**10.5**	5.0
−15	16.5	16.5	18.0	15.5	**13.0**	**8.0**	**9.0**	11.5
−30	17.5	15.5	15.5	7.5	**10.0**	**14.0**	**16.5**	17.0
−45	14.0	13.5	11.5	7.0	**12.0**	**16.5**	**17.0**	15.5
Listener J.H.								
	5.0	6.0	7.0	8.0	**8.5**	**9.0**	**9.5**	10.0
60	16.0	14.0	16.0	13.0	**12.5**	**12.0**	**11.0**	9.0
45	17.0	14.0	17.0	14.0	**14.5**	**14.0**	**10.0**	9.5
30	17.0	15.0	18.5	16.0	**14.5**	**13.5**	**13.0**	8.0
15	18.5	14.5	17.5	16.5	**14.0**	**10.0**	**9.5**	6.0
0	19.0	15.5	17.5	14.5	**14.0**	**10.0**	**8.0**	3.0
−15	17.5	14.0	15.5	12.0	**9.5**	**6.0**	**5.0**	10.0
−30	17.0	12.5	13.5	5.0	**8.0**	**12.0**	**14.0**	15.0
−45	14.0	10.0	6.0	9.0	**9.5**	**15.5**	**16.5**	15.0
Listener T.J.								
	5.0	6.0	7.0	**7.5**	**8.0**	**8.5**	9.0	10.0
60	14.5	12.5	14.5	**14.5**	**13.0**	**11.0**	8.0	–
45	15.0	13.0	15.0	**16.0**	**13.5**	**12.0**	10.0	–
30	16.0	15.0	18.0	**18.0**	**16.0**	**13.5**	10.0	–
15	17.0	15.5	17.0	**18.5**	**16.0**	**10.0**	7.5	–
0	18.0	16.0	16.0	**15.0**	**12.0**	**8.0**	7.0	–
−15	17.0	15.0	13.0	**11.0**	**7.5**	**9.5**	10.5	–
−30	16.5	13.0	10.0	**11.5**	**14.0**	**16.0**	15.5	–
−45	13.0	10.5	11.0	**14.0**	**16.5**	**16.0**	15.0	–

Note—The dB readings in bold print bracket the breakpoint.

$p < .01$], but Tukey's HSD test failed to reveal a significant difference ($p > .05$) between the individual mean ESs. We next analyzed the HP data in terms of the number of front–back reversals recorded for each stimulus condition and found that the differences among conditions were again significant [$F(3,30) = 5.1, p < .01$]. Tukey's HSD test indicated that a statistically significant difference ($p < .05$) in the number of front–back reversals between the various stimulus conditions occurred only between that tabulated for the broadband stimulus and that tabulated for the T stimulus.

Since the influence of different bandstop-filtered noise bands was greater for the VP location judgments, these data were further segmented. The upper panel of Figure 6 shows mean ESs in degree when the sound originated from 30° and 45° in the VP ("high"), and the lower panel shows mean ESs in degree when the sound originated from −30° and −45° in the VP ("low"). A comparison between ESs for the T − 2k stimulus and the T + 2k stimulus is of interest. Greater ESs for the T − 2k stimulus than for the T + 2k stimulus were recorded when the sound originated from "high." When the sound originated from "low," the converse obtained. The differences in ESs

for the two stimulus conditions, although in the expected direction (based on the sound pressure measurements described in the Method section), did not reach statistical significance. The reason all ESs for 30° and 45° ("high") were greater than those for −30° and −45° ("low") was that the mean apparent location of all stimuli deviated progressively from the actual location of the sounds when the elevation of the latter moved from 15° to 60°. The slopes for the broadband and T + 2k stimuli overlapped and deviated less from actual location of the sound source than did the slopes for T and T − 2k stimuli. The slopes of the latter two functions also corresponded closely to each other.

As noted, the magnitude of ESs in the vertical plane for the T stimulus and the T − 2k stimulus differed significantly from the magnitude of the ESs for the broadband stimulus. Two questions arise. (1) Were these differences due to an increase in bias brought about by the attenuation of selected frequency bands, or were they due to an overall increase in response variability? We examined the frequency distribution of responses to the various loudspeaker elevations for the broadband stimulus, the T stimulus, and the T − 2k stimulus. We calculated the ratio of the variances for the T stimulus/broadband stimulus and for T − 2k stimulus/broadband stimulus. In neither case were the F values significant for a one-tailed test ($p > .05$). (2) Were these increases in ESs concentrated at or near −90° in the HP—the coordinate where the sound pressure levels were measured that then served to tailor the stimuli—Or were they distributed more or less evenly across the HP? The upper panel of Figure 7, shows VP performance data for the broadband stimulus—the reference stimulus condition. Here, mean ESs (in degrees) in the VP are shown as a function of HP source location. Save for the 0° HP location, ESs tended to increase as the source location was moved toward the rear. The middle panel shows the ratio for ESs (the T stimulus/broadband stimulus); the lower panel shows the ratio for ESs (the T − 2k stimulus/broadband stimulus). With performances on the broadband stimulus serving as a reference, ES increases in the VP for the T and T − 2k stimuli were distributed across HP locations. And whereas the T/broadband ratio peaked at 90° azimuth, the magnitude of the ratio did not differ significantly from those of the neighboring azimuthal positions.

DISCUSSION

This study goes a step beyond the general statement that interaural temporal and level differences predominantly govern localization of sound in the HP and spectral cues predominantly govern localization of sound in the VP. Our results call for a refinement in thinking about the role of spectral cues in VP localization. All segments within the higher auditory frequencies do not exert equal influence on location judgments. The frequency segment encompassing the 5.0-kHz to 10-kHz range is especially important for localization accuracy, for when 2.0-kHz-wide
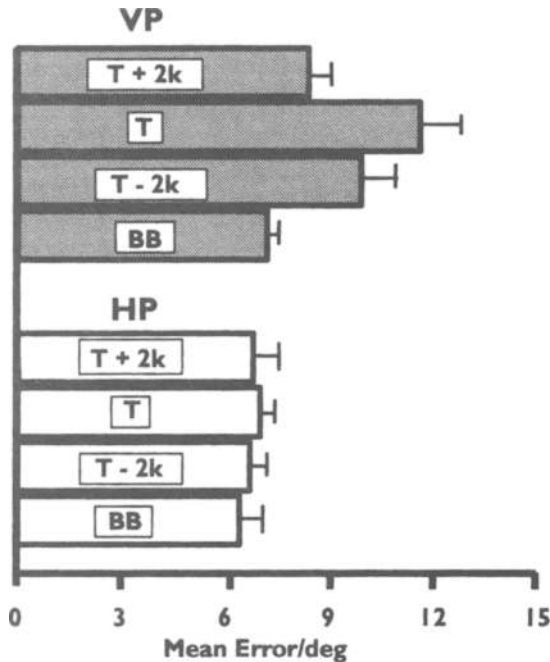
VP

Figure 5. Upper panel: Mean error scores (ESs) in degrees when localizing the four stimuli in the vertical plane (VP). Lower panel: Mean ESs in degrees when localizing the four stimuli in the horizontal plane (HP). See text for description of stimuli. Standard error of the mean is extended from each bar.

segments within this range were filtered, performance declined significantly. What must be taken into account is the spatial referents of the various frequency segments. The filtered segment of the T stimulus possessed spatial referents that lay in the transition zone between frequency segments whose spatial referents were "high," on the one hand, and "low," on the other. Hence, the spectral information about the sound's location when it originated both above and below 0° in the VP was degraded. In our view, degradation of spectral information is tantamount to a loss in the clarity of a spatial referent. The frequency segment that serves as the basis of the spatial referent was no longer prominent since the sound pressure level of the segment had been attenuated through filtering. Performance proficiency suffered accordingly. For the T − 2k stimulus, spectral information when the sound emanated from above 0° was degraded; localization accuracy declined for the sounds originating in the upper region of the VP. For the T + 2k stimulus, spectral information on sound coming from the lower region in the VP was somewhat degraded in that localization accuracy, for this spatial area declined when the higher audio frequencies were filtered. This interpretation is consonant with the performance data shown in Figure 6. It should be pointed out, however, that overall accuracy in localizing the T + 2k stimulus approached that recorded for the broadband stimulus. As mentioned in connection with Figure 6 in the Results section, apparent location of the

sounds deviated progressively from their actual location as the latter reached the higher elevations. This event has been reported before (Butler, Humanski, & Musicant, 1990), and we assume that spectral cues provided by the pinnae are simply less effective at these higher elevations.

Recall that all sound pressure measurements from which the various bandstop stimuli were constructed were carried out at one fixed HP location—that is, −90° azimuth. Yet, the ESs in the VP were distributed across HP locations, not just concentrated at or near −90° azimuth. Perusal of Figures 2, 3, and 4 suggests an explanation for the widespread distribution of VP errors across azimuth positions; the pattern of the sound pressure levels as a function of source elevation is similar irrespective of the HP coordinate of the source. Hence, attenuating a particular frequency segment whose spatial referent is "high" in space diminishes the influence of this segment on VP location judgments for most, if not all, sounds originating in the upper regions of the VP. By the same token, attenuating a frequency segment whose spatial referent is "low" in space affects adversely location judgments of sounds originating in the lower regions of the VP.

As mentioned in the Results section, there were no differences in ESs among stimulus conditions when listeners reported the HP coordinate of the sounds. We attribute this lack of effect to the fact that both interaural temporal and level differences were available for location judgments. We prefer not to speculate on why the number of front–back reversals was greatest for the T stimulus, since we are unaware of other data that would imply that front–back reversals are directly linked with the stimulus condition that characterized the T stimulus. The fact that the T stimulus was more difficult to localize in the vertical plane cannot, in itself, account for an increase in front–back reversals, but it may have contributed the confusion.

Several researchers have proposed that the trough in the sound spectrum is the most important cue for the perceived elevation of a sound. An illusion of a change in source elevation can be accomplished simply by manipulating the stimulus either in the time domain or the frequency domain. The former was accomplished by means of a two delay and add system (Bloom, 1977; Watkins, 1978); the latter was accomplished through the use of filters (Hebrank & Wright, 1974). What happens is that as the center frequency of a spectral notch embedded in a wideband noise is increased, the listener perceives the

Table 2
Tukey's HSD Post Hoc Analysis of
Mean Error Scores (in Degrees) in the Vertical Plane

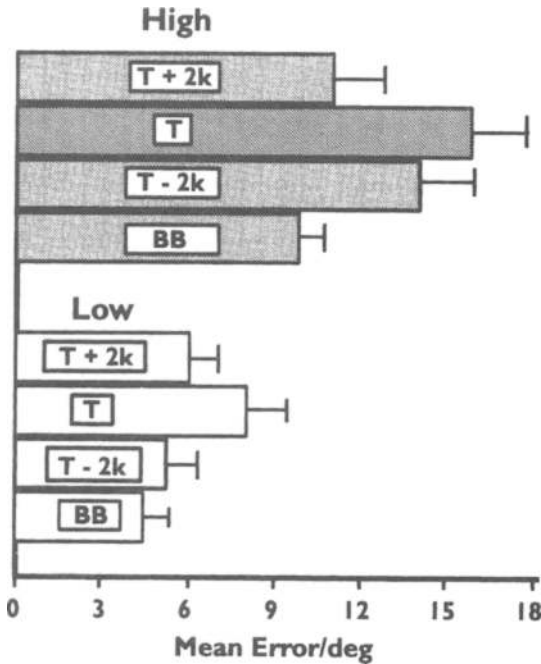| Stimulus Condition	Stimulus Condition			
	Broadband	T − 2k	T	T + 2k
Broadband	–	2.7*	4.2*	1.1
T − 2k		–	1.5	1.6
T			–	3.1*
T + 2k				–

*p < .05.

Figure 6. Upper panel: Mean ESs for the four stimuli when the sound originated from 30° and 45° ("high") in the vertical plane (VP). Lower panel: Mean ESs for the four stimuli when the sound originated from −30° and −45° ("low") in the VP. Standard error of the mean is extended from each bar.

sound as increasing in elevation. Butler and Belendiuk (1977), recording the output of a miniature microphone placed in the external ear canal, also observed a comparable relation between the center frequency of the spectral notch embedded in a broadband noise and the actual elevation of the stimulus. Was it likely that the spectral notch that we fabricated in the current study produced false cues for location, thereby resulting in decreased performance accuracy? We think not. First, a notch, by definition, is a place between two peaks in the frequency domain; changing the frequency position of the notch changes the frequency position of the peaks accordingly. Consider the frequency range from 6.0 to 12.0 kHz. When the center frequency of a notch is increased from 6.0 kHz, the spectral energy initially is greater for the higher frequencies—those whose spatial referents are below 0° in the VP. But as the center frequency is further increased, the higher frequencies within this range are progressively attenuated while the spectral energy at the lower frequencies is relatively increased. In the context of our study, attenuating the lower frequencies is tantamount to degrading the spatial referents for stimuli originating above 0° in the VP. Such changes in the relative influence of spatial referents are consistent with the finding that increasing the center frequency of a spectral notch generates an illusion that the sound is emanating from progressively higher locations in the VP. It is well documented that a pronounced spectral notch normally

occurs as a consequence of directional filtering by the pinna. We propose that the importance of the notch is that of informing the listener where the sound is *not*. For the listener, information provided by a notch is of equal importance to that provided by a peak. Perceptually, the listener homes in on a restricted region of space, and localization errors are eliminate or greatly reduced.

We have discussed spatial referents as located "high" or "low" in the VP, however, we have not considered the frequency segment whose spatial referent is in the middle— 0° in the VP. Figure 1, upper panel, shows that narrow noise bands centered at 4.0 and 5.0 kHz are perceived to originate around 0° as well as narrow noise bands centered at 13.0 kHz and above. An early study by Roffler and Butler (1968) indicated that tonal frequencies less than 2.0 kHz were perceived to lie at or near 0° elevation.
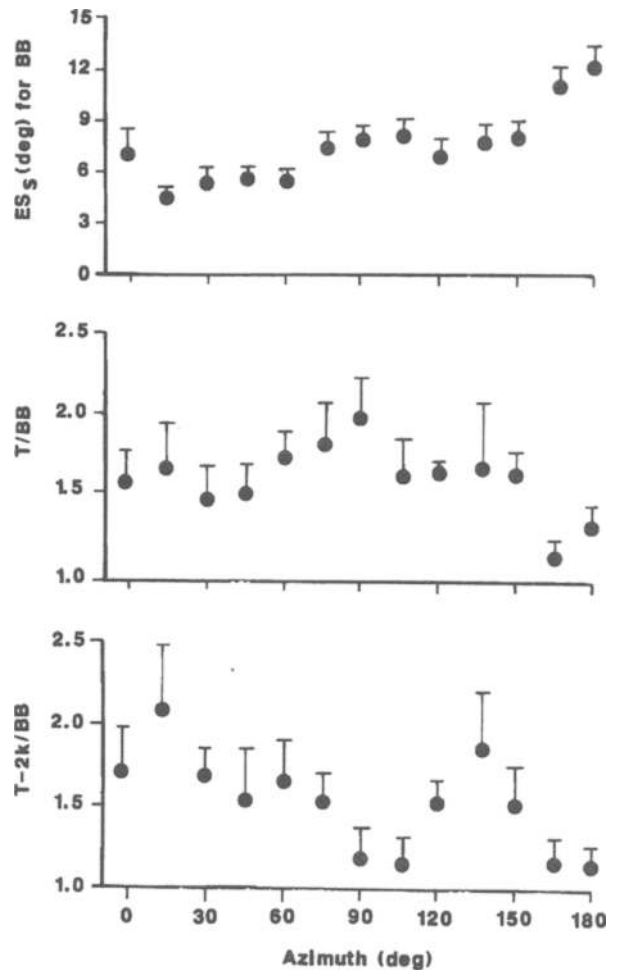

Figure 7. Upper panel: Mean error scores (ESs) in the vertical plane for the broadband (BB) stimulus as a function of source location in the horizontal plane (HP). Middle panel: Ratio of mean ESs (T/BB) in the vertical plane (VP) as a function of source location in the HP. Lower panel: The ratio of mean ESs [(T − 2k)/ BB] in the VP as a function of source location in the HP. Standard error of the mean is extended above each point.

From one point of view, spatial referents of stimulus frequency are annoying perceptual biases that confound any direct relation between actual and perceived location of a sound in space. In effect, they consist of another set of auditory illusions, since under monaural localization conditions, as well as binaural conditions where the stimuli emanate in the medial sagittal plane, the sound is perceived to come from a location other than its place of origin. It was because of the ubiquity of this auditory phenomenon that one of the authors (R.A.B.) thought that spatial referents must be directly implicated in the sound localization process. With respect to the biological underpinnings of sound localization, interaural temporal and level differences have been the object of intense neuroanatomical and neurophysiological study. Stimulus frequency, on the other hand, has been placed in a different category of auditory experience—pitch. The discovery of multiple tonotopic maps in the auditory cortex lent further support for the place theory of pitch perception. But frequency also has a referent in space. Perhaps one of these tonotopic maps is dedicated to processing the input of stimulus frequency in the localization of sound.

REFERENCES

BLAUERT, J. (1969). Sound localization in the median plane. *Acoustica*, 22, 205-213.

BLOOM, P. J. (1977). Creating source elevation illusions by spectral manipulation. *Journal of the Audio Engineering Society*, 25, 560-565.

BUTLER, R. A. (1970). The effect of hearing impairment on locating sound in the vertical plane. *International Audiology*, 9, 117-126.

BUTLER, R. A. (1997). Spatial referents of stimulus frequencies: Their role in sound localization. In R. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing* (pp. 99-115). Hillsdale, NJ: Erlbaum.

BUTLER, R. A., & BELENDIUK, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *Journal of the Acoustical Society of America*, 61, 1264-1269.

BUTLER, R. A., & FLANNERY, R. (1980). The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane. *Perception & Psychophysics*, 28, 449-457.

BUTLER, R. A., & HUMANSKI, R. A. (1992). Localization of sound in the vertical plane with and without high-frequency spectral cues. *Perception & Psychophysics*, 51, 182-186.

BUTLER, R. A., HUMANSKI, R. A., & MUSICANT, A. M. (1990). Binaural and monaural localization of sound in two-dimensional space. *Perception*, 19, 241-256.

BUTLER, R. A., & MUSICANT, A. M. (1993). Binaural localization: In-

fluence of stimulus frequency and the linkage to covert peak areas. *Hearing Research*, 67, 220-229.

CARLILE, S. (1990). The auditory periphery of the ferret. II: The spectral transformations of the external ear and their implications for sound localization. *Journal of the Acoustical Society of America*, 88, 2196-2204.

DAVIS, H., & KRANTZ, F. W. (1964). The international standard reference zero for pure tone audiometers and its relation to the evaluation of impairment of hearing. *Journal of Speech & Hearing Research*, 7, 7-16.

HEBRANK, J., & WRIGHT, D. (1974). Spectral cues used in the localization of sound sources on the median plane. *Journal of the Acoustical Society of America*, 56, 1829-1834.

MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, 42, 135-159.

MIDDLEBROOKS, J. C., MAKOUS, J. C., & GREEN, D. M. (1989). Directional sensitivity of sound-pressure levels in the human ear canal. *Journal of the Acoustical Society of America*, 86, 89-108.

MUSICANT, A. M., CHAN, J. C. K., & HIND, J. E. (1990). Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons. *Journal of the Acoustical Society of America*, 87, 757-781.

PRATT, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, 13, 278-285.

ROFFLER, S. K., & BUTLER, R. A. (1968). Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America*, 43, 1260-1266.

ROGERS, M. E., & BUTLER, R. A. (1992). The linkage between stimulus frequency and covert peak areas as it relates to monaural localization. *Perception & Psychophysics*, 52, 536-546.

WATKINS, A. J. (1978). Psychoacoustical aspects of synthesized vertical locale cues. *Journal of the Acoustical Society of America*, 63, 1152-1165.

NOTES

1. Our definition of mean CPA is identical to the area enclosed by a 1-dB iso-level contour used to describe the physical acoustic properties of the pinna (Carlile, 1990; Middlebrooks, Makous, & Green, 1989; Musicant, Chan, & Hind, 1990). Our reason for employing the term *covert* in *covert peak area* is that the peaks of the functions when SPL is plotted against frequency (spatial location as the parameter) are not always evident. They can be made apparent when SPL is plotted against spatial location with frequency serving as the parameter.

2. Data less representative of the group are plots in which the trough is not as deep as that illustrated in Figure 2.

3. This description of the notched stimulus was based on the output of a B&K half-inch microphone placed at the location normally occupied by the listener's head during testing.