Spectral cues used in the localization of sound sources on the median plane

Jack Hebrank and D. Wright

Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27706
(Received 10 June 1974; revised 16 September 1974)

The spectral cues used for median-plane localization are described by three experiments. First, the frequency spectrum necessary for localization is measured by noting the accuracy of subjects localizing low- and high-pass-filtered white noise. Second, several high-pass, low-pass, bandpass, and bandstop filters are associated with the subjective impression of direction by observing what directions are most frequently perceived by subjects localizing white noise colored by each filter. Third, the frequency responses of several artificial ears are measured for different angles of median-plane sound incidence. Results show that sound spectra from 4 to 16 kHz are necessary for localization. Frontal cues are a 1-octave notch, with a lower-frequency cutoff between 4 and 10 kHz and increased energy above 13 kHz. The "above" cue is a 1/4-octave peak between 7 and 9 kHz, with a high-frequency cutoff at 10 kHz. The "behind" cue is a small peak from 10 to 12 kHz. Increases in frontal elevation are signaled by an increase in the lower cutoff frequency of the 1-octave notch. This notch appears to be generated by time-delayed reflections off the posterior concha wall interfering with sound directly entering the external auditory canal.

Subject Classification: 65.62, 65.52.

INTRODUCTION

Auditory perception of elevation involves a complex interaction of several localization subsystems. For sounds located off the median plane (positions having nonzero azimuth angles), auditory localization cues comprise interaural time differences (ITD), interaural amplitude differences (IAD), and directional filtering by the external ears, as well as the changes in all these cues during head motion. Though the IAD and ITD mechanisms have been investigated and explained through azimuthal localization experiments, the generation and processing of external ear filtering cues are poorly understood.

An obvious situation for studying the role of external ear filtering in auditory localization is the localization of sounds on the median plane. Because sounds on the median plane (MP) are equidistant from both ears, IAD and ITD cues are invariant. Consequently, directional filtering by the external ears provides the sole localization cues. The necessity of external ear cues for MP localization is demonstrated by Gardner and Gardner's (1973) experiments that show that as the convolutions of the pinna are increasingly covered, MP localization ability of subjects drops to mere guessing.

We hypothesize that MP localization is accomplished by the decoding of spectral cues encoded by the direction-dependent filtering of the external ear. Further, we hypothesize that the spectral cues are decoded as combinations of simple features, i.e., low- and high-pass cutoffs, bandpass peaks and bandstop notches. These simple features are deeply filtered by the external ear, making them easily decodable for a variety of sounds.

This hypothesis is suggested by results from several other researchers. Probe-tube measurements by Shaw and Terishini (1968) show that the external ear (including the pinna and canal) performs complex spectral filtering of sounds above 5 kHz, with directional features spanning up to 20 dB. Furthermore, in additional experiments by Batteaù (1966), Condamines (1965), and Thurlow and Runge (1967), white noise is shown to be localized better than any other sound, implying that spectral cues which are optimally conveyed by the uniform spectral energy of white noise are used for MP localization. The notion that the cues are combinations of simple features is supported by the experiments of Blauert (1969–70) and Roffler and Butler (1968), which indicate that bandpass and low-pass filtering of sounds can induce the perception of MP locations independent of a sound source's actual location. Finally, Hebrank and Wright (1974) show subjects are less accurate monaurally than binaurally in MP localization of white noise, but have the same relative loss in accuracy when localizing ripple spectrum noise, suggesting MP localization does not involve processing differential cues generated by asymmetric pinna filtering.

Three experiments are reported in this paper which determine the features encoded by the external ear that are used in median-plane localization. The first experiment defines the frequency band containing MP localization cues by measuring subject's accuracy in localizing a white noise signal as its bandwidth is decreased by either sharp low-pass or sharp high-pass filtering. In the second experiment, subjects localize white noise deeply filtered to have low-pass, high-pass, bandpass, or bandstop characteristics in the localization band determined in the first experiment. A subjective external ear filtering pattern for each direction is created by noting the filters which bias the subjects' responses to certain directions. The third experiment verifies these subjective filtering patterns by measuring the actual external ear filtering for different directions. This is accomplished using three artificial ears, with a microphone mounted at the position of the ear drum to obtain the direction-dependent frequency response of the external ear. Together, these three experiments explain the role of the external ear in median plane localization and in the auditory perception of elevation.

I. EXPERIMENT I

A. Method

The spectral band used in MP localization was identified by measuring the accuracy of ten subjects localizing white noise in which the bandwidth was changed by sharp low- or high-pass filtering. The high- and low-pass limited white noise was generated by filtering white noise from an Elgenco 602A (±1 dB, 20 Hz to 20 kHz) with a seven-pole Cauer–Chebychev filter (±2-dB maximum passband ripple, 50-dB minimum stopband attenuation, and 24% transition band) capable of seven low-pass cutoff frequencies (3.9, 6.0, 8.0, 10.3, 12.0, 14.5, and 16.0 kHz) and six high-pass cutoff frequencies (3.8, 5.8, 7.5, 10.0, 13.2, and 15.3 kHz). The band-limited noise was fed to a speaker (Electrovoice T-35, ±5 dB, 2.5–20 kHz) which could be moved silently in a 0.9-m radius around the subject's clamped head in a darkened anechoic chamber. During the experiment, the speaker moved among nine positions (30° apart) from −30° to +210° (0° being frontal horizon), emitting the filtered noise for 1 sec at 8-sec intervals. Subjects indicated their perceived location of the sound by pressing one of nine buttons on a lighted hand-held box. The white noise was electronically sequenced and gated (5-msec rise and fall times). Each filter condition was presented at 18 positions (twice at each of the nine positions in random order) so that a total of 252 trials was presented to each subject (seven low-pass cutoffs, six high-pass cutoffs, and unfiltered white noise). Sound pressure level at the subject's ear was 50 dBA for the white noise, with all system gains remaining the same for each filter condition.

B. Results

Figure 1 shows percent correct localization responses as a function of the cutoff frequencies for both high- and low-pass-filtered noise. Results, plotted for coarse (±45°) as well as fine (±15°) performance, show that the absence of spectral energy above 16 kHz or below 3.8 kHz does not affect localization performance, implying that the features used in MP localization are between 3.8 and 16 kHz.

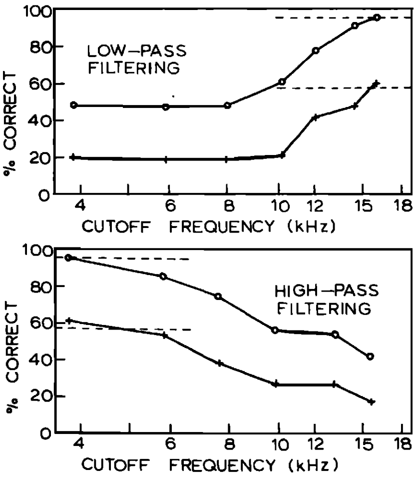

FIG. 1. Mean percent correct responses for ten subjects localizing low- and high-pass-filtered white noise with accuracies of ±45° (0) and ±15° (+). Dotted lines show mean accuracies for 180 trials of unfiltered noise. Graphs based on 180 trials for each filter cutoff frequency.

These results may appear in conflict with Roffler and Butler's (1968) conclusion that, for MP localization, the stimulus "must include frequencies above 7.0 kHz." However, our results are for localization judgments over a 240° arc, whereas Roffler and Butler only considered a 33° arc in the frontal quadrant. In the discussion, their results are interpreted in light of the features we find responsible for frontal elevation perception.

II. EXPERIMENT II

A. Method

Subjective external ear filtering for different directions was determined in Experiment II by associating simple filters of different cutoff or center frequencies with the perception of specific median-plane directions. For instance, if the cue to a sound located above (+90) were a peak near 8 kHz, then a subject trying to localize bandpass-filtered white noise, that contains energy only at 8 kHz, would interpret the prefiltered sound to be overhead, in spite of its actual location. By having subjects localize each type of filtered noise several times from different locations and noting which filters bias the subjects' responses toward specific directions, the individual features of external ear filtering for each direction were established. Of course, the data from this experiment were muddied by both differences among subjects' ear geometry and by conflicts between the electronically filtered cues and the external ear cues in the passband of the prefiltered noise. Nevertheless, it should be possible to predict the external ear filtering for each direction by combining the features moderately or strongly associated with each direction. The features localized by subjects included high-pass cutoffs at six different frequencies, low-pass cutoffs at seven frequencies, 1/12-octave bandpass filters at 12 frequencies, and bandstop filters at 12 frequencies.

Three different groups of subjects were tested in this experiment. The first group of ten subjects localized white noise which was filtered to have one of six high-pass cutoffs (3.8, 5.8, 7.5, 10.0, 13.2, and 15.3 kHz), one of seven low-pass cutoffs (3.9, 6.0, 8.0, 10.3, 12.0, 14.5, and 16.0 kHz), or unfiltered white noise located among the nine positions between −30° and 210° on the median plane. Each filter type was localized 18 times by each subject. The second group of ten subjects localized white noise which was filtered to have one of 12 1/12-octave bandpass peaks (4.0, 4.5, 5.1, 5.7, 6.4, 7.2, 8.1, 9.1, 10.2, 11.4, 12.8, 14.5 kHz). Each bandpass filter was presented nine times, once at each position, to each subject. The third group of eight subjects localized bandstop notches at 12 frequencies (6.2, 6.6, 7.0, 7.4, 7.9, 8.8, 9.8, 10.8, 12.0, 13.3, 15.0, and 17.8 kHz) generated by summing white noise with a time delay (28–81 μsec) of itself.

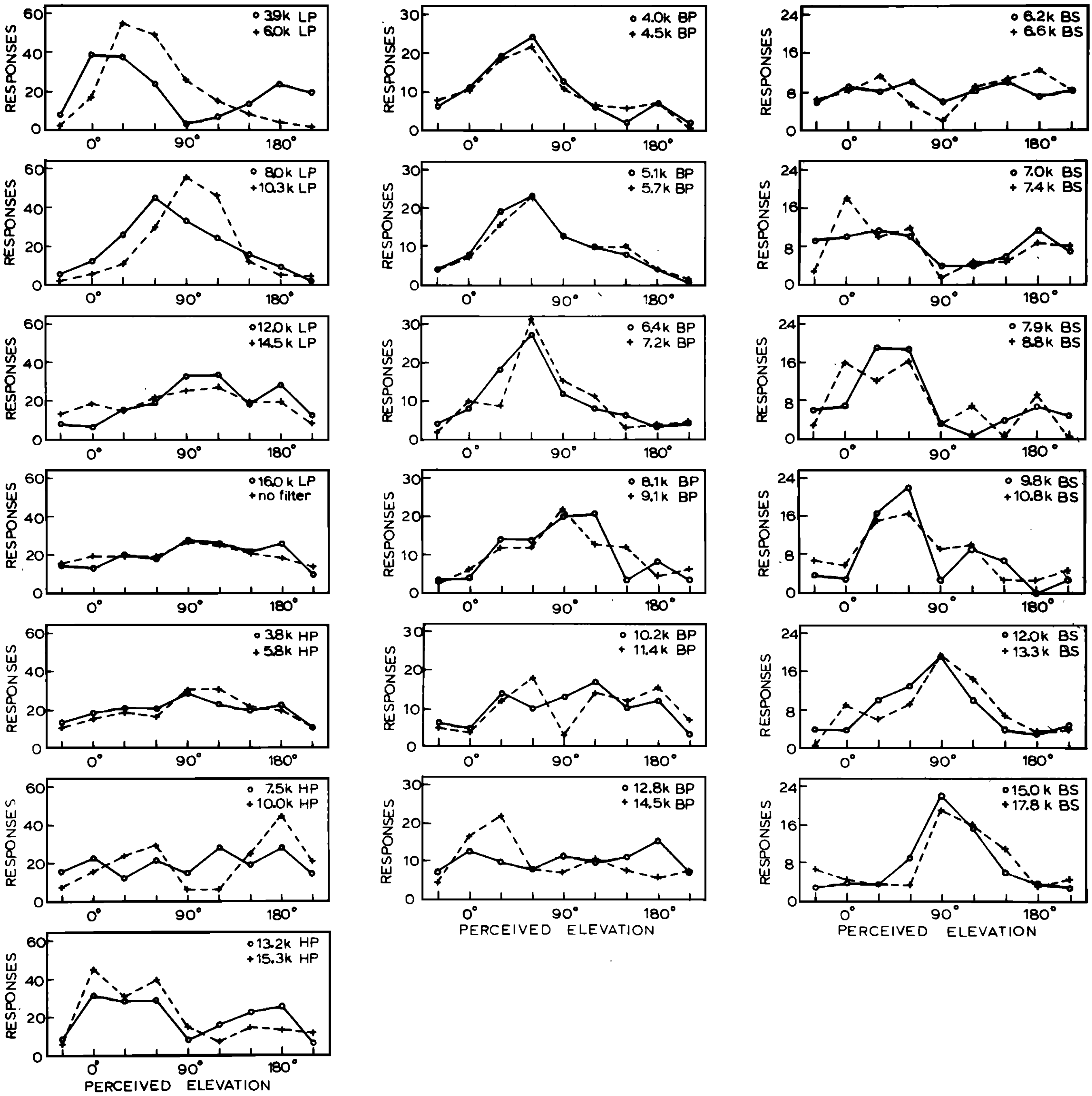

FIG. 2. Responses of subjects localizing white noise filtered to contain features similar to external ear filtering. Four types of filters used are low-pass (LP) and high-pass (HP) in (a), bandpass (BP) in (b), and bandstop (BS) in (c). Distribution of localization responses for unfiltered white noise is shown in (a).

In all three groups of this experiment, an equal number of stimuli was presented at each position. However, both sequencing of filters and speaker positions were "random," in that the ordering was by lottery. Other experimental conditions, such as sound level and gating, clamped head, and darkened anechoic chamber, were the same as described in Experiment I.

B. Results

The results of Experiment II [shown in Figs. 2(a), 2(b), and 2(c)] indicate that some of the filters clearly cause perception of particular directions. In the seven graphs of Fig. 2, the number of responses as a function of perceived direction is plotted for the different high- and low-pass filters. Also plotted are the responses to unfiltered noise. The strongest tendencies seen in Fig. 2(a) are: (1) 3.9–8.0-kHz low-pass cutoffs are associated with frontal elevation (increasing cutoff frequency is related to the perception of increasing elevation); (2) the 10.3-kHz low-pass-filtered noise is perceived above; (3) 10.0-kHz high-pass cutoffs appear to be behind; and (4) the 13.2- and 15.3-kHz high-pass cutoffs are perceived in front.

In Fig. 2(b), localization responses to the 12 bandpass-filtered signals are plotted. Again, some of the filters induce strongly localized impressions. Specifically, the localization impressions caused by these filters are: (1) 4.0–7.2-kHz bandpass filtering induces elevated "frontness"; (2) 8.1–9.1-kHz bandpass filtering induces "aboveness"; and (3) 14.5-kHz bandpass noise is associated with "frontness." To a lesser extent, 10.2–12.8-kHz bandpass noise is perceived behind.

In Fig. 2(c), responses to notch filtering are plotted. Here, 7.4–10.8-kHz notches induce "frontness" and 12.0–17.8-kHz notches are perceived as above. In both Figs. 2(a) and 2(c), increases in the cutoff frequencies of the low-pass filters and increases in the center frequencies of the notch filters cause the perception of increased elevation.

By graphing all those filters causing the perception of a particular direction, the ear's spectral filtering for a direction can be predicted. In Fig. 3, the patterns for 0°, 90°, and 180° are graphed by drawing the frequency responses of filters associated with the perception of each direction. Filters associated with each direction are arbitrarily defined as those filters causing responses of that direction to be 1.5 times the responses of that direction for unfiltered noise. In Fig. 3, the interrelationships between low- or high-pass cutoffs and peaks or notches are clear. It appears that a peak or notch can convey the sensation of a low- or high-pass

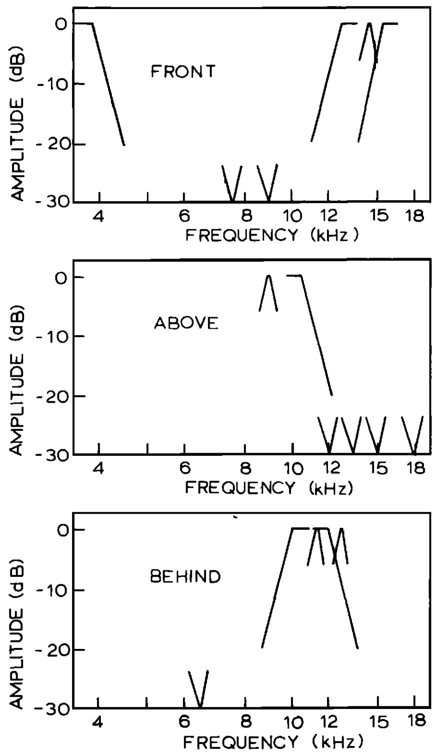

FIG. 3. Frequency responses of all filters which generate colored noise that is localized at 0° (a), 90° (b), or 180° (c). For clarity, only the filter response about its center or cutoff frequency is plotted. The overall pattern of external ear filtering used for localizing each of these three directions can be predicted from the combined frequency responses of the simple filters.

cutoff at frequencies just above or below its center frequency; similarly, high- or low-pass filtering can give the impression of a peak at the edge of its passband or a valley in its stopband. This phenomenon is analogous to Mach bands in the visual system. The good fit of peaks, notches, and cutoffs into patterns lends credence to the predicted subjective filtering.

The major features involved in MP localization can be predicted from Figs. 2 and 3. Frontal sounds are cued by the low-pass cutoff between 4 and 8 kHz of a notch and increased energy above 13 kHz. Overhead sounds are cued by a peak near 8–9 kHz and its accompanying 10-kHz low-pass slope. Posterior sounds appear to be cued by a 10-kHz high-pass slope leading to a peak near 12 kHz. From Fig. 2, it is apparent that increasing frontal elevation is cued by the increasing low-frequency cutoff of a notch of increasing frequency.

III. EXPERIMENT III

A. Method

In Experiment III, actual external ear directional filtering was measured in order to verify the localization cues determined in Experiment II. The filtering was measured by inserting a microphone (Brüel and Kjaer 4134, 1/2-in. condenser microphone) at the position of the eardrum in several artificial external ears. The "ears" were made by making molds of three subjects' ears with dental "geletin" and casting epoxy in these molds. Artificial canals, 1-in. long and 3/16-in. in diameter, were attached to the epoxy castings and terminated by the microphone element. Exact matching of the model's eardrum impedance was not attempted, since Shaw and Teranishi's (1968) measurements show that the "height" and "position" (frequency) of directional features above 5 kHz are relatively independent of eardrum impedance. Frequency response measurements between 4 and 16 kHz were made with each of the three ears for the nine directions used in Experiments I and II.

B. Results

Figure 4 shows the nine frequency responses of each of the three ears. Examination of these frequency responses, in light of the features found in Experiment II, indicates that all three ears perform very similar filtering. The most obvious features are (1) the increasing low-frequency cutoff between 4 and 10 kHz of a 1-octave notch, corresponding to increasing frontal elevation, (2) a 7–9-kHz 1/4-octave peak and its 10-kHz cutoff slope, corresponding to overhead sounds, and (3) the difference in energy above 13 kHz for anterior and posterior directions. The peak at 12-kHz signaling posterior sounds is present in all three ears, but less obvious.

IV. DISCUSSION

The localization cues we have found are substantiated by two other reported experiments. Blauert (1969–70) shows that subjects trying to localize 1/3-octave noise

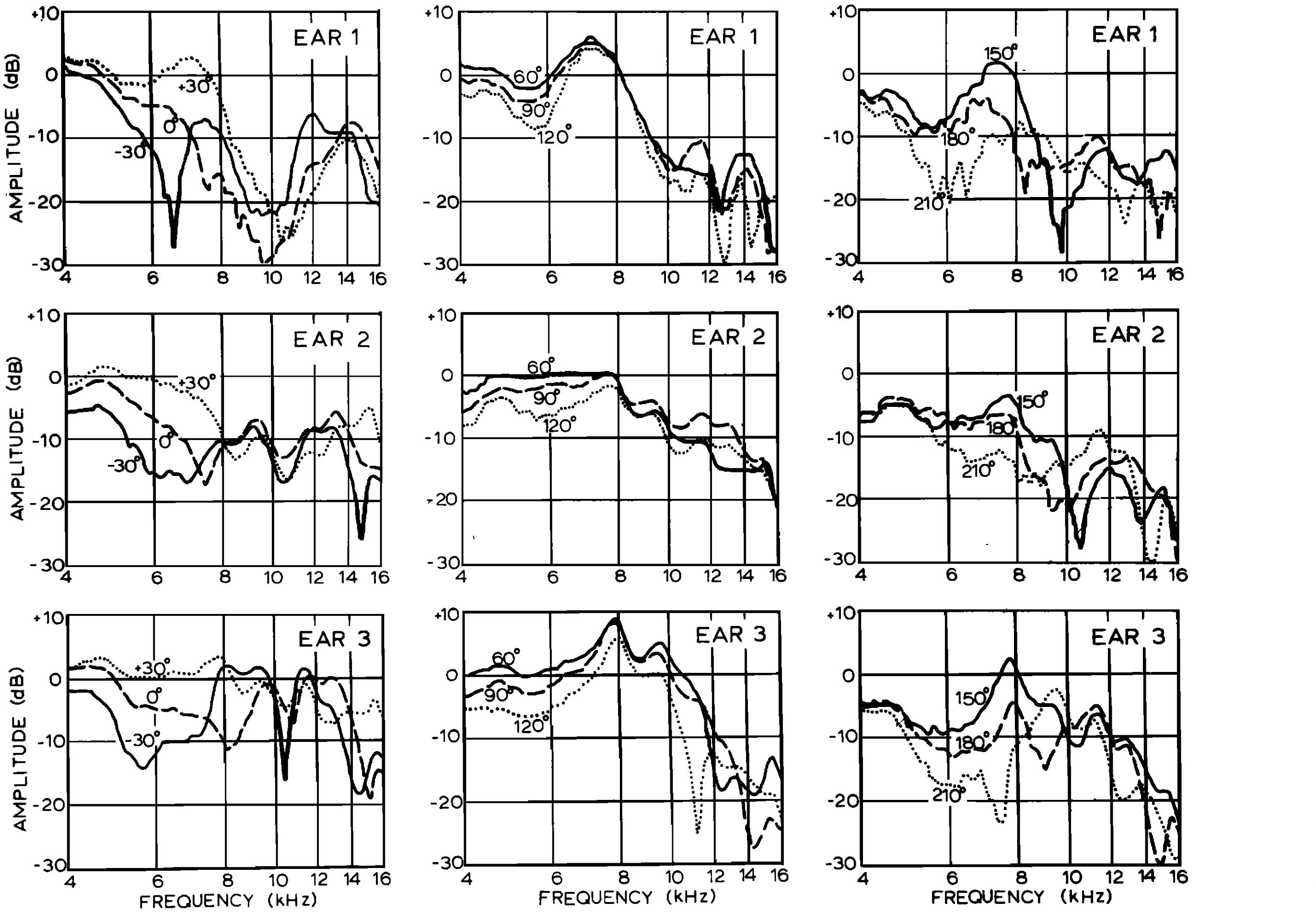

FIG. 4. Frequency response at the eardrum of three artificial ears for nine angles of median plane incidence (0° is frontal horizon). Note (1) the correspondence of increasing frontal elevation and the increasing lower cutoff frequency of a 1-octave notch, (2) the peak near 7–9 kHz for overhead directions, and (3) the difference in attenuation above 13 kHz for front and behind directions.

—as front, above, or behind—make localization judgments independent of the source's position but highly dependent upon the center frequency of the 1/3-octave noise. Blauert observed the following: third-octave noise between 2.5 and 6.3 kHz is localized in front (these 1/3-octave signals have low-pass cutoffs in the 4–8 kHz band); 8-kHz 1/3-octave noise is localized above (this noise band causes a peak at 8 kHz); noise bands between 10 and 12.5 kHz are localized behind (corresponding to the inducement of a peak around 10 to 12 kHz or a high-pass slope at 10 kHz); and 16-kHz 1/3-octave noise is localized in front (again corresponding to our observation that increased energy above 13 kHz is a frontal cue.) Therefore, the cues reported by Blauert are consistent with our hypothesis. Roffler and Butler's (1968) experiments can also be related to our work. They observed that subjects localizing low-pass-filtered noise in the frontal quadrant (−13° to 20°) respond accurately for cutoff frequencies equal to or greater than 8 kHz, but make errors for cutoff frequencies below 8 kHz. Our finding that 4–8-kHz low-pass cutoffs convey frontal elevation cues suggest that low-pass filtering below 8 kHz either eliminates the external ear cues used for frontal elevation or creates false cues. For low-pass filtering above 8 kHz, the external ear elevation cues are conveyed intact, so that subjects can localize accurately.

In addition to explaining the results of other median-plane localization investigators, our findings shed light on two long-standing questions of sensory psychology.

(1) How does the odd but characteristic shape of the concha relate to its function?

(2) What physiological factor causes the spatial terms "high" and "low" to be used as descriptions of auditory qualities?

We hypothesize that the concha's shape results in part from the necessity for elevation cues to be encoded by the external ear. The physical acoustics of the pinna canal are, admittedly, a complex combination of diffractions, resonances, and reflections. However, a strong case can be made for the hypothesis that frontal elevation cues are generated primarily by the reflection of a sound off the posterior concha wall interfering with the same sound as it directly enters the external auditory canal (see Fig. 5). Attenuation greater than 6 dB occurs at frequencies where the time delay between direct and reflected sounds causes phase shifts from 120° to 240°. For higher elevations, the reflected path length is shorter, so that the notch

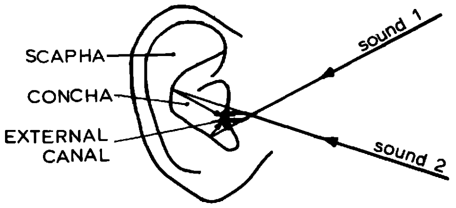

FIG. 5. Schematic diagram of the external ear. The posterior wall of the concha reflects frontal sounds, so that both direct and time-delayed sound from source enter the external auditory canal. Because of the concha's geometry, sounds at higher elevations have shorter reflected component path lengths, hence, less time delay, so that the perceived sound's spectrum will have a notch at a higher frequency.

in a sound's spectrum is at an increased frequency. To confirm the reflection hypothesis, the actual path lengths of concha reflections can be compared with the frequencies of notches in the external ear response. For example, the distance d from the external auditory canal to the posterior concha wall, at a MP sound incidence ϕ, is related to the time delay τ of the reflected sound by

$$\tau(\phi) = 2d(\phi)/\text{velocity of sound},$$

so the predicted notch in a sound's spectrum should be centered at frequency f:

$$f(\phi) = 1/2\tau(\phi) = \text{velocity of sound}/4d(\phi).$$

In ear model 1 of Experiment III, the distance from the canal to concha wall is roughly 1/2 in. for $-30°$ MP sound incidence, 3/8 in. for $0°$ incidence, and 1/4 in. for $+30°$ incidence. These path lengths would cause delays of 76, 57, and 38 μsec, predicting notch center frequencies of 6.6, 8.8, and 13.2 kHz, respectively. The center frequencies of ear 1, as seen in Fig. 4, are 6.5, 9.5, and 10.5 kHz for $-30°$, $0°$, and $+30°$ elevations, respectively. The agreement of the notches predicted from the physical dimensions and the notches in the frequency response data supports the hypothesis that concha geometry is responsible for the production of elevation cues by the variable path-length reflections that occur of the posterior wall.

We also speculate that the association between auditory and spatial "highness" results from the frontal elevation cue. The application of the spatial terms "high" and "low" to a tone's pitch was shown to have a phenomenological basis by Pratt (1930), who demonstrated that subjects localize higher-frequency pure tones above lower-frequency pure tones. The frontal elevation cue, a 1-octave notch, may effectively function as a low-pass cutoff, so that lower elevation sounds are filtered to have less bandwidth, hence, fewer high-frequency components. The hint that the frontal elevation cue, a 1-octave notch, may act as a low-pass cutoff comes from Experiment II, in which both low-pass and bandstop filtering are shown to induce perception of elevated sounds in front of the subject. Possibly, the 20-dB/octave rolloff in the auditory system's sensitivity above 5 kHz (see Hirsh, 1952), combined with the octave-wide notch, creates a functional low-pass filter. Although the arguments presented here, implying that notch elevation cues are effectively low-pass cutoffs, are tenuous, the increase in frequency of the elevation cue with increasing elevation is probably the source of the link between spatial and auditory height.

V. SUMMARY

We have shown that median-plane localization of white noise is based on simple but deep spectral cues generated by the directional filtering of the external ear. These cues are between 4 and 16 kHz. The three general directions of front, above, and behind are represented by three patterns.

(1) Front: a 1-octave notch having a lower cutoff frequency between 4 and 8 kHz and increased energy above 13 kHz.

(2) Above: a 1/4-octave peak between 7 and 9 kHz.

(3) Behind: a small peak between 10 and 12 kHz with a decrease of energy above and below the peak.

Elevation in the frontal quadrant is cued by the lower cutoff frequency of a 1-octave notch centered from 5 to 11 kHz. The notch, whose center-frequency and low-frequency cutoff increase as elevation increases, is caused by summation of the direct and concha-reflected components of a sound entering the external auditory canal. We speculate that the relationship between the notch frequency and elevation is responsible for the classic association of spatial and auditory "highness."

Batteau, D. W. (1961). "Characteristics of Human Localization of Sound," Final Rep. May 1961, prepared under contract No. N123-(60530) 23545A for U. S. Naval Ordinance Test Station, China Lake, Calif.

Blauert, J. (1969/70). "Sound Localization in the Median Plane," Acustica 22, 205–213.

Condamines (1965). "Le Reperage Directionnel Des Sons En Ecoute Mono et Biauriculaire," 5eme Cong. Int. D'Acoust. Liege, Belgium. Paper B63.

Gardner, M. B., and Gardner, R. S. (1973). "Problem of Localization in the Median Plane: Effect of Pinna Cavity Occlusion, J. Acoust. Soc. Am. 53, 400–408.

Hebrank, J. H., and Wright, D. (1974). "Are Two Ears Necessary for Localization of Sounds on the Median Plane?" J. Acoust. Soc. Am. 56, 935–938.

Hirsh, I. J., *The Measurement of Hearing* (McGraw-Hill, New York, 1952), pp. 105–110.

Pratt, C. C. (1930). "The Spatial Character of High and Low Tones," J. Exp. Psychol. 13, 278–285.

Roffler, S. K., and Butler, R. A. (1968). "Factors that Influence the Localization of Sound in the Vertical Plane," J. Acoust. Soc. Am. 43, 1255–1259.

Shaw, E. A. G., and Teranishi, R. (1968). "Sound Pressure Generated in an External Ear Replica and Real Human Ears by a Nearby Point Source," J. Acoust. Soc. Am. 44, 240–249.

Thurlow, W. R., and Runge, P. S. (1967). "Effect of Induced Head Movement on the Localization of the Direction of Sounds," J. Acoust. Soc. Am. 42, 480–488.