

# The contribution of head motion cues to localization of low-pass noise

STEPHEN PERRETT and WILLIAM NOBLE

University of New England, Armidale, New South Wales, Australia

Localization of low-pass sounds was tested in relation to aspects of Wallach's (1939, 1940) hypotheses about the role of head movement in front/back and elevation discrimination. With a 3-sec signal, free movement of the head offered only small advantage over a single rotation through  $45^\circ$  for detecting elevation differences. Very slight rotation, as observed using a 0.5-sec signal, seemed sufficient to prevent front/back confusion. Cluster analysis showed that, in detecting elevation, some listeners benefited from rotation, some benefited from natural movement, and some from both. Evidence was found indicating that a moving auditory system generates information for the whereabouts of sounds, even when the movement does not result in the listener facing the source. Results offer significant if partial support for Wallach's hypotheses.

Classical theory of auditory localization suggests that, for motionless listening, interaural cues, at least for simple sounds, provide spatial information that is essentially ambiguous. Interaural cues take the form of time differences for low-frequency acoustic energy and level differences for high-frequency energy. A classical model assumes the head to be a perfect sphere, with holes, centered on opposite sides, for ears. In this model, interaural cues specify the angle of horizontal displacement of a sound source from the median vertical plane (MVP), with the vertex of the angle at the center of the listener's head. From here on, this will be referred to as the *azimuth angle*. Its properties are consistent with the azimuth angle defined in the double-pole coordinate system used by Middlebrooks, Makous, and Green (1989). Figure 1 illustrates this system and shows that a particular azimuth angle encompasses a range of directions, including the direction of the source. An azimuth angle of  $60^\circ$  to the left, for example, specifies a cone-shaped locus covering positions forward, rearward, above, and below the interaural axis—a *cone of confusion* (Woodworth & Schlosberg, 1954). It can thus be seen that, considered purely in terms of geometry, interaural cues alone do not specify the elevation of the source, nor whether it is forward or rearward.

Other localization cues can be obtained from the shape of the sound spectrum. Direction-dependent spectral changes furnished by the pinnae allow fairly precise front/back-up/down localization. However, for the auditory system to make use of pinnae cues, the sound must comprise energy spanning a range of frequencies within the 4- to 12-kHz region (Hebrank & Wright, 1974). When acoustic energy above 4 kHz is absent, as in noise that is low-pass filtered at 2 kHz, localization errors occur of a sort expected from cone-of-confusion principles (Perrett & Noble, 1995).

Interaural cues become far more complex in the context of head movements. Wallach (1939, 1940) proposed that head movements could be used to pinpoint the location of a sound anywhere in auditory space. For example, if a sound is on the horizontal plane (HP) at  $60^\circ$  to the left of the straight-ahead position, a leftward rotation of the head about a vertical axis initially produces a decrease in the azimuth angle specified by interaural cues. In contrast, if a sound is produced in the rear/left quadrant of the HP at  $60^\circ$ , the same movement would cause an initial increase in azimuth angle. The different patterns of change in interaural cues brought about through head rotation might enable front sources to be distinguished from back sources.

Wallach (1939, 1940) further argued that within a hemisphere above or below the HP, the rate of change in azimuth angle relative to the change in head orientation could be used to specify source elevation. When a sound source is located on the HP, a given head rotation brings about the same change in azimuth angle; for example,  $15^\circ$  of head rotation will bring about a  $15^\circ$  change in this angle. On the other hand, rotation produces no change in azimuth angle when a sound is directly overhead. That angle remains  $0^\circ$  for any amount of head rotation. For intermediate elevations, the change in azimuth angle is somewhat greater than zero but somewhat less than the degree of head rotation. Wallach proposed that the dif-

---

We wish to thank Malcolm A. Perrett, who helped provide mathematical solutions for processing of the head tracker data and whose development of the DSP software and digital signals used in the experiment formed part of his Honours (Electrical Engineering) project. Appreciation is extended to Frank Niebling for construction of the supporting framework for the loudspeakers and the spherical screen, and to Dean Davidson for computer programming. We also thank the reviewers of earlier drafts of the paper, whose comments led to significant developments in our thinking and practice. The research reported here was supported in part by grants from the Australian Research Council and the UNE Arts Faculty Research Fund. Correspondence should be addressed to W. Noble, University of New England, Department of Psychology, Armidale, NSW 2351, Australia (e-mail: wnoble@metz.une.edu.au).

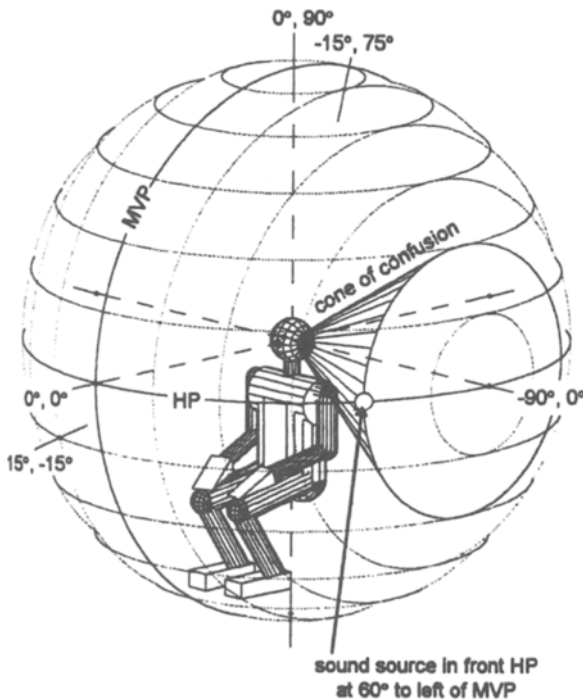


Figure 1. Schematic of spatial references relative to a listener, showing a cone of confusion at a leftward azimuth angle of  $60^\circ$ . Increments by  $15^\circ$  in azimuth angle are represented on a sphere as circles perpendicular to the horizontal plane (HP). Azimuth angles are positive when rightward, negative when leftward. Increments by  $15^\circ$  in elevation angle are shown as circles perpendicular to the median vertical plane (MVP). These are positive when upward, negative when downward. Examples are shown of coordinates using this double-pole system.

ference between the change in azimuth angle and the change in head orientation could allow elevation to be distinguished. Furthermore, the ambiguity between above and below could be eliminated by then tilting the head from side to side.

Wallach (1939, 1940) successfully simulated different source elevations using loudspeakers arrayed in the HP, along with a rotary switching device attached to the listener's head. The switching device caused the loudspeakers to be activated sequentially with rotation of the listener's head. By varying the separation between loudspeakers, the same head rotation could be used to generate different apparent changes in azimuth angle, thus simulating a variety of elevations. Listeners did report apparent sources at elevations broadly consistent with the rate of change in azimuth angle relative to the change in head orientation. This result suggests that the changes in interaural information produced by head rotation may act as an elevation as well as a front/back cue.

Subsequent experiments have provided only limited support for Wallach's (1939, 1940) theory. Pollack and Rose (1967) tested localization of broad-band signals in the front HP and concluded that head movement contributed little except when signal duration was at least 3 sec,

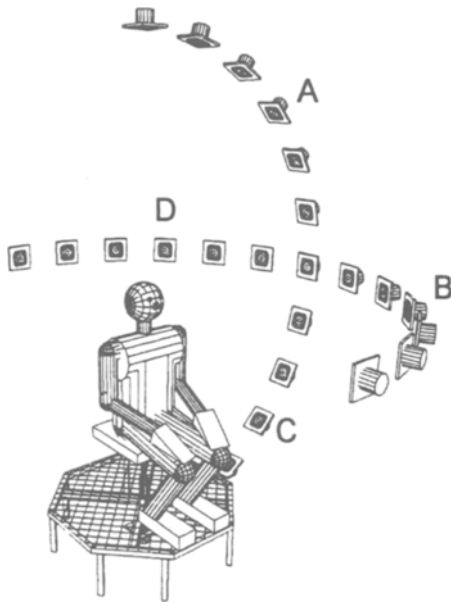
and sources were at lateral positions. Thurlow and Runge (1967) tested various mechanically induced head movements and found that movement produced only very slight increase in the accuracy of localization of 5-sec low-frequency noise bursts. Thurlow and Mergener (1970), using the same loudspeaker positions, concluded that free head movement assists localization of low-frequency noise when the signal is at least 1 sec, while performance approaches an optimal (although not very accurate) level with 2-sec signals.

These studies have certain limitations. Pollack and Rose (1967) restricted testing to visible sources in the front HP; therefore, front/back-up/down localization was untested. Thurlow and colleagues (Thurlow & Mergener, 1970; Thurlow & Runge, 1967) did not test localization of sources elevated more than  $41^\circ$  from the HP. Furthermore, experimental conditions have often featured movement that enabled listeners to zero in on sources, leaving the actual function of a moving auditory system in relation to an ongoing signal unexamined. These limitations mean that the contribution of head motion to sound localization remains unclear. As noted by Middlebrooks and Green (1991, p. 153), "In light of all the evidence, a defensible argument is that unless the sound duration is sufficient to allow the listener to turn to face the source, thereby obtaining the optimum static localization cues, moving one's head may indeed be a poor strategy for improving the accuracy of localizing short-duration sources." The present report covers an experiment designed to investigate the role of head motion and to offer a test of aspects of Wallach's (1939, 1940) theory, using signals of short as well as longer duration, and restricted as well as free movement regimes.

## METHOD

The experiment involved 12 people (5 females, 7 males), all of similar background and reporting normal hearing. They were presented with low-pass noise bursts from 25 loudspeakers (Realistic midrange/tweeters) at  $15^\circ$  intervals, forming two intersecting arcs, one spanning the left HP from directly in front to directly behind, the other spanning the leftmost lateral vertical plane (LVP) from directly above to directly below. The loudspeakers were mounted on curved frames, constructed of 2-cm box-section steel tube, at a distance of 1.25 m from, and facing the center of, the listener's head (Figure 2). The apparatus was housed in a semi-anechoic, sound-isolated room.

Perrett and Noble (1995) showed that altering the range of feasible response choices profoundly affects localization decisions, so that constraints on response choice can give a biased picture of localization proficiency. To achieve as low as possible a constraint on responses, a 1.2-m radius spherical screen was constructed, using acoustically transparent fine-weave fabric, and suspended in a framework of hoops made of 2-cm diameter PVC tubing, thus visually masking the locations of the loudspeakers. The PVC frame for the screen was designed so as not to obstruct any of the loudspeakers, which were arrayed just beyond it. A hinged section of the screen could be swung open to allow access. Inside, a rotatable seat, height adjustable, was mounted on a platform made of 2-cm box-section steel tubing and weldmesh, which allowed the portion of the screen lying underneath to be seen. The legs of the platform pierced the screen to connect with the laboratory floor. Additional



**Figure 2.** Apparatus used, showing listener position and loudspeaker positions at a radius of 1.25 m. Region A is the upper lateral vertical plane (LVP), Region C the lower LVP; Region B is the front horizontal plane (HP), Region D the back HP. Not represented is the 1.2-m-radius spherical mask positioned around the listener and in front of the arrays of loudspeakers (Figure 1 allows it to be visualized).

screening outside the sphere prevented participants from gaining knowledge of actual loudspeaker placement. As far as they were aware, potential sound sources could lie in virtually any direction.

Polhemus Isotrak II head tracking equipment was used to register head motion. The head tracker transmitter was positioned 480 mm to the right of the center of the listener's head, mounted on the end of a length of 30-mm-diameter PVC tubing inserted through the screen. An unobtrusive adjustable head harness held the head tracker receiver and a laser pointer; the pointer was to assist listeners in making response decisions. TV monitoring was also used; the camera lens accommodated by making a small aperture in the screen.

The signal was created digitally, using the Matlab RANDN function, to produce 3.08- and 0.58-sec samples of white noise (power density spectrum constant at all frequencies) with sampling rate of 44.1 kHz. Each sample was digitally filtered with a 255th-order finite impulse response (FIR) filter to convert it to pink noise (equal energy per octave), then a 4,096th-order FIR low-pass filter to produce a 2.0-kHz cutoff with approximately 90 dB/octave rejection slope. Using custom-written digital signal processing (DSP) software, each signal was inverse filtered (4,096th-order FIR) to minimize effects of individual loudspeaker transfer functions. The start and finish of each sample was truncated by 0.04 sec to eliminate transients produced by the filtering process. To reduce onset/offset transients occurring at playback, the DSP software also applied a cosine squared windowing function with 20-msec onset/offset ramps. Thus separate 3- and 0.5-sec samples were produced for each loudspeaker and stored as Microsoft Windows WAVE-format sound files on the hard disk drive of an 80386 DX/40 personal computer. The computer was fitted with a Sound Blaster 16 Vibra audio card, which was used to convert the sound files into analog signals.

Custom-written software running on a second (80286) computer, linked to the 80386, was used to control sound delivery during the experiment. The audiocard analog output was passed through a cus-

tom-built manually adjustable attenuator, used to set the mean overall signal level. Signals were then routed through a custom-built computer-controlled attenuator, which provided for random-level variations and also compensated for differences in individual loudspeaker efficiency. The signal was then amplified by a custom-built power amplifier and switched to 1 of the 25 loudspeakers via an array of self-cleaning relays; switching occurred approximately 100 msec before and after each sound delivery, thus preventing audible switching transients. The mean signal level was 55 dBA; the actual level was randomly adjusted from trial to trial in multiples of .375 dB over a range of  $\pm 3$  dB to prevent identification of loudspeaker positions based on minor overall loudness differences.

Listeners were required to localize the sound sources in each of three different conditions: (1) reacting as they normally would when searching for the source of a sound; (2) rotating the head 45° leftward on the HP, after the onset of the signal, and thereafter remaining motionless; and (3) remaining motionless throughout. Participation in each of the three movement/no-movement conditions was in two separate sessions, one with the 0.5-sec signal and the other with the 3-sec one. Movement and signal duration conditions were counterbalanced across participants.

Before each trial, listeners sat facing ahead, aiming the laser pointer at a spot of light at 0° azimuth in the HP. When ready, they pressed a button on a hand-held module that initiated head tracker data collection and, after a random delay of between 1 and 2 sec, a noise burst was produced from one of the loudspeakers. In response to the onset of the noise burst, listeners moved or remained still, as the condition required. In the rotation condition, the head was rotated leftward until the pointer was aligned with a light spot at -45° on the HP. At signal offset, in all conditions, listeners were permitted to move as they wished so as to aim the laser pointer where they judged the sound to have come from. Once satisfied that the pointer was aiming in the appropriate direction, the listener made a further button-push, which stopped head tracking data collection. The coordinates recorded by the head tracker at that moment identified the judged direction of the source. For each listening condition, there were 7 practice trials followed by 50 experimental trials (2 per loudspeaker position). The order of loudspeaker activation was random. No feedback on accuracy was given. From inspection of head tracker recordings and from TV monitoring at the time of the experiment, it was evident that the different conditions were successfully complied with on almost every trial.

## RESULTS

Localization performance as a function of movement and signal duration is considered first, and in relation to Wallach's (1939, 1940) claims, through analyses of four features of the data: (1) absolute accuracy, (2) front/back errors, (3) absolute elevation error, and (4) apparent elevation. Absolute accuracy is expressed as the source-head-response (SHR) angle—the angle describing the relation between the position of the source, the center of the listener's head, and the position of the response, on any trial. Front/back errors are counted as responses occurring in the hemisphere behind or in front of the interaural axis to sounds located in the opposite hemisphere (thus responses to sources on the LVP were not included in this analysis). Absolute elevation error is the vertical component of the SHR angle and is used in statistical analysis of elevation judgments since it avoids the cancellation effect from adding error magnitudes of opposite sign (responses above and below actual positions) that occurs with signed elevation error. Apparent eleva-

**Table 1**  
Average SHR Angle, Absolute Elevation Error, and Standard Deviations for Each Condition

| Signal (Seconds) | Listening Condition | SHR Angle |           | Elevation Error |           |
|------------------|---------------------|-----------|-----------|-----------------|-----------|
|                  |                     | <i>M</i>  | <i>SD</i> | <i>M</i>        | <i>SD</i> |
| 3.0              | natural             | 22°       | 6.6†      | 21°             | 5.5†      |
| 3.0              | rotation            | 26°       | 3.1†      | 23°             | 3.1*      |
| 3.0              | motionless          | 42°       | 7.1       | 28°             | 4.4       |
| 0.5              | natural             | 37°       | 8.3*      | 26°             | 4.6       |
| 0.5              | rotation            | 30°       | 4.1       | 25°             | 4.4       |
| 0.5              | motionless          | 41°       | 8.1†      | 28°             | 5.4       |

Note—SHR, source-head-response. Tukey's HSD test: \* $p < .05$ . † $p < .01$ .

tion is simply the elevation of a listener's pointing response and is relied on for graphically representing listeners' elevation judgments. A subsequent analysis of the horizontal component of SHR angle was made to scrutinize accuracy for sources in different spatial regions as a function of movement condition.

### SHR Angle

Average SHR angles in the six experimental conditions are shown in Table 1. A one-way repeated measures analysis of variance (ANOVA) revealed a significant difference across conditions [ $F(5,55) = 32.47, p < .001$ ]. Generally, localization was more accurate when head motion occurred. Post hoc testing (Tukey's HSD multiple comparisons) confirmed that the 3-sec natural and rotation conditions produced significantly smaller SHR angles than did the 3-sec motionless condition ( $p < .01$ ) and that the 0.5-sec rotation condition produced significantly smaller SHR angles than either the 0.5-sec motionless condition ( $p < .01$ ) or the 0.5-sec natural condition ( $p < .05$ ). The SHR angles for 3-sec natural and rotation

conditions were not significantly different, nor were the SHR angles for 0.5-sec natural and motionless conditions. The latter result arises because movement was less likely to be initiated in the 0.5-sec natural movement condition prior to signal offset; hence that condition has features in common with motionless listening.

### Front/Back Errors

ANOVA on percentages of front/back errors revealed significant differences across the six conditions [ $F(5,55) = 20.61, p < .001$ ]. Figure 3 shows the percentage of front/back errors occurring in HP trials for each condition. Tukey's HSD tests established that the 3-sec natural and rotation conditions produced significantly fewer front/back confusions than did the 3-sec motionless condition ( $p < .01$ ), and that the 0.5-sec rotation condition produced significantly fewer front/back confusions than did either the 0.5-sec motionless condition ( $p < .01$ ) or the 0.5-sec natural condition ( $p < .05$ ). The percentages of front/back confusions for 3-sec natural and rotation conditions were not significantly different, nor were the percentages of front/back confusions for 0.5-sec natural and motionless conditions. Again, the latter result arises because movement is less likely to be initiated before the offset of the 0.5-sec signal under the natural movement condition. As derived from head tracker coordinates at the moment of responding, there were a few front/back errors in the 3-sec natural movement condition. These occurred with 1 listener who opted not to move throughout the period of the signal. A few front/back errors occurred in natural and rotation conditions with 0.5-sec signals, in cases in which movement was less than 5° from the initial listening orientation. Other front/back errors occurred in the rotation and the 0.5-sec natural movement conditions, even where movement was greater than 5°, but these were confined to sources in a spatial region around the interaural axis.

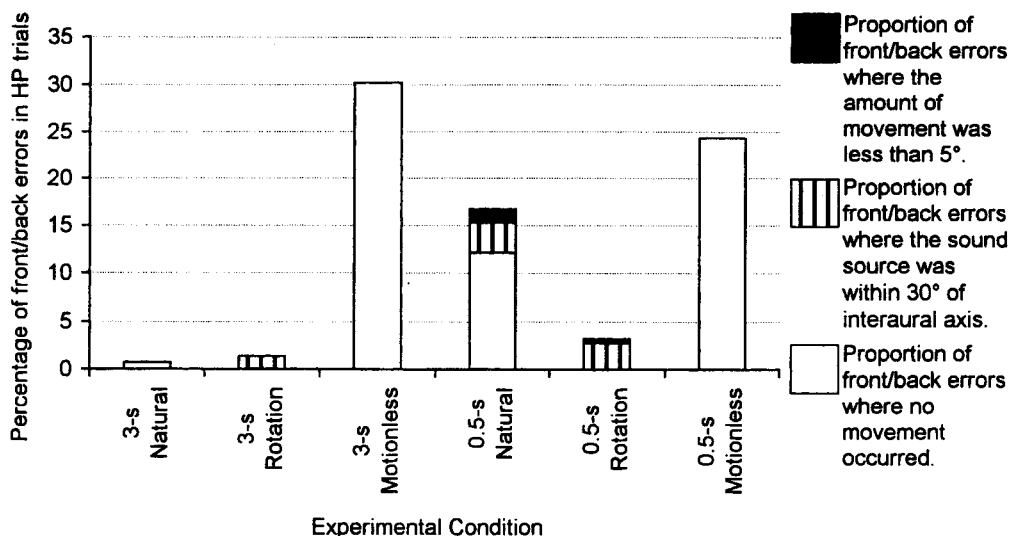


Figure 3. Front/back errors under different conditions.

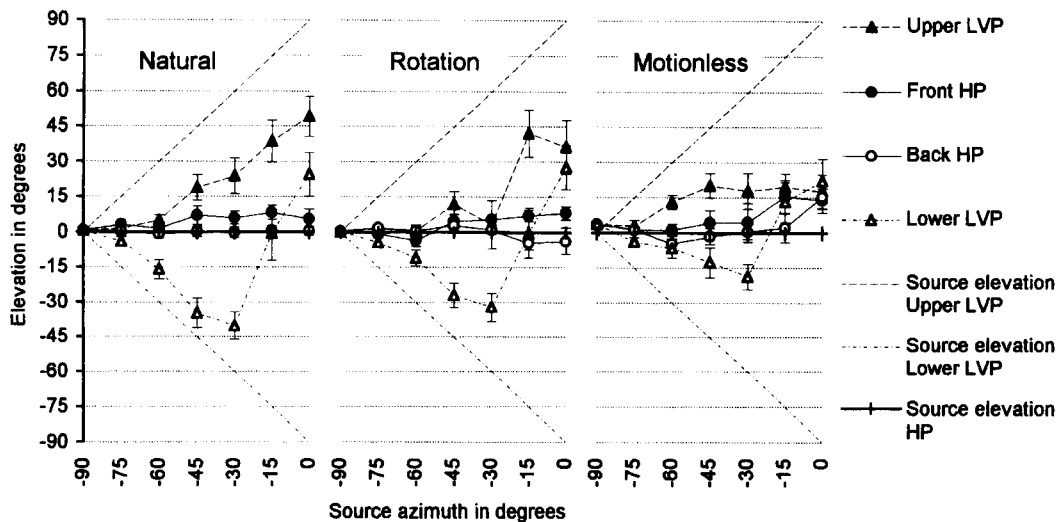


Figure 4. Apparent elevation of sources in four spatial quadrants under three test conditions for the 3-sec signal (error bars indicate standard error of the mean). HP, horizontal plane; LVP, lateral vertical plane.

### Absolute Elevation Error

To assess the effect of movement on the elevation component of localization accuracy, an ANOVA was performed on absolute elevation error. A significant difference was observed across conditions [ $F(5,55) = 9.71$ ,  $p < .001$ ]. Mean values are shown in Table 1. Post hoc testing showed that the 3-sec natural and rotation conditions produced significantly less error than did the 3-sec motionless condition ( $p < .01$  and  $p < .05$ , respectively). There were no significant differences between the 3-sec natural and rotation conditions nor among the various 0.5-sec conditions.

### Apparent Elevation

The average apparent elevation of each sound source is shown in Figure 4 (3-sec signal only) for sources in the LVP above the horizon, LVP below, HP in front of the interaural axis, and HP behind. It may be seen that natural movement offers little advantage over rotation in the perception of displacements above and below the horizon, although the upper LVP function is smoother for natural. Sources below the horizon, at least to a limit of  $-60^\circ$  elevation ( $-30^\circ$  azimuth), were distinguished in both movement conditions, and with a slight advantage over motionless listening in the  $-45^\circ$  to  $-60^\circ$  region of the lower LVP. In all conditions, there was a marked reversal of judged elevation for sources beneath the listener. From  $0^\circ$  to  $\pm 30^\circ$  in the LVP, and throughout the HP, there were no differences among conditions. (For the 0.5-sec signal, there were also no differences among conditions across the range of locations in either plane, with patterns of responses in the upper LVP akin to those in the motionless condition shown in Figure 4, and patterns in the lower LVP akin to those in the rotation condition shown here.)

### Cluster Analysis

There were noticeable variations in the performance of different listeners, especially with respect to elevation judgment. Hierarchical cluster analysis was used to explore the existence of a typology of listeners based on absolute elevation error in each of the three movement/no-movement conditions with 3-sec signals. Ward's minimum variance cluster analysis (Blashfield, 1976) was applied with dissimilarities between listeners' profiles being defined by squared Euclidean distance ( $D^2$ ). The first large increment in aggregate  $D^2$  occurred at the merging of three clusters into two, suggesting that three was an appropriate number to interpret. Clusters 1, 2, and 3 contained 5, 4, and 3 listeners, respectively. Their performance patterns, in the form of average apparent elevation judgments, are given in Figures 5a, 5b, and 5c.

The first cluster showed an advantage for natural movement over rotation or motionless conditions in discriminating upper LVP positions; they showed no clear performance difference between movement conditions in the lower LVP, and virtually no deviation from the horizon for sources on the HP. This relates to the point that Cluster 1 showed little attribution of lower LVP sources to locations above the horizon.

The second cluster showed an advantage for both natural and rotation movements in maintaining fairly proficient performance in the upper LVP compared with motionless listening. There was a substantial attribution of the lowermost LVP source to above the horizon. An "upward" bias in this group may partly explain the higher average apparent elevation for upper LVP sources in this cluster compared with that in Cluster 1. The third cluster showed an advantage for rotation over both natural movement and motionless listening in the uppermost region of the upper LVP. There are signs that the source overhead

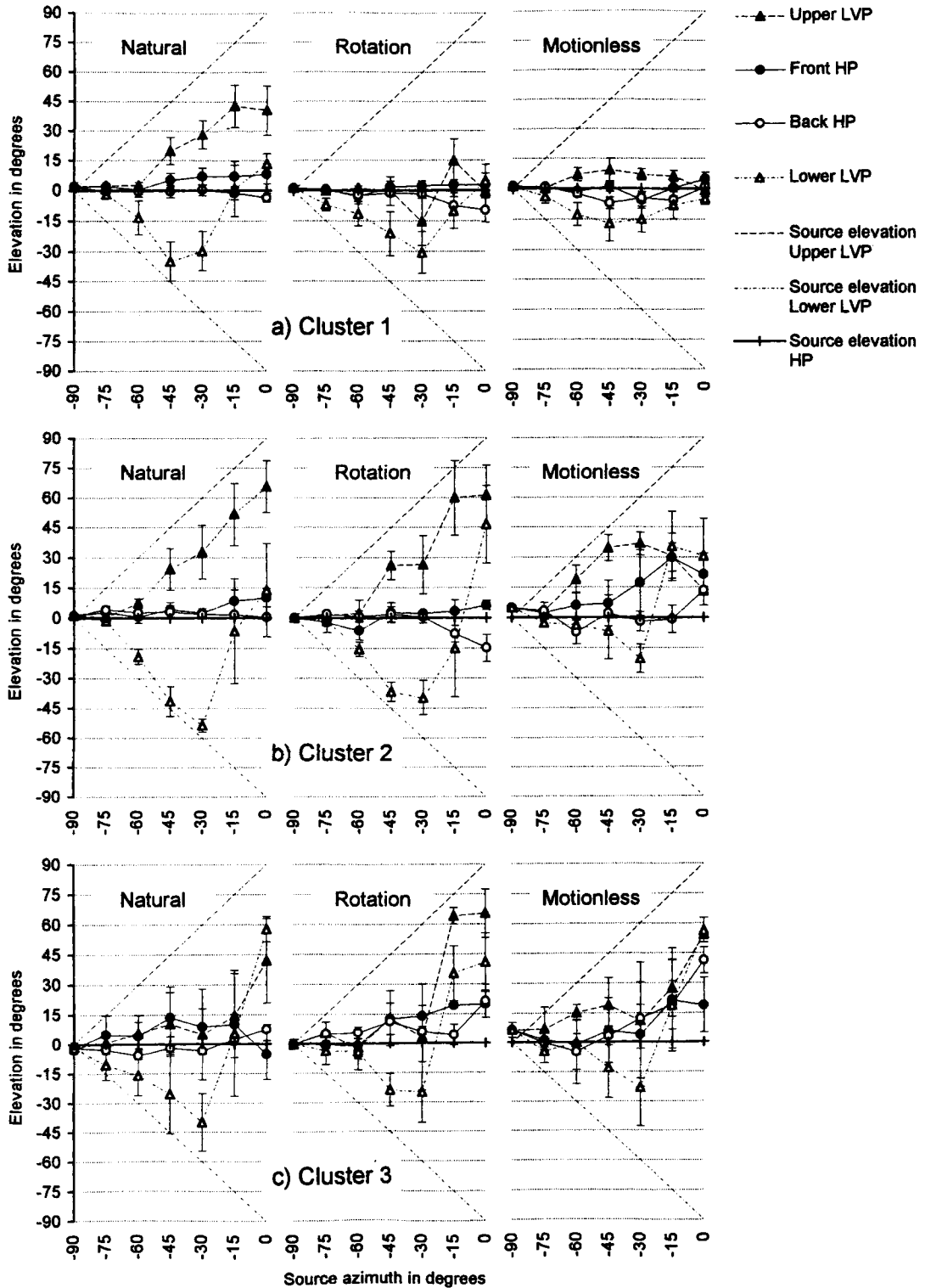


Figure 5. Apparent elevation in the same conditions as for Figure 4, but showing responses of the three cluster groups separately, in 5a, 5b, and 5c. HP, horizontal plane; LVP, lateral vertical plane.

was heard as at or toward that location in all conditions for this cluster, which also showed a distinct attribution of the lowermost source to uppermost positions.

An examination of responses to the overhead source showed two distinct patterns following instructed or natural movement: Either there was a sense of this source being within  $\pm 45^\circ$  of the horizon, or there was a sense of it being at least  $60^\circ$  above the horizon. Using  $+60^\circ$  elevation as a criterion, we noted that 3 of the 12 listeners located the overhead source correctly on each of the four 3-sec movement trials they underwent with respect to that source; 2 others were correct both times in the rotation condition, and 2 others were correct both times under natural movement conditions. Of the 5 listeners achieving success in natural movement, 1 was successful both times under those conditions with the 0.5-sec signal and 3 others were successful once under those conditions. Rotation with the 0.5-sec signal led to success for 1 listener on one trial. On trials in which the source was not detected as overhead, the head tracker records showed that a common response was to identify its whereabouts as either directly in front of or directly behind where the listener was facing at the point of signal offset.

Inspection of the forms of movement made in the natural condition (3-sec signal) showed that in response to HP signals displaced from the MVP, the typical first phase of movement was a single sweeping leftward rotation. In many cases this overshot the actual source location, especially in response to sources at  $-15^\circ$  and  $-30^\circ$ , and was followed by a return to the true position or a rapidly damped oscillation about it. In other cases, more noticeably in the region around the interaural axis and the back HP, a rapid initial leftward rotation was slowed before the true position was reached. When the source was directly in front, there was either no or virtually no movement initiated, or there was an up/down nodding of the head or an oscillatory rotation. In the case of LVP sources at more than  $45^\circ$  from the HP, the initial rotation was often followed by substantial downward and/or upward nodding of the head, usually initiated while rotation to face the LVP region was still being completed.

### Horizontal Error

The conditions of this experiment throw some light on the matter raised by Middlebrooks and Green (1991), that head movement may have little effect except as it allows recruitment of the optimum static localization cues that become available when the source is faced. (A brief account of relevant results is made here; a more detailed analysis will be incorporated in a paper currently in preparation.) In motionless listening conditions, with a 3-sec signal, and when front/back errors are removed from consideration, a substantial ( $10^\circ$ – $20^\circ$ ) horizontal component remains in the SHR angle for most HP sources, especially for those in the front HP. In other words, listeners are not able to accurately face the source once the signal has ceased. By contrast, in both the rotation and natural movement conditions, the horizontal component for all sources in the front HP was always

less than  $4^\circ$ , and there was no difference between the two conditions. Thus, even when movement is limited to a single rotation to a fixed point, other locations can be accurately oriented to, following signal offset.

With the 0.5-sec signal there was a high incidence of nonmovement in the natural condition prior to signal offset, whereas listeners were able to execute at least a part of the instructed rotation while the signal was still on. The different outcome between rotation and motionless conditions, noted above for the 3-sec signal, is preserved for the briefer signal, whereas that for the natural condition is not. When trials are excluded in which no or virtually no movement occurred in the 0.5-sec natural condition, the pattern observed for the natural condition with the 3-sec signal is restored, but only for sources in the front HP, evidently because the back HP is not able to be covered by the movement.

The 3-sec rotation condition provides both facing and nonfacing conditions, in the sense that performance with the loudspeaker at  $-45^\circ$  on the HP represents a "face-on" posture, and performance at other horizontal angles represents nonfacing. Similarly, the motionless condition represents face-on to the loudspeaker at  $0^\circ$ . Horizontal error was very slightly lower for the source at  $-45^\circ$  in the rotation condition than at other front HP positions, and substantially lower at  $0^\circ$  in the motionless conditions than at other positions. Along with the lessened accuracy for back HP sources in the 0.5-sec natural condition, the data indicate that the "face-on" argument can also be supported.

## DISCUSSION AND CONCLUSIONS

### Absolute Accuracy

The overall result shows clearly that, with a signal of 3-sec duration, both natural head movement and a single  $45^\circ$  rotation provide almost identical proficiency as regards the magnitude of angular deviation of responses from source positions. Whereas the 3-sec natural condition may be one that enables a listener to initiate movement so as then to rely on static cues, the rotation condition limits the system, in most cases, to a set of changing interaural relationships. Wallach's (1939, 1940) claim that such a changing dynamic may be an information source gains support from this general result. The outcome of movement versus nonmovement with a short-duration signal includes the complicating factor that little movement was observed in the 0.5-sec natural condition. That last point is discussed later.

### Front/Back Errors

The experiment shows very clearly that head movement contributes significantly to localization of sources in terms of resolving front/back ambiguity, even when the duration of the signal is quite short. In noting this it must also be considered that because a listener needs time to react to the onset of a noise burst, the head will not rotate for the full duration of a 0.5-sec signal. Analysis of records for the 0.5-sec rotation condition revealed

that the mean duration under which head rotation and noise burst occurred simultaneously was 290 msec. This suggests that the auditory system is indeed sensitive to spatial cues brought about by head rotation. That conclusion is further supported by the finding, in the natural movement condition, that front/back errors did not occur with a 0.5-sec signal if a movement through as little as 5° was accomplished before signal offset. Again, Wallach's (1939, 1940) claim that there is information in the different interaction between head rotation and the position of a source forward or rearward of the interaural axis is supported by these outcomes.

There were occasional front/back errors in the region  $\pm 30^\circ$  from the interaural axis. Inspection of the distribution of front/back errors showed a peak for their occurrence in that region under all conditions. Given the reduced acuity for spatial discrimination in the region around the interaural axis (Mills, 1972), some errors in that area, which were counted as front/back, may be better seen as instances of localization "blur."

### Elevation

Above the horizon in the LVP, shoulder and torso reflections may help to account for the small amount of elevation discrimination even in motionless conditions (Perrett & Noble, 1995; and Figure 4 here). Not all listeners could use movement cues to distinguish the elevation of sources in the upper LVP, and the case of the signal overhead may explain why. Interaural differences for that signal remain null throughout any head rotation. Some listeners are sensitive to the unchanging geometry of interaural events in the face of changing head position and detect the signal as being above them. The situation may be one in which, for a signal of 3-sec duration, the state of interaural differences at the start of a trial, prior to head movement, can be compared to their state following the cessation of such movement. Other listeners seem to attend only to final input conditions. For a source overhead, the final input condition is that the signal is in the (repositioned) MVP, and, in the absence of pinna cues for elevation, the signal is heard, relative to the listener's final orientation, as straight ahead or straight behind. We may contrast these as integrative versus non-integrative solutions. With a 0.5-sec signal, the head tracker records showed that the sound had ceased before movement was completed; hence the same comparison was unavailable. This may explain the markedly fewer signs of elevation detection with the shorter signal.

With sources in the lower LVP, listeners' judgments were fairly proficient, with rotation or natural movement, from 0° to -60° elevation (-90° to -60° azimuth). Even in motionless conditions, listeners could perceive the elevation of lower LVP sources to some extent—we comment presently on the marked apparent elevation of the lowermost source. The appearance of (slight) detection of the elevation of sources above and below the horizon in the absence of movement and pinna cues suggests a role for the shoulders and torso in altering the spectrum

of the signal (Gardner, 1973; Kuhn, 1987). The improvement with 45° rotation suggests the possibility of an interaction between such bodily spectral effects and changes in interaural differences. Such a possibility is outside the scope of Wallach's (1939, 1940) hypotheses, in which a side-to-side head tilt was thought necessary to resolve locations above and below the horizon.

There was little sign of greater accuracy in elevation detection, with movement as opposed to no movement, up to at least  $\pm 30^\circ$ —the effect of movement becoming more clearly seen at greater distances from the horizon. According to Wallach (1939, 1940), this finding would be consistent with his theory because differences in the rate of change in azimuth angle relative to head orientation, for positions within 30° of the HP, are minor compared to those at greater elevations. This point may also explain why Thurlow and Runge (1967) found little effect for movement over no movement, since their sources were only up to 41° above and below the horizon.

For the lowermost source, there was often observed a "lowest-to-highest" effect. This is a coherent outcome in one sense: The listener's seat acts to disperse energy from the lowermost region; hence there are no body-related cues along with the null interaural difference. Hence, further, there are no cues to anchor the sound to the lowermost point, and it is as plausibly heard overhead as below. This suggests, by contrast, that there are cues that anchor sources to locations in front of or behind the listener, since, in general, these were not heard as overhead or below, even in motionless listening. Shoulder and body reflections may offer such an anchor. We note that Cluster 1 heard the lowermost source as at the horizon, and that this cluster also derived no benefit from rotation for the signal overhead; under that condition they heard the signal overhead as also being on the horizon in front of or behind them. For the sample as a whole, sources overhead were usually not heard as below. This may reflect a bias derived from conditions in the everyday environment.

### Horizontal Error

The present study provides evidence in general that head motion can contribute to localization accuracy even when the sound is not of sufficient duration to allow the listener to reorient so as to face the source, or, as in the rotation condition, the instruction acts to prevent this from occurring. The bulk of the effect is in the dispelling of front/back errors, but when trials involving front/back errors are discounted, listeners are still generally more accurate in rotating to face HP sources compared with motionless listening. That said, there is also some evidence in favor of the point made by Middlebrooks and Green (1991). For instance, highest accuracy in horizontal localization is observed in the motionless condition at the place which is face-on throughout, and horizontal errors for brief signals, associated with truncated forms of natural movement, are greatest in spatial regions well away from the path of movement traversed while the source was active.



## General

The present study suggests that the major contribution to localization of rotational head motion cues may be in eliminating large (front/back) errors. Furthermore, head movement can indicate the direction of sources lying directly overhead and offers improved horizontal and vertical localization compared with motionless listening. There seems little additional benefit from natural movement over rotation. The information for elevation is therefore the different transformations derived from movement angle relative to azimuth angle—what may be dubbed the “Wallach cue.” Evidence bearing that out is in the head tracker records in 3-sec natural movement for some listeners. These show that rotational movement is predominant for HP sources off the MVP, but, in contrast, there is substantial up/down nodding movements, superimposed upon the ongoing rotation, for LVP sources greater than 45° from the horizon. The appreciation that the source may be elevated, but with uncertainty as to the direction of that elevation, are both arguably derived from the rotational Wallach cue.

Allowing for individual variability, it was generally found that low-pass LVP sources tend to be heard at equivalent cone-of-confusion loci on the horizon in motionless listening. This is consistent with an earlier report (Perrett & Noble, 1995) that the system takes the horizon as its default plane when there is no strong information for placement outside of that.

The cluster groups reflected the different test orders to some extent: Three of the 5 in Cluster 1 had natural movement as their first condition, 3 of the 4 in Cluster 2 began with no movement, and all 3 in Cluster 3 began with rotation. We were (and remain) unsure how the different orders of test would influence behavior across conditions; counterbalancing was used precisely to control for any effects. It is not unexpected that there may be interactions between performance and orders of testing, and that these interactions will be observed in different people to different degrees.

A faster reaction time (RT) occurred with rotation alone compared with natural listening (on average, 0.3 vs. 0.4 sec in the 3-sec conditions; 0.2 vs. 0.5 sec in the 0.5-sec conditions). It may be reasoned that the natural listening task required the listener to give priority to localization cues before moving the head, whereas a preinstruction to rotate the head as soon as possible after the onset of the sound required the listener merely to hear the sound before moving. Attending to localization cues may have induced a greater processing load, resulting in the extra time taken before reacting.

It could further be argued that a preinstruction to move as soon as a sound is heard is an artificial listening condition. On the other hand, it may have features in common with reactions called for in the real world—for example, rapid reaction to a brief warning signal. In any case, it was effective. In everyday situations, listeners are likely to encounter sounds that fail to provide pinna cues.

Sounds can lack the frequencies necessary for pinna cue production because their initial generation has excluded them or because physical barriers have had a low-pass filtering effect. The indications from the present study are that head rotation, occurring at the same time as a sound lacking high-frequency energy, is likely to increase localization accuracy, even if the sound is of short duration.

Further investigation is needed to enable a more comprehensive examination of Wallach's (1939, 1940) classic proposals, but the evidence from the experiment reported here is that front/back ambiguity is undoubtedly resolved by head rotation, and there are signs that, for some listeners in some conditions, movement contributes to elevation judgments. These outcomes give support to Wallach's position. The role of movement cues has become overshadowed by more recent advances in detailed knowledge about the contributions of the pinnae to auditory spatial perception; the results of the study reported here suggest that it is fruitful to give experimental attention to dynamic cues.

## REFERENCES

- BLASHFIELD, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. *Psychological Bulletin*, **83**, 377-388.
- GARDNER, M. B. (1973). Some monaural and binaural facets of median plane localization. *Journal of the Acoustical Society of America*, **54**, 1489-1495.
- 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.
- KUHN, G. F. (1987). Physical acoustics and measurements pertaining to directional hearing. In W. A. Yost & G. Gourevitch (Eds.), *Directional hearing* (pp. 3-25). New York: Springer-Verlag.
- 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.
- MILLS, A. W. (1972). Auditory localization. In J. V. Tobias (Ed.), *Foundations of modern auditory theory* (pp. 303-348). New York: Academic Press.
- PERRETT, S., & NOBLE, W. (1995). Available response choices affect localization of sound. *Perception & Psychophysics*, **57**, 150-158.
- POLLACK, I., & ROSE, M. (1967). Effect of head movement on the localization of sounds in the equatorial plane. *Perception & Psychophysics*, **2**, 591-596.
- THURLOW, W. R., & MERGENER, J. R. (1970). Effect on stimulus duration on localization of direction of noise stimuli. *Journal of Speech & Hearing Research*, **13**, 826-838.
- THURLOW, W. R., & RUNGE, P. S. (1967). Effect of induced head movements on localization of direction of sounds. *Journal of the Acoustical Society of America*, **42**, 480-488.
- WALLACH, H. (1939). On sound localization. *Journal of the Acoustical Society of America*, **10**, 270-274.
- WALLACH, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, **27**, 339-368.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental psychology* (2nd ed.). London: Methuen.