

The auditory motion aftereffect: Its tuning and specificity in the spatial and frequency domains

CHANG-JIANG DONG, NICHOLAS V. SWINDALE, and PIERRE ZAKARASKAS
University of British Columbia, Vancouver, British Columbia, Canada

VINCENT HAYWARD
McGill University, Montreal, Quebec, Canada

and

MAX S. CYNADER
University of British Columbia, Vancouver, British Columbia, Canada

In this paper, the auditory motion aftereffect (aMAE) was studied, using real moving sound as both the adapting and the test stimulus. The sound was generated by a loudspeaker mounted on a robot arm that was able to move quietly in three-dimensional space. A total of 7 subjects with normal hearing were tested in three experiments. The results from Experiment 1 showed a robust and reliable negative aMAE in all the subjects. After listening to a sound source moving repeatedly to the right, a stationary sound source was perceived to move to the left. The magnitude of the aMAE tended to increase with adapting velocity up to the highest velocity tested (20°/sec). The aftereffect was largest when the adapting and the test stimuli had similar spatial location and frequency content. Offsetting the locations of the adapting and the test stimuli by 20° reduced the size of the effect by about 50%. A similar decline occurred when the frequency of the adapting and the test stimuli differed by one octave. Our results suggest that the human auditory system possesses specialized mechanisms for detecting auditory motion in the spatial domain.

In vision, the motion aftereffect (MAE) is a common phenomenon that can be often observed in the natural environment and easily demonstrated experimentally. The well-known waterfall illusion and the spiral aftereffect are just two examples. After a few minutes of viewing an object moving in a single direction (the adapting stimulus), a stationary object (the test stimulus) appears to move in the opposite direction. This visual MAE (vMAE) has been extensively studied for more than a century and has been taken as psychophysical evidence for the existence of specialized motion detection channels in the visual system (e.g., Gates, 1934; Wade, 1994; Wohlgenuth, 1911).

In the visual studies, the duration of the vMAE and the apparent velocity of the test stimulus have commonly been used as measures of its magnitude. It has been shown that the duration of the vMAE increases with the duration of the adapting stimulus and can last as long as 10–30 sec (Herschenson, 1989; Sekuler & Ganz, 1963). Herschen-

son (1993) reported that the duration of the vMAE increased in proportion to the square root of adaptation time up to 90 sec. By using a nulling method, in which the subject continuously adjusts the velocity of a test stimulus moving in the direction opposite to that of the vMAE until the test stimulus appears to be stationary, the apparent velocity of the vMAE over time can be recorded. It has been found that, within a limited range, the velocity of the vMAE increases with the velocity of the adapting stimulus (Brigner, 1986; Taylor, 1963; Wohlgenuth, 1911). In addition, the vMAE is restricted to the retinal area stimulated by adapting motion (Wohlgenuth, 1911), displays a spatial frequency specificity (Brigner, 1982; Cameron, Baker, & Boulton, 1992; Over, Broerse, Crassini, & Lovegrove, 1973), and transfers between eyes (Holland, 1957; O'Shea & Crassini, 1981; Wohlgenuth, 1911).

A commonly held explanation for the vMAE is that the visual system contains specialized motion detection channels and that these channels are distinct from those that process stationary position or depth (Beverley & Regan, 1973; Pantle & Sekuler, 1968; Regan & Beverley, 1973; Regan, Beverley, & Cynader, 1979; Sekuler & Ganz, 1963; Sekuler & Pantle, 1967). During prolonged exposure to one direction of motion, neurons sensitive to that direction of motion adapt (Barlow & Hill, 1963; Giaschi, Douglas, Marlin, & Cynader, 1993; Marlin, Hasan, & Cynader, 1988; Saul & Cynader, 1989a, 1989b). Then, when a static stimulus is presented, the activities of

Part of this work was presented at the 134th Meeting of the Acoustical Society of America, San Diego, December 1997. The present study was supported by a NSERC grant to M.S.C. We thank Eric De Silva for his contribution to the design and construction of the acoustical stimulation system used in the study and all the subjects who kindly volunteered to participate in the study. Correspondence concerning this article should be addressed to C.-J. Dong, Department of Ophthalmology, University of British Columbia, 2550 Willow Street, Vancouver, BC V5Z 3N9, Canada (e-mail: cdong@interchg.ubc.ca).

the neurons sensitive to the opposite direction dominate, causing the stationary object to be perceived as moving in the direction opposite to that of the adapting motion.

Although specialized motion-sensitive channels are well established in vision, their existence in the auditory system is still a subject of debate. In principle, there are two different interpretations for the perception of sound source motion by the auditory system (Middlebrooks & Green, 1991). One, the *snapshot* theory, is that the auditory system simply uses the same neural mechanism that is used to localize stationary sound sources to register the positions of the sources at different times and then compares them to extract the velocity. In this case, a specialized system for motion detection is not needed or employed. The second interpretation is that there exist specialized motion detection channels in the auditory system. These specialized channels are most sensitive to the dynamic aspects of localization cues—for example, time-varying interaural intensity difference (IID), interaural time difference (ITD), or time-varying spectral information available to each ear. As Lappin, Bell, Harm, and Kottas (1975) proposed, source velocity might be a “directly perceived attribute of moving stimuli” (p. 393).

Compared with studies of the vMAE, studies of analogous auditory MAE (aMAE) are relatively scarce (e.g., Ehrenstein, 1994; Grantham, 1989; Grantham & Wightman, 1979; Reinhardt-Rutland, 1992). In these studies, simulated moving sound was used as the stimulus. In the Reinhardt-Rutland study, simulated moving sound was generated by simultaneously changing the sound level of 1000-Hz sinusoids at the two ears presented over headphones in opposite directions. After adaptation, a changing-loudness aftereffect was reported for the monaural test stimulus by all the subjects. However, when tested binaurally, no subjects observed any MAE or loudness aftereffect. Ehrenstein reported that after 90 sec of listening to a simulated moving sound produced by dynamically varying the IID or the ITD, a stationary test stimulus was displaced in the direction opposite to that of adaptation, but no apparent motion was observed. However, when simulated moving sound was used as both the adapting and the test stimuli (Grantham & Wightman, 1979) or just as the test stimulus (Grantham, 1989), an aMAE was found, although it lasted only 1–3 sec and was not observed in all the subjects tested. In the study by Grantham and Wightman, an aMAE was reported for 500-Hz sinusoids but was virtually absent for 2000-Hz pure tones. When the velocity of an adapting stimulus (a 500-Hz low-pass noise) was below 30°/sec, an aMAE was observed only for 2 out of 4 subjects tested (Grantham, 1989).

Simulated moving sound may be a less-than-ideal stimulus to use, because it provides the auditory system with incomplete motion cues. For example, when simulated moving stimuli are presented through headphones and the motion is achieved by dynamically varying the IIDs or the ITDs, the modifying effects of the head and pinna on the incoming signals are bypassed altogether. Thus, the auditory motion detection channels may not be adequately stimulated. However, a real moving sound can pro-

vide a more natural stimulus and, potentially, more localization cues, including a time-varying frequency spectrum, as well as time-varying intensity and phase information. Here, we report the results of experiments in which a real moving sound was used as both the adapting and the test stimuli. Sound stimuli were generated by a loudspeaker mounted on a robot arm that was able to move smoothly and quietly at velocities of up to 30°/sec on the surface of an imaginary sphere centered on the subject's head. With this stimulus, a robust and repeatable aMAE was demonstrated in all (7/7) the subjects tested. Experiment 1 confirmed the existence of the aMAE and its dependence on adapting velocity. Experiments 2 and 3 examined the tuning and specificity of the aMAE in the spatial and frequency domains.¹

GENERAL METHOD

Subjects

A total of 7 subjects (T.S., R.C., M.H., M.L., J.Q., P.Z., and C.D.) with clinically normal hearing participated in these three experiments. All the subjects, 23–39 years of age, were recruited from members of the Ophthalmology Research Lab or were students at the University of British Columbia. Except for the two authors (P.Z. and C.D.), all the subjects were unaware of the purpose of the experiments.

Apparatus

In order to study the aMAE, an acoustical stimulus system (Figure 1) was designed that allows a sound source to move with a given velocity (up to 30°/sec) along a specified trajectory. The stimulus system consisted of a robot arm and control circuits and software. The robot arm was designed and constructed at the Department of Electrical Engineering at McGill University. Control software for the robot arm was developed by one of the authors (N.V.S.). In this system, a loudspeaker (LCS-150, Labtec) was mounted at the *tip* (the end effector) of the robot arm, which is a servo-controlled mechanism of a closed-loop kinematic chain (see the Appendix for further details of the design). Servo motors were interfaced with an IBM PC computer through a PC motion control interface card (MFIO-3A, Precision MicroDynamics). The system was designed to move the loudspeaker smoothly, quietly, and safely so that the trajectory of the loudspeaker was constrained to lie on the surface of a sphere with a radius of 0.8 m. The frequency spectrum of the noise produced by the motion of the robot arm is shown in Figure 2.² When the subject's head was located at the center of this sphere, the loudspeaker was always oriented toward the subject's head.

The area of the surface that could be reached by the loudspeaker was about 63% of the total surface area of the sphere, covering almost all of the subject's frontal hemifield. The loudspeaker was connected to a soundblaster card (Sound Blaster 16, Creative Labs), which was programmed to generate different kinds of sound, including white noise, band-pass noise, pure tones, and clicks. These sounds were synchronized with the movement of the robot arm. In this way, the sound was activated after a given latency when the loudspeaker moved into a given spatial region. This acoustical stimulation system was positioned in an acoustically treated soundproof chamber, which was modified from an IAC sound-insulated chamber (Industrial Acoustics Company). Taking advantage of this unique sound stimulator, a series of psychoacoustical experiments was carried out to study the aMAE.

Procedure

A two-alternative forced-choice paradigm was used to measure the aMAE. In the initial adaptation period (160–200 sec), a sound

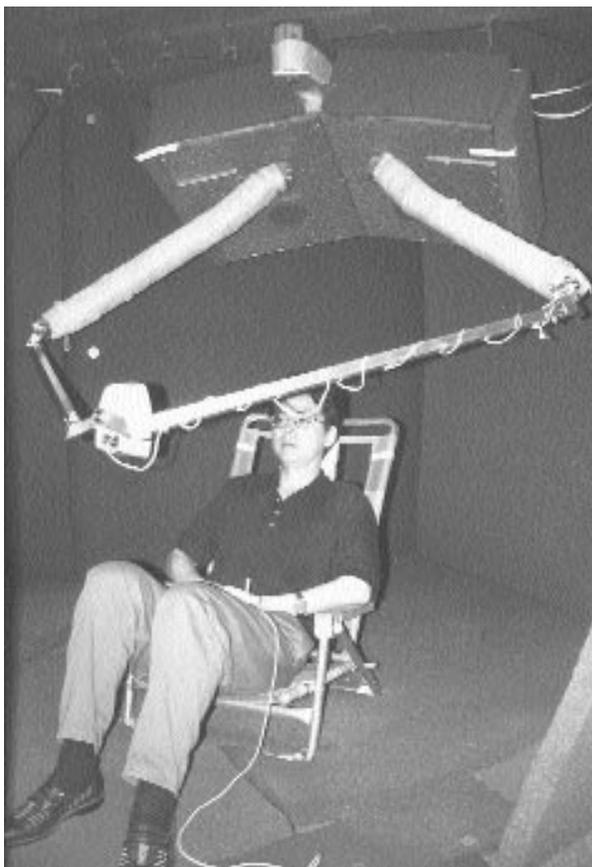


Figure 1. The robot arm used in this study. The arm is positioned in an acoustically treated soundproof chamber. A loudspeaker (LCS-150, Labtec) mounted on the arm can move smoothly and quietly on the imaginary surface of a sphere with a radius of 0.8 m. In the picture, a subject is seated in the chamber with his head at the center of the sphere.

source repeatedly swept a specified arc in a single direction. Between two successive adapting sweeps, there was a silent interval of about 1 sec, which was required to reposition the loudspeaker to the start position and to accelerate it to the desired speed (Figure 3). By varying the number of adapting sweeps, the total duration of actual sound motion during the initial adaptation was kept constant (120 sec) for all of the adapting velocities tested. During the test period, the subject's task was to indicate the direction (i.e., left or right) of a brief test sound by pressing one of two buttons.

The velocity of the test stimulus was randomized from trial to trial. In order to maintain the subject in an adapted state, following each test, there was a readaptation period (8–10 sec), within which the subject was exposed to the adapting sound again for a total moving sound exposure of 6 sec (Figure 3). The range of test velocities was always centered on zero and was chosen so that the extremes of the range were almost always correctly identified by the subject. A psychometric function was generated by plotting the percentage of *left* responses to the different test velocities presented during a block of trials. Probit analysis (Finney, 1971) was used to estimate the 50% response rate on the resulting psychometric function. This was the stimulus velocity that sounded stationary to the subjects. If there was an aMAE, the test velocity that appeared to be stationary would shift in the direction of the adapting velocity. In our studies,

this mean velocity was used as a measure of the magnitude of the MAE.

Close inspection of the experiments done by Grantham and Wightman (1979; Grantham, 1989) reveals a potential drawback in their experimental design. In their studies, test stimuli moved symmetrically about the midline. Therefore, the direction of motion of a test stimulus could be decided on the basis of the start or end position of the test stimulus, relative to the midline. For example, if the start position of a test stimulus was on the right side of the midline, there may have been a bias to report that the test stimulus moved to the left. The subjects in these studies might have used the localization cues for the start or end position of a test stimulus to judge the direction of the moving test stimulus. To overcome this potential drawback in our experiments, for each start position, a test stimulus could move randomly in either direction, left or right. Thus, the start and/or the end positions could no longer provide direction cues for the subjects, except in extreme cases.

All the experiments were conducted in the darkened soundproof chamber described above. The subjects were seated at the center of the sphere defined by the motion of the loudspeaker and were instructed to keep their eyes closed and maintain a steady upright posture throughout the course of experimentation. A headrest was provided to prevent the subjects from tilting their heads either sideways or forward. Before data were collected, at least 2 h of training were provided to each subject until the performance of the subject was stable. The order of trial blocks with different stimulus parameters (i.e., adapting velocity, stimulus frequency, or spatial region) was randomized for each of the three experiments in order to minimize any order effects. Each session lasted about 1 h, which typically included three individual trial blocks. For example, in Experiment 1, each session consisted of three trial blocks with different adapting velocities: $+20^\circ/\text{sec}$, $0^\circ/\text{sec}$ (control), and $-20^\circ/\text{sec}$, where positive velocities indicate that the direction of the moving stimulus is to the right, whereas negative velocities indicate that the direction is to the left. After each block of trials, a 5-min break was provided in order to prevent the subject from fatiguing. Each experiment was spread over 3 weeks.

EXPERIMENT 1

This experiment was designed to demonstrate the aMAE and to analyze the magnitude of the aMAE as a function of adapting velocity.

Method

Four subjects (T.S., R.C., P.Z., and C.D.) were tested. Broadband noise (500–14000 Hz, Figure 4) was used as both the adapting and the test stimuli. The average level of each sound, presented from a stationary loudspeaker at 0° azimuth at the subject's ear level and measured at the subject's head position (0.8 m distant), was about 75 dBA. In each block of trials, the adapting velocity was constant. During adaptation, the adapting stimulus repeatedly traversed an arc of 30° ($\pm 15^\circ$ centered on the midline) at the subject's ear level with one of six velocities ($\pm 10^\circ/\text{sec}$, $\pm 15^\circ/\text{sec}$, and $\pm 20^\circ/\text{sec}$). The duration of each sweep was 3, 2, and 1.5 sec for adapting velocities of $10^\circ/\text{sec}$, $15^\circ/\text{sec}$, and $20^\circ/\text{sec}$, respectively. In order to keep the total duration of exposure to the moving sound equal (120 sec), different numbers of adapting sweeps (40, 60, and 80) were presented for different adapting velocities ($10^\circ/\text{sec}$, $15^\circ/\text{sec}$, and $20^\circ/\text{sec}$, respectively). During testing, the test stimulus (1 sec of broadband noise) moved at one of eight velocities ($\pm 1^\circ/\text{sec}$, $\pm 3^\circ/\text{sec}$, $\pm 5^\circ/\text{sec}$, and $\pm 7^\circ/\text{sec}$) with one of eight possible start positions ($\pm 1^\circ$, $\pm 3^\circ$, $\pm 5^\circ$, and $\pm 7^\circ$). Therefore, a total of 64 test presentations was given on each block of trials (8 test velocities \times 8 start positions), and the whole adapting range (from -15° to 15° azimuth) was covered by the motion of these test stimuli. For each adapting velocity, there

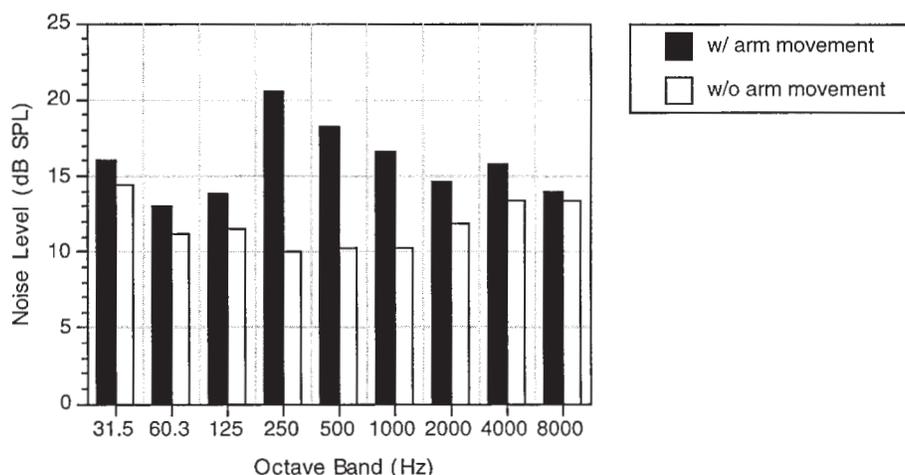


Figure 2. Background noise level measured with the MLSSA system at the subject's position in the acoustically treated soundproof chamber when the robot arm was still and when it was in motion.

was a control trial block, in which the adapting stimulus was a stationary sound (i.e., one with a velocity of $0^\circ/\text{sec}$) presented directly in front of the subject (0.8 m distant) with the same time course as the corresponding moving adaptors. As was described earlier, in each block of trials, percentage of *left* responses was plotted as a function of test velocity (eight responses per test velocity), and the velocity that appeared stationary was obtained from the 50% point on the resulting psychometric function. This value (the test velocity that appeared stationary) served as the measure of aMAE magnitude for a given condition. The mean aMAE and the standard error of the mean were calculated on the basis of at least three separate blocks of trials obtained with each adaptation velocity.

Results and Discussion

In Figure 5, the magnitude of the aMAE from all 4 subjects in Experiment 1 is plotted as a function of adapting velocity. Different panels represent the results from different individual subjects. The results showed clear aMAEs for all the subjects for adapting velocities of $10^\circ/\text{sec}$, $15^\circ/\text{sec}$, and $20^\circ/\text{sec}$. After adaptation, the point of subjective stationarity shifted in the expected direction. That is, when the adapting velocity was positive (i.e., to the right), the test velocity that appeared to be stationary changed in a positive direction. Thus, a positive velocity was judged by the subject as stationary, and correspondingly a stationary sound was heard as moving to the left.

In Figure 6, the aMAE averaged over all 4 subjects tested in Experiment 1 is displayed as a function of adapting velocity. In the top panel, as in Figure 4, the magnitude of the aMAE is measured as the mean perceived stationary velocity, whereas in the bottom panel, in order to analyze the aMAE in terms of gain, the aMAE is expressed as a percentage of the adapting velocity. Although, as was expected, in the top panel of Figure 5 the magnitude of the aMAE increased with the adapting velocity, the results in the bottom panel show that the gain of the aMAE decreased from 17% to 12.8% as the adapting velocity increased from $10^\circ/\text{sec}$ to $20^\circ/\text{sec}$.

EXPERIMENT 2

This experiment was designed to study the tuning and specificity of the aMAE in the spatial domain.

Method

Three subjects (M.H., T.S., and C.D.) participated in Experiment 2, 2 of whom were also subjects in Experiment 1. In this experiment, an arc of 70° ($\pm 35^\circ$ centered on the midline) at the subject's ear level was divided into seven equal subregions, each 10° of arc (-35° to -25° , -25° to -15° , -15° to -5° , -5° to 5° , 5° to 15° , 15° to 25° , and 25° to 35°). Two experimental conditions were used. In Condition 1, in order to obtain the spatial tuning of the aMAE, the adapting and test regions were the same, and the magnitude of the aMAE was measured as a function of spatial region. In Condition 2, in order to determine whether the aMAE was spatially specific, the adapting and the test stimuli were presented in separate spatial regions, and the magnitude of the aMAE was analyzed as a function of the distance between the two separate loci. In this condition, the region swept by the adapting stimulus was a 10° arc ($\pm 5^\circ$ centered on the midline), whereas the test region was chosen from one of six separate regions (-35° to -25° , -25° to -15° , -15° to -5° , 5° to 15° , 15° to 25° , and 25° to 35°). Only one adapting velocity (i.e., $20^\circ/\text{sec}$) was used in this experiment. The test stimulus was always a 0.5-sec presentation of broadband noise, moving at one of six velocities ($\pm 1^\circ/\text{sec}$, $\pm 3^\circ/\text{sec}$, and $\pm 5^\circ/\text{sec}$), with one of four start positions, two of which were 1° on either side of the center of the test region and the other two 3° . Taking the central spatial region (-5° to 5°) as an example, the start positions of the test stimulus were $\pm 1^\circ$ and $\pm 3^\circ$ azimuths. During each run, eight responses were collected for each test velocity, two at each start position. As in Experiment 1, a psychometric function was constructed plotting percentage of *left* responses (out of eight) as a function of test velocity. For each condition, three or four runs were obtained, and the mean and standard error of the aMAE magnitude for each condition was computed from these three or four functions, as was described for Experiment 1.

Results and Discussion

The results obtained from Conditions 1 and 2 for all the subjects in Experiment 2 are displayed in Figures 7 and

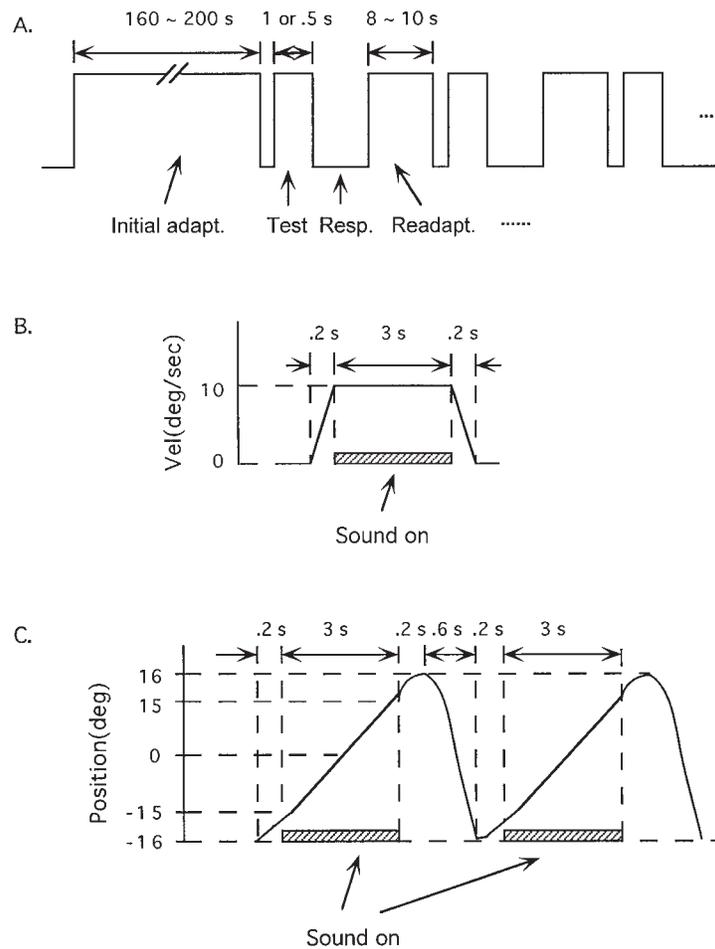


Figure 3. (A) Time sequence of the stimuli. See text for details. (B) Speaker velocity as a function of time during one adapting sweep. In this example, the adapting velocity is 10°/sec. After 0.2 sec of acceleration, the speaker reaches its target velocity of 10°/sec. Then, the speaker moves at this speed (10°/sec) for 3 sec with the sound on (highlighted by a filled bar), followed by 0.2 sec of deceleration. (C) Speaker position as a function of time during the adaptation period. Taking an adapting velocity of 10°/sec as an example, since the adapting region is an arc of 30° centered at 0° azimuth, the duration of one sweep is 3 sec (highlighted by filled bars). Note that between sweeps, there is a silent interval of 1 sec, consisting of 0.2 sec of deceleration of the speaker immediately following each sweep, a period of 0.6 sec for repositioning the speaker to its start position, and 0.2 sec of acceleration to get the speaker to the velocity of 10°/sec immediately preceding the next sweep.

8, respectively. In these figures, the results from different subjects are shown in different panels. In Figure 7, the magnitude of the aMAE was plotted as a function of spatial region. Results showed that within the spatial region tested (from -35° to 35°), the spatial tuning curves were relatively independent of position. That the aMAE is spatially specific is clearly shown in Figure 8, in which the magnitude of the aMAE was plotted as a function of location of the test region. When the test region coincided with the adapting region (from -5° to 5°), the aMAE elicited was largest. As the distance between the test and the adapting regions increased, the aMAE diminished,

resulting in an inverted-V shape of the spatial tuning function. Since the measure of the space constant can provide some information about the size of the receptive field of a *channel*, a space constant was determined from the spatial tuning curves in Figure 8, for each subject and for the average over all 3 subjects tested in Experiment 2, by finding the best-fitting Gaussian function and taking the standard deviation as a measure of the space constant. These were 15.3°, 14.9°, and 19.5° for Subjects T.S., M.H., and C.D., respectively. The space constant obtained from the average spatial tuning curve of these 3 subjects was 16.4°.

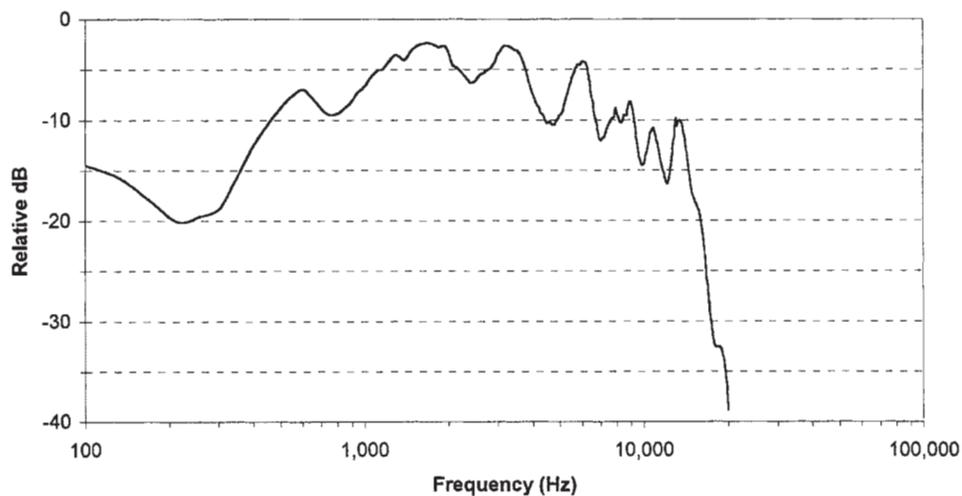


Figure 4. The frequency spectrum of the broadband noise, measured at the position of the subject.

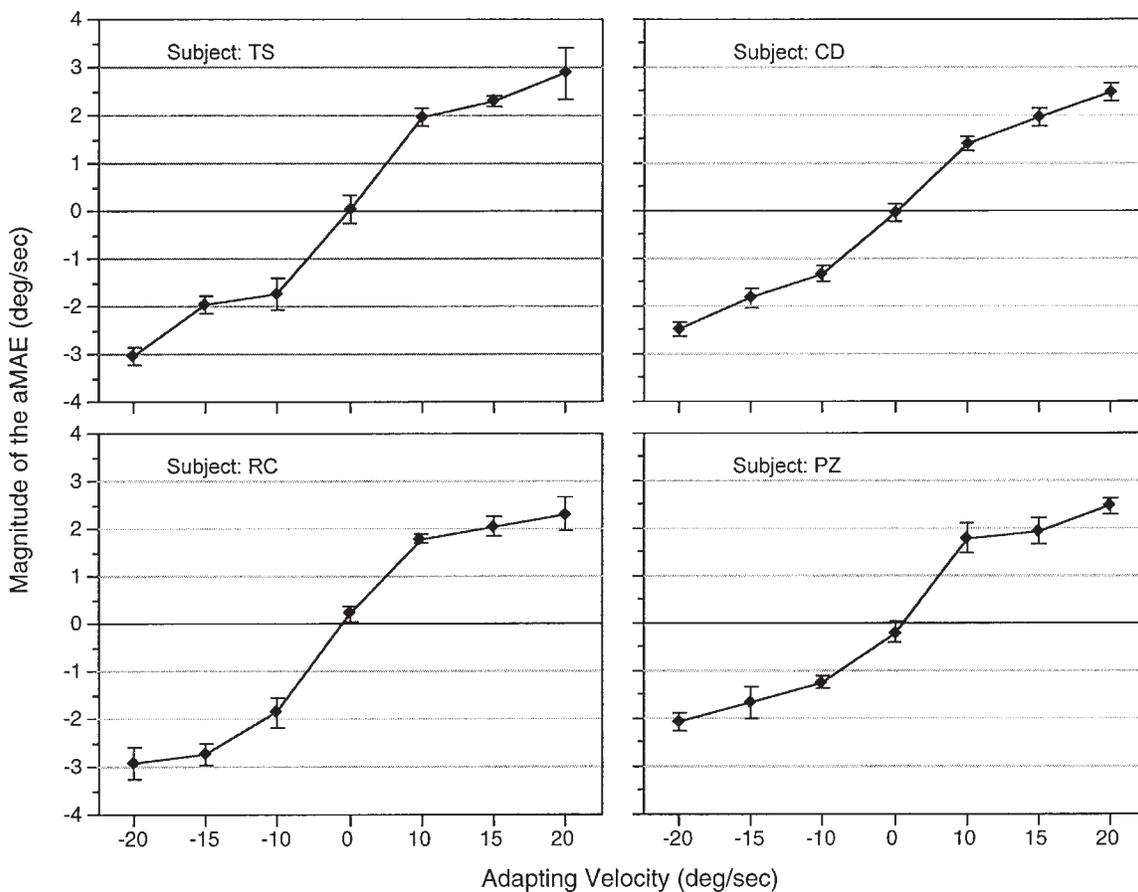


Figure 5. The magnitude of the auditory motion aftereffect (aMAE) measured as a function of adapting velocity. Different panels show the mean sizes of the aMAE for different subjects. The velocity of 0°/sec indicates the control condition in which the adapting sound was presented from a stationary loudspeaker directly in front of the subject (0.8 m distant). Note that there is a clear aMAE for all the subjects tested. The magnitude of the aMAE increased up to the highest velocity tested. The error bars indicate ± 1 standard errors in the means of 3–4 separate blocks of trials.

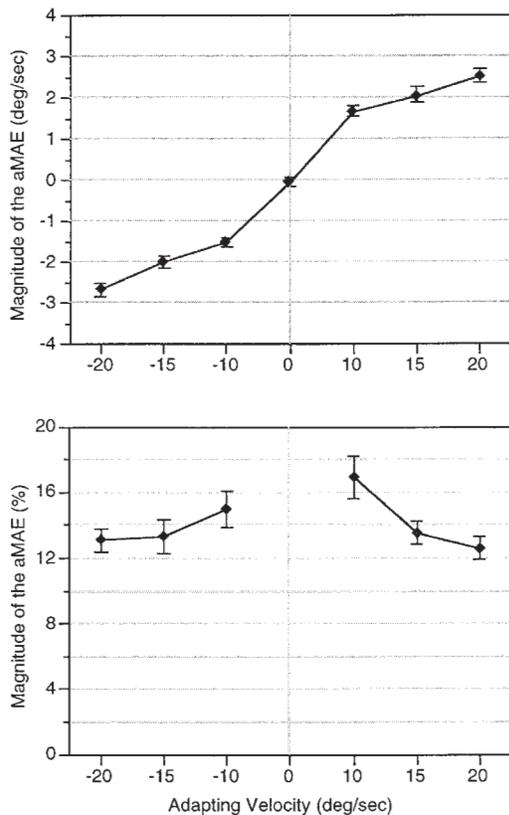


Figure 6. Grand average of the auditory motion aftereffect (aMAE) over all 4 subjects tested in Experiment 1 as a function of adapting velocity. The top panel shows the mean magnitude of the aMAE (the mean of the panels shown in Figure 4). The bottom panel displays the size of the aMAE as a percentage of the adapting velocity—that is, as the gain of the aMAE. The error bar indicates ± 1 standard error of the mean across 4 subjects. Note that although the magnitude of the aMAE increased with increases in adapting velocity, the gain of the aMAE tended to decrease as the adapting velocity increased.

EXPERIMENT 3

This experiment was designed to study the tuning and specificity of the aMAE in the frequency domain.

Method

Three subjects (M.L., J.Q., and C.D.) were tested. The adapting region, velocities, and start positions of the test stimulus and the stimulus intensity were the same as those in Experiment 1. As in Experiment 2, only one adapting velocity (20°/sec) was tested. In this experiment, instead of using broadband noise, five one-octave band-pass noises with different center frequencies (f_c s; i.e., 500, 1000, 2000, 4000, and 8000 Hz) were used as stimuli. As in Experiment 2, there were two conditions. In the first condition, in order to study the frequency tuning of the aMAE, the adapting and the test stimuli had the same spectrum. In the other condition, in order to investigate whether the aMAE is frequency specific, the adapting and the test stimuli were in different frequency spectral ranges. A total of 11 combinations of the adapting and the test stimuli were tested in these two conditions (i.e., 500/500, 1000/1000, 2000/2000, 4000/4000, 8000/8000, 500/2000, 500/8000, 2000/500, 2000/8000, 8000/500, and 8000/8000 Hz, where the first figure indicates the f_c

of the adapting stimulus and the second indicates that of the test stimulus). Owing to limited time available from each subject, each band-pass noise and each combination of adapting and test stimuli was tested only once. Thus, each point on the psychometric function was based on eight responses.

Results and Discussion

The results from Experiment 3 are shown in Figures 9 and 10. In Figure 9, the magnitude of the aMAE was plotted as a function of frequency band. The results showed that the magnitude of the aMAE was slightly larger for low-frequency (<1000 Hz) sound than for middle- (2000–4000 Hz) or high-frequency (>4000 Hz) sounds, which is consistent with the results reported by Grantham and Wightman (1979). Sound in the middle-frequency range (2000–4000 Hz) was least efficient in inducing the aMAE. The same pattern was shown in all 3 subjects tested. Figure 10 compares the aMAEs measured with nine combinations of the adapting and the test stimuli (i.e., 500/500, 500/2000, 500/8000, 2000/500, 2000/2000, 2000/8000, 8000/500, 8000/2000, and 8000/8000 Hz, where the first figure indicates the spectral content of the adapting stimulus and the second one indicates that of the test stimulus). In three panels (top right, bottom left, and bottom right, respectively) of Figure 10, low- (f_c , 500 Hz), middle (f_c , 2000 Hz), and high- (f_c , 8000 Hz) frequency sounds served as test stimuli. In these three panels, the adapting stimulus was a low- (f_c , 500 Hz), medium- (f_c , 2000 Hz), or high- (f_c , 8000 Hz) frequency sound. The results showed that the aMAE was largest when the adapting and the test stimuli were identical in frequency—that is, when both of them were low-, middle-, or high-frequency sound. When the adapting and the test stimuli were different in frequency, the magnitude of the aMAE decreased. The extent of the decrease depended on the difference in frequency between the adapting and the test stimuli. The larger the difference in frequency, the weaker the aMAE, suggesting that the aMAE is frequency specific. These results are summarized in the top left panel of Figure 10, in which the size of the circle represents the magnitude of the aMAE averaged over all 3 subjects tested in nine combinations of adapting and test stimuli. In addition to the frequency specificity of the aMAE, which is reflected by the larger sizes of the circles along the diagonal (lower left corner to upper right corner), the pattern in this panel shows an interesting asymmetry of the frequency distribution of the aftereffect. The circles below the diagonal are larger than those above it, indicating that the aMAE is stronger when the test stimulus is higher in frequency than the adapting frequency and is weaker in the opposite conditions.

GENERAL DISCUSSION

Speculations About Specialized Motion Detection Mechanisms

The existence of an MAE is often regarded as evidence for specialized motion detection mechanisms in a modality. An aftereffect caused by adaptation to auditory spec-

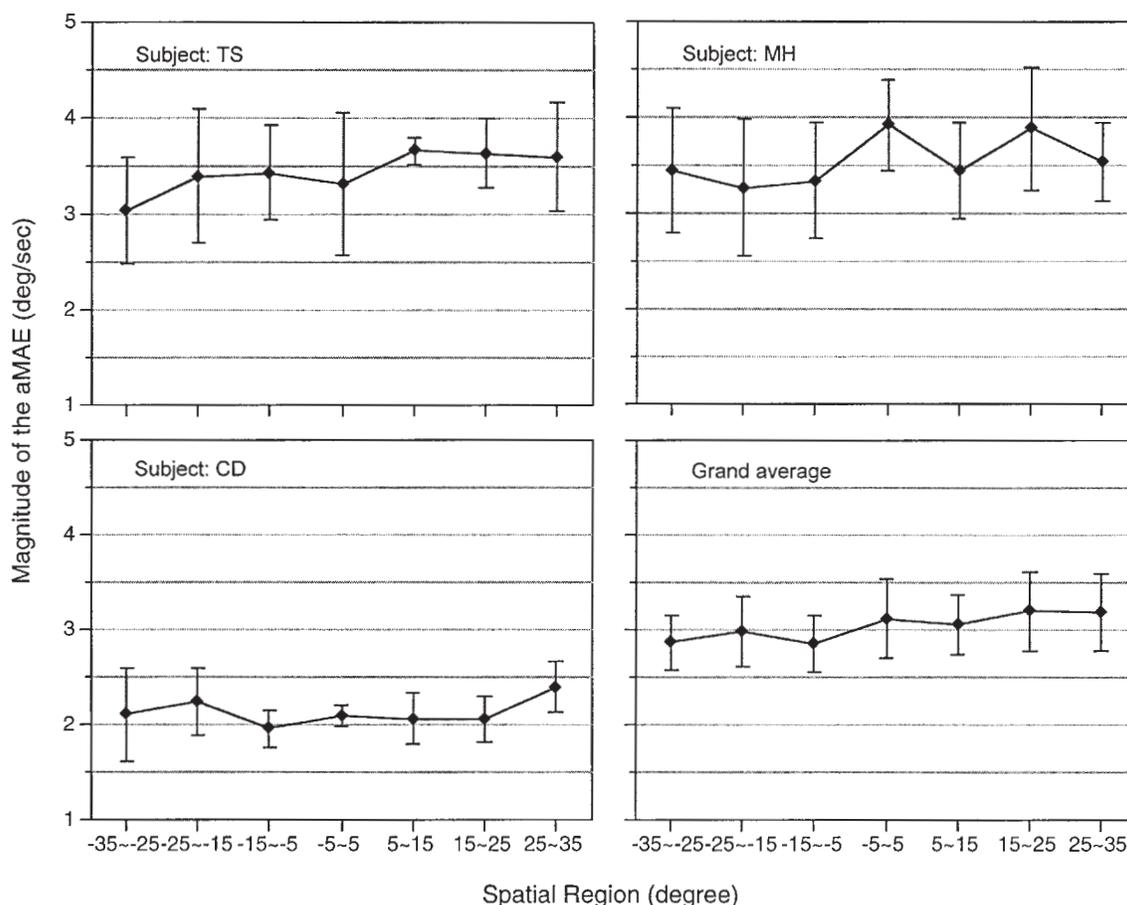


Figure 7. Spatial tuning of the auditory motion aftereffect (aMAE) for 3 subjects, and the average over all 3 subjects (Experiment 2, Condition 1). In this condition, the adapting and the test regions were always the same, and the magnitude of the aMAE was measured as a function of spatial region. Note that in this experimental condition, within the spatial region tested, the aMAE was largely invariant with azimuth. The error bars indicate ± 1 standard errors in the means of 3–4 separate blocks of trials.

tral motion was earlier demonstrated in our laboratory (Shu, Swindale, & Cynader, 1993). After listening to a simple spectral pattern (a spectral peak or a spectral notch) moving upward or downward in frequency for a few minutes, the same pattern was perceived as moving in the opposite direction even though it was actually stationary. This result suggests that there exist specialized channels in the auditory system that process dynamic spectral cues.

In the present study, by using real moving sound as both adapting and test stimuli, a robust aMAE was observed even with adapting velocities as low as $10^\circ/\text{sec}$. When the magnitude of the aMAE obtained in the present study is expressed as gain—that is, as a percentage of the adapting velocity—it is comparable with that of the vMAE. By using a nulling procedure, Taylor (1963) measured the velocity of the vMAE as a function of adapting velocity. In order to compare the MAEs between these two modalities, we recalculated the magnitude of the vMAE measured by Taylor in terms of gain. Consistent with our results for the aMAE, his results showed that although the

velocity of the vMAE increased with the adapting velocity, the gain of the vMAE expressed as a ratio with adapting velocity decreased from 23.3% to 4.8% when the adapting velocity increased from $9^\circ/\text{sec}$ to $108^\circ/\text{sec}$. Our results suggest that, like the visual system, the auditory system contains specialized mechanisms for auditory spatial motion detection.³

Comparison With Previous Studies of the Auditory Motion Aftereffect

Previous studies of the aMAE have used simulated moving sound as the stimulus. Simulated sound source motion has been generated by modulating IID (Ehrenstein, 1994; Reinhardt-Rutland, 1992), ITD (Ehrenstein, 1994), or both of these cues (Grantham, 1989; Grantham & Wightman, 1979). When the simulated moving stimulus used only one of these dynamic cues, no aMAE was elicited (Ehrenstein, 1994; Reinhardt-Rutland, 1992). Instead, a monaural changing-loudness aftereffect or a displacement aftereffect was observed. When both ITD

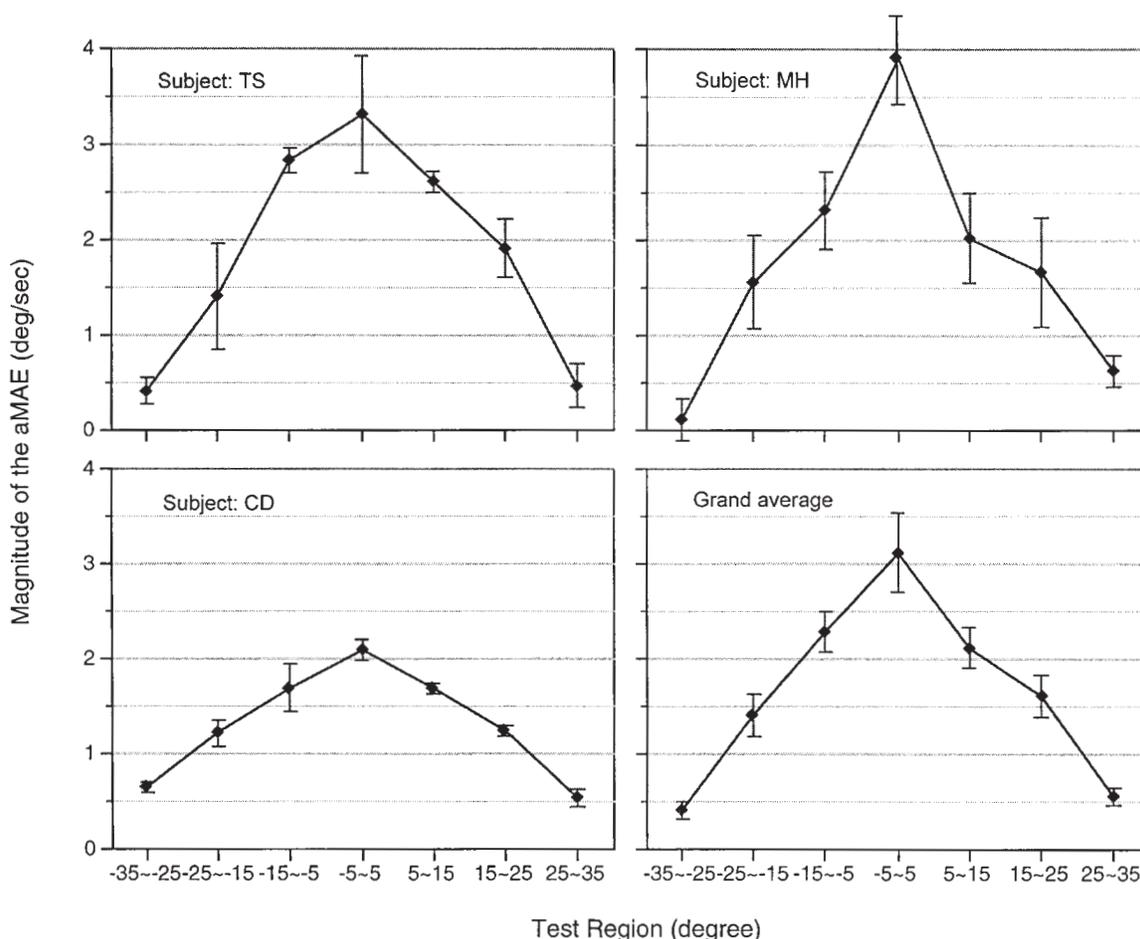


Figure 8. Magnitude of the auditory motion aftereffect (aMAE) in the case in which the adapting and the test regions were separated (Experiment 2, Condition 2). Data are shown for 3 individual subjects, as well as the average over all 3 subjects. Note that when the distance between the adapting and the test regions increased, the magnitude of the aMAE decreased. The error bars indicate standard errors in the means of 3–4 separate blocks of trials.

and IID cues were used (Grantham, 1989; Grantham & Wightman, 1979), an aMAE was found at low frequencies (500 Hz), but for only some of the subjects. In contrast, by using real moving sound as the stimulus, we observed an aMAE for all the subjects in our sample and for all the conditions tested, including relatively low adapting velocities and low and high frequencies. This suggests that the aMAE requires as many motion cues as possible. A real moving stimulus will provide monaural spectral cues, as well as temporal and level information. In order to have the aMAE explicitly expressed, not only do the auditory motion detection channels need to be adequately adapted, but it may also be necessary to use a stimulus rich in motion cues as a test.

More recently, Grantham (1998) published a study in which moving sound stimuli were first recorded with two microphones put in the ear canals of a KEMAR manikin. These recorded sounds were later played back to the subject through headphones during experiments. Thus, in

his experiments, sound waves reaching the subject's tympanic membrane consisted of the dynamic cues provided by the KEMAR manikin's head and pinnae, such as time-varying ITD, IID, and monaural spectral shape cues. Four of the 5 subjects tested in his study reported that the sounds were externalized, with a distance to their heads of less than 30 cm. But 1 subject perceived the stimuli to be inside her head. In the Grantham (1998) study, the magnitude of the aMAE was defined as a percentage of the maximum possible area between the two psychometric functions obtained when the adapting stimuli had the same velocity but moved in opposite directions. Using this kind of virtual moving sound as the stimulus, Grantham (1998) demonstrated an aMAE for both low- (1000 Hz low-pass filtered) and high- (5000–8000 Hz band-pass filtered) frequency signals within a range of velocities from 10°/sec to 180°/sec. No significant velocity effect was found on the magnitude of the aMAE. He also showed that the adaptor and the probe must have the

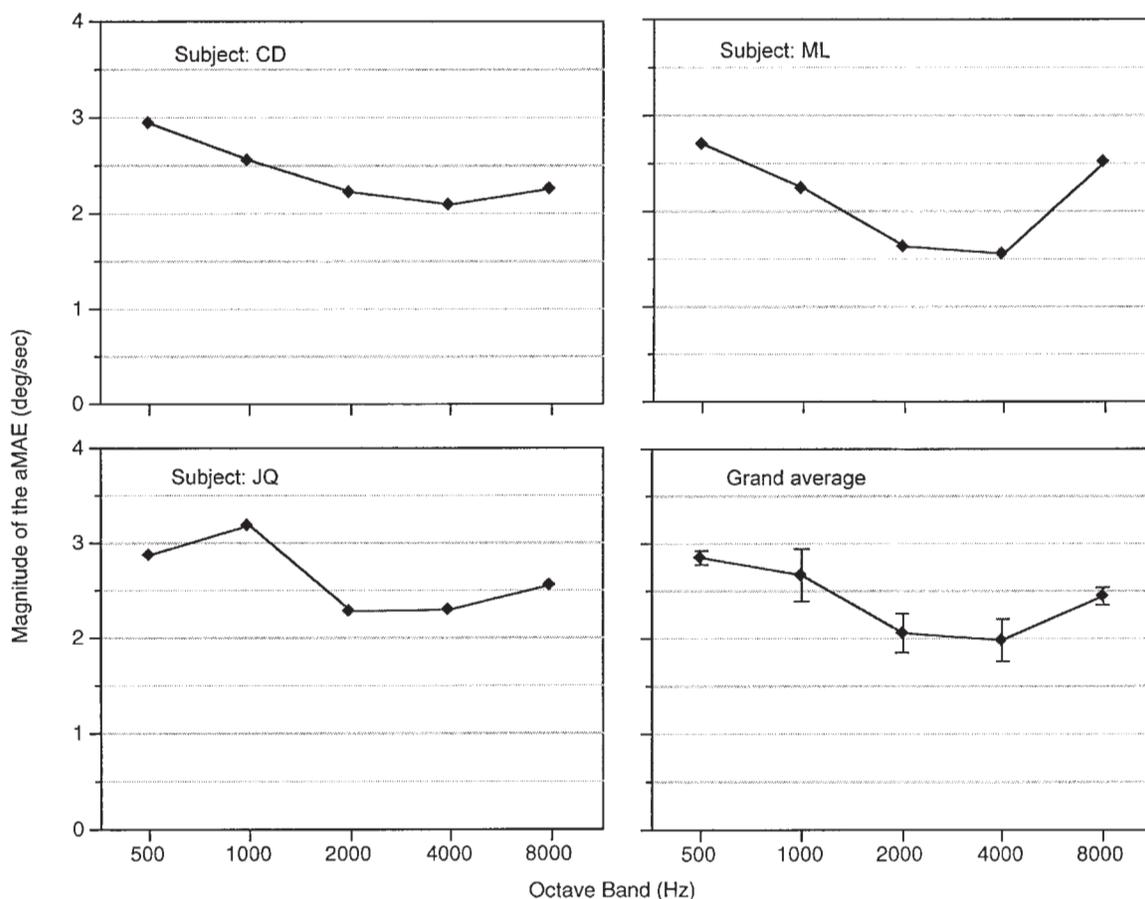


Figure 9. Magnitude of the auditory motion aftereffect (aMAE) as a function of frequency band in the case when the adapting and the test stimuli had the same spectrum (Experiment 3, Condition 1). Different panels show results for 3 individual subjects, as well as the average over all 3 subjects (lower right panel). Note that the aMAE was slightly larger for low-frequency sound than for high- and middle-frequency sounds. The middle-frequency stimulus was least efficient in inducing the aMAE. The error bars indicate standard errors of the means of 3–4 separate blocks of trials.

same spectrum and share the same azimuthal region in order to obtain a measurable aMAE, suggesting that the aMAE was frequency and spatially specific.

Since different methods were used to measure the aMAE in Grantham's (1998) and our studies, we cannot directly compare the strengths of the aMAEs observed. But in general, our results agree with Grantham's (1998) findings, in that, with sounds moving in the real world as stimuli, an aMAE could be elicited by low, middle, or high frequencies and the magnitude of the effect depended on the extent of the overlapping spatial region, as well as the spectral content of the adapting and test stimuli. In addition, our study extends the Grantham (1998) study by further investigating the tuning of the aMAE in the frequency and spatial domains. Our results (Figure 10) show that the aMAE is stronger when adapting stimuli contain lower frequency sound relative to that contained in the test stimuli than when the frequency relation between the adapting and the test stimuli is reversed. These results suggest that the frequency tuning of the motion-sensitive channels may spread further to the high end than to the

low end of the frequency spectrum. Our finding that, across the spatial range tested (from -35° to $+35^\circ$ azimuth), the aftereffects are largely invariant with eccentricity is consistent with research on minimum audible movement angle, which has shown that dynamic acuity is relatively constant throughout this range of azimuths (Chandler & Grantham, 1992; Grantham, 1986; Strybel, Manligas, & Perrott, 1992).

In contrast to the absence of a significant effect of adapting velocity reported by Grantham (1998), our results show a clear velocity dependence of the aMAE (Figures 5 and 6). Grantham (1998) showed that there was a considerable variability across subjects, to which his failure to find the velocity effect on the aMAE may be attributed. Taking into account the features of the virtual moving sound used by Grantham (1998), intersubject variation in his results should not be surprising. In his experiments, sound stimuli delivered to the subject over headphones were filtered by the KEMAR manikin's head-related transfer functions (HRTFs). However, studies have shown that the dimensions of the head and pinnae can vary

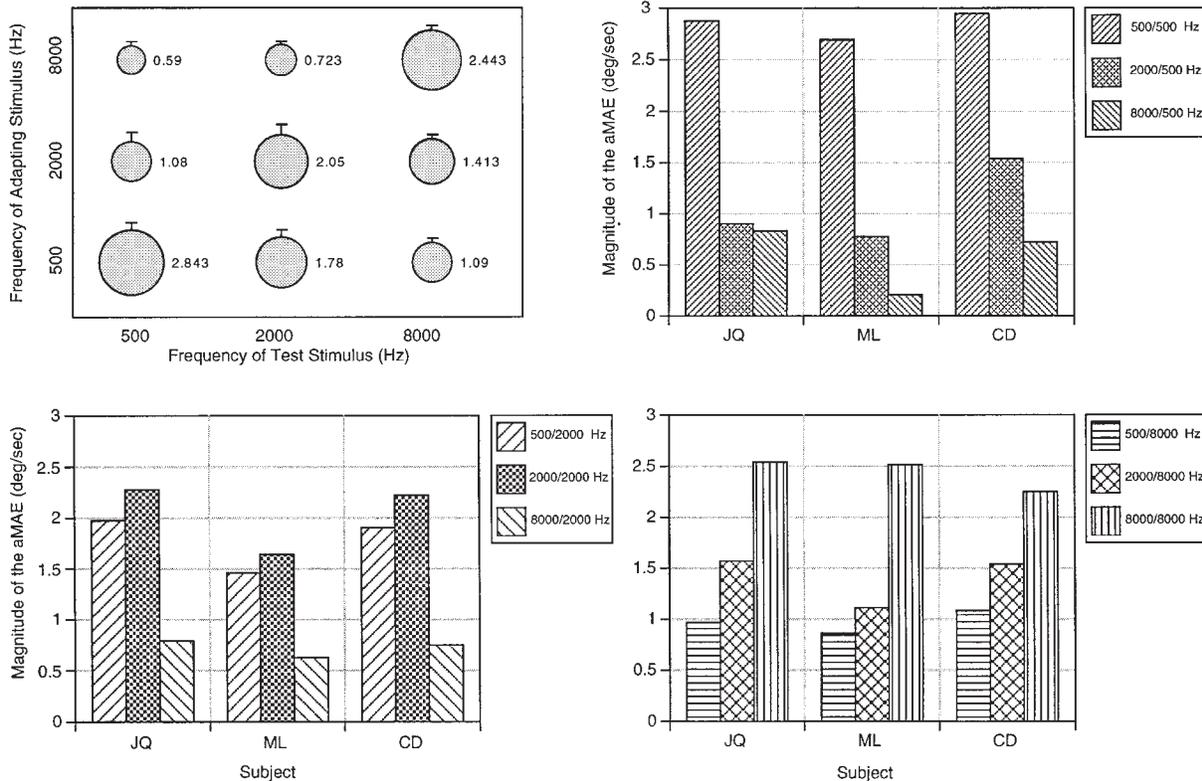


Figure 10. Frequency specificity of the auditory motion aftereffect (aMAE) for 3 subjects. In the top right, bottom left, and bottom right panels, the center frequency of the test stimulus is 500, 2000, and 8000 Hz, respectively. In each panel, the legend lists combinations of adapting and test stimuli, where the first figure indicates the center frequency of the adapting stimulus and the second that of the test stimulus. Note that the aMAE was largest when the adapting and the test stimuli were identical in frequency. When the adapting and the test stimuli were different in frequency, the magnitude of the aMAE decreased. The larger the difference in frequency, the weaker the aMAE. In the top left panel, the results for all 3 subjects are summarized. The size of the filled circle represents the magnitude of the aMAE averaged over all 3 subjects tested in nine combinations of adapting and test stimuli. The numbers next to the circles show the magnitude of the aftereffect for each combination. Note that the sizes of the circles below the diagonal (lower left corner to upper right corner) are larger than those above it, showing that the aMAEs for the combinations of 500/2000, 500/8000, and 2000/8000 Hz are stronger than those for the 2000/500-, 8000/500-, and 8000/2000-Hz combinations.

considerably among adults. Since these anatomical dimensions are roughly comparable with the wavelength of sounds in the range of human hearing, individual differences in anatomical dimensions transform into individual differences in potential localization cues (Burkhard & Sachs, 1975; Middlebrooks & Green, 1990; Shaw & Teranishi, 1968; Wenzel, Wightman, & Kistler, 1991). As a result, different individuals have different HRTFs. Some of the subjects tested in the Grantham (1998) study might have HRTFs similar to KEMAR's, whereas others might have quite distinct HRTFs and their motion detection pathways might be inadequately stimulated. Most of these drawbacks are overcome by using real moving sound as a stimulus. It should be noted, however, that the sounds produced by our sound stimulus system lack any significant Doppler shift, which might be a potent cue for detecting sound source motion in directions that are not circumferential to the head. Rosenblum, Carello, and Pastore (1987) showed that listeners could use the Doppler effect, as well as ITDs and amplitude changes, to lo-

cate a moving sound source, although the Doppler effect was the least effective of these cues.

Evidence for Hierarchical Organization for Auditory Motion Processing

A comparison of results from previous studies and the present study suggests that auditory motion processing channels respond optimally to a combination of several cues and that these channels may operate at more than one stage of analysis. In vision, this kind of psychophysical model has been proposed for information processing of motion in depth, in which monocularly driven changing-size channels and stereoscopic-motion channels, driven by interactions of signals from the two eyes, feed a motion-in-depth stage (Regan et al., 1979). In that study, an impression of motion in depth was generated by either changing the size of a retinal image or adjusting relative velocities of the left and right retinal images (stereoscopic-motion stimulus). Using a selective adaptation procedure, they demonstrated that there exist changing-size channels and

stereoscopic-motion channels in the visual system. Their finding that the impression of motion in depth elicited by changing-size stimuli can be canceled by appropriate stereoscopic-motion stimuli indicates that changing-size stimulation and stereoscopic-motion stimulation generate visual signals that converge on the same motion-in-depth stage.

Studies of the auditory aftereffects show that although adaptation to an auditory stimulus providing one cue (e.g., changing ITD or IID) produced only a loudness or a displacement aftereffect (Ehrenstein, 1994; Reinhardt-Rutland, 1992), adaptation to a moving sound containing two or more cues, including dynamic level, temporal, and spectral information, did elicit a repeatable aMAE. Thus, it is reasonable to conceive a similar hierarchical organization for auditory motion processing (Reinhardt-Rutland, 1992). That is, channels responsible for perceiving changes in sound level, spectrum, and temporal cues precede motion detection channels, and motion detection channels are fed by outputs from earlier analysis at these preceding channels. Simulated sound source movement generated by varying only one of these localization cues is unable to adequately stimulate motion detection channels and thus fails to elicit an aMAE. However, more peripheral channels (e.g., changing-level channels) may be sufficiently stimulated by such stimuli to produce an aftereffect (e.g., a loudness aftereffect).

Summary

In contrast to previous studies, we observed a robust and reliable aMAE when real moving sound was used as the adapting and test stimuli. Within the spatial region tested ($\pm 35^\circ$ centered on the midline), the aMAE was relatively independent of azimuth. The aMAE was larger for low-frequency sound than for middle- and high-frequency sounds, but it could be observed for both low- and high-frequency stimuli. The aMAE observed was spatially and frequency specific. These results suggest that the auditory system contains specialized motion detection channels and that the processing of motion information may proceed through a hierarchy of stages.

REFERENCES

- BARLOW, H. B., & HILL, R. M. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figure after-effects. *Nature*, **200**, 1345-1347.
- BEVERLEY, K. I., & REGAN, D. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *Journal of Physiology*, **235**, 17-29.
- BRIGNER, W. L. (1982). Spatial frequency selectivity of spiral aftereffect. *Perceptual & Motor Skills*, **55**, 1129-1130.
- BRIGNER, W. L. (1986). Is velocity of motion aftereffect proportional to velocity of induction? *Perceptual & Motor Skills*, **63**, 362.
- BURKHARD, M. D., & SACHS, R. M. (1975). Anthropometric manikin for acoustic research. *Journal of the Acoustical Society of America*, **58**, 214-222.
- CAMERON, E. L., BAKER, C. L., & BOULTON, J. C. (1992). Spatial frequency selective mechanisms underlying the motion aftereffect. *Vision Research*, **32**, 561-568.
- CHANDLER, D. W., & GRANTHAM, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *Journal of the Acoustical Society of America*, **91**, 1624-1636.
- EHRENSTEIN, W. H. (1994). Auditory aftereffects following simulated motion produced by varying interaural intensity or time. *Perception*, **23**, 1249-1255.
- FINNEY, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge: Cambridge University Press.
- GATES, L. W. (1934). The after-effect of visually observed movement. *American Journal of Psychology*, **46**, 34-46.
- GIASCHI, D., DOUGLAS, R., MARLIN, S., & CYNADER, M. S. (1993). The time course of direction-selective adaptation in simple and complex cells in cat striate cortex. *Journal of Neurophysiology*, **70**, 2024-2034.
- GRANTHAM, D. W. (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *Journal of the Acoustical Society of America*, **79**, 1939-1949.
- GRANTHAM, D. W. (1989). Motion aftereffects with horizontally moving sound sources in the free field. *Perception & Psychophysics*, **45**, 129-136.
- GRANTHAM, D. W. (1998). Auditory motion aftereffects in the horizontal plane: The effects of spectral region, spatial sector, and spatial richness. *Acustica/Acta Acustica*, **84**, 337-347.
- GRANTHAM, D. W., & WIGHTMAN, F. L. (1979). Auditory motion aftereffects. *Perception & Psychophysics*, **26**, 403-408.
- GRIFFITHS, T. D., REES, A., WITTON, C., SHAKIR, R. A., HENNING, G. B., & GREEN, G. G. (1996). Evidence for a sound movement area in the human cerebral cortex. *Nature*, **383**, 425-427.
- HERSCHENSON, M. (1989). Duration, time constant, and decay of the linear motion aftereffect as a function of inspection duration. *Perception & Psychophysics*, **45**, 251-257.
- HERSCHENSON, M. (1993). Linear and rotation aftereffects as a function of inspection duration. *Vision Research*, **33**, 1913-1919.
- HOLLAND, H. C. (1957). The Archimedes spiral. *Nature*, **179**, 432-433.
- LAPPIN, J. S., BELL, H. H., HARM, O. J., & KOTTAS, B. (1975). On the relation between time and space in the visual discrimination of velocity. *Journal of Experimental Psychology*, **1**, 383-394.
- MARLIN, S. G., HASAN, S. J., & CYNADER, M. S. (1988). Direction-selective adaptation in simple and complex cells in cat striate cortex. *Journal of Neurophysiology*, **59**, 1314-1330.
- MIDDLEBROOKS, J. C., & GREEN, D. M. (1990). Directional dependence of interaural envelope delays. *Journal of the Acoustical Society of America*, **87**, 2149-2162.
- MIDDLEBROOKS, J. C., & GREEN, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, **42**, 135-159.
- O'SHEA, R. P., & CRASSINI, B. (1981). Interocular transfer of the motion after-effect is not reduced by rivalry. *Vision Research*, **21**, 801-804.
- OVER, R., BROERSE, J., CRASSINI, B., & LOVEGROVE, W. (1973). Spatial determinants of the aftereffect of seen motion. *Vision Research*, **13**, 1681-1690.
- PANTLE, A. J., & SEKULER, R. W. (1968). Velocity-sensitive elements in human vision: Initial psychophysical evidence. *Vision Research*, **8**, 445-450.
- REGAN, D., & BEVERLEY, K. I. (1973). Disparity detectors in human depth perception: Evidence for directional selectivity. *Science*, **181**, 877-879.
- REGAN, D., BEVERLEY, K. I., & CYNADER, M. S. (1979). Stereoscopic subsystems for position in depth and for motion in depth. *Proceedings of the Royal Society of London: Series B*, **204**, 485-501.
- REINHARDT-RUTLAND, A. H. (1992). Changing-loudness aftereffect follow simulated movement: Implications for channel hypotheses concerning sound level change and movement. *Journal of General Psychology*, **119**, 113-121.
- ROSENBLUM, L. D., CARELLO, C., & PASTORE, R. E. (1987). Relative effectiveness of three stimulus variables for locating a moving sound source. *Perception*, **16**, 175-186.
- SAUL, A. B., & CYNADER, M. S. (1989a). Adaptation in single units in visual cortex: The tuning of aftereffects in the spatial domain. *Visual Neuroscience*, **2**, 593-607.
- SAUL, A. B., & CYNADER, M. S. (1989b). Adaptation in single units in visual cortex: The tuning of aftereffects in the temporal domain. *Visual Neuroscience*, **2**, 609-620.
- SEKULER, R. W., & GANZ, L. (1963). Aftereffect of seen motion with a stabilized retinal image. *Science*, **139**, 419-420.

- SEKULER, R. W., & PANTLE, A. J. (1967). A model of after-effects of seen movement. *Vision Research*, **88**, 1-11.
- SHAW, E. G., & TERANISHI, R. (1968). Sound pressure generated in an external ear replica and real human ears by a nearby sound source. *Journal of the Acoustical Society of America*, **44**, 240-249.
- SHU, Z. J., SWINDALE, N. V., & CYNADER, M. S. (1993). Spectral motion produces an auditory after-effect. *Nature*, **364**, 721-723.
- STRYBEL, T. Z., MANLIGAS, C. L., & PERROTT, D. R. (1992). Minimum audible movement angle as a function of the azimuth and elevation of the source. *Human Factors*, **34**, 265-275.
- TAYLOR, M. M. (1963). Tracking the decay of the after-effect of seen rotary movement. *Perceptual & Motor Skills*, **16**, 119-129.
- WADE, N. J. (1994). A selective history of the study of visual motion aftereffects. *Perception*, **23**, 1111-1134.
- WENZEL, E. M., WIGHTMAN, F. L., & KISTLER, D. J. (1991). Localization with non-individualized virtual acoustic display cues. In S. B. Robertson, G. M. Olson, & J. S. Olson (Eds.), *Proceedings of CHI'91* (pp. 351-359). New York: ACM Press.
- WOHLGEMUTH, A. (1911). On the after-effect of seen movement. *British Journal of Psychology*, **1**, (Monograph Suppl. 1), 1-117.

NOTES

1. Shortly before this paper was submitted, Grantham (1998) published another, independent study of the aMAE. We include discussions of the results of his recent work in relation to ours in the General Discussion section of this paper.
2. A control experiment showed that the noise produced by the movement of the robot arm did not provide a direction cue. In this experiment, the loudspeaker was turned off, and the arm moved randomly at one of eight velocities (1°/sec, 3°/sec, 5°/sec, and 7°/sec, either leftward or rightward) with one of eight start positions (1°, 3°, 5°, and 7°, either to the left or the right of the midline). Therefore, there was a total of 64 trials (8 velocities × 8 start positions) in the test. This test condition is identical to that used in our adaptation experiments, except for the absence of any loudspeaker-generated sound. We found that for each test velocity, the subject's response was at chance level—that is, the percentage of *left* or *right* responses was about 50%.
3. This notion gains further support from a neurological study (Griffiths et al., 1996), which described a patient who had a specific deficit in auditory motion detection, although his ability to localize stationary sounds was much less impaired. An fMRI examination of this patient

revealed damage in a cortical area that was distinct from the primary auditory cortex, suggesting that moving sound may be processed independently from stationary sound.

APPENDIX

The design of the robot arm was governed by the following requirements: maximum coverage of space, low noise, high speed and acceleration, high structural resonant modes, safety, low bulk, and of course, low complexity and cost. The resulting design used a five-bar closed-loop spherical mechanism in which all five joint axes meet at one end point, where the loudspeaker is located. It is a property of these mechanisms that all points of their links are constrained to move on the surfaces of concentric spheres. This means that no part of the mechanism can penetrate the space in which the subject is located. If the actuators have coinciding axes, the workspace can cover the entire sphere, except the antipodes. In our system the two actuated joints were placed 20° apart, thereby approaching the *coinciding* condition and reducing bulk. The angular design of the other links was optimized to maximize the region within which the system preserves high-acceleration capabilities. This allowed about 63% of the surface of the sphere to be covered. The device was constructed of aluminum and achieved fairly high structural resonance despite the long reach. One link, the base link, was mechanically grounded and supported two actuated and instrumented joints. The other three joints were free. The base link was supported by a rigid overhead gantry and was located behind and above the subject's head. It included two soundproof motion reduction boxes driven by DC motors via elastomeric belts, with a reduction of about 1:60. This maximized acceleration capability (roughly proportional to the square root of the ratio of the mechanism's inertia to that of the motors). Computer control of the joint servo motors allowed arbitrary trajectories to be generated within the arm's workspace.

(Manuscript received September 2, 1998;
revision accepted for publication July 19, 1999.)