

Orientation dependence of spatial memory acquired from auditory experience

NAOHIDE YAMAMOTO AND AMY L. SHELTON
Johns Hopkins University, Baltimore, Maryland

The present study investigated whether memory for a room-sized spatial layout learned through auditory localization of sounds exhibits orientation dependence similar to that observed for spatial memory acquired from stationary viewing of the environment. Participants learned spatial layouts by viewing objects or localizing sounds and then performed judgments of relative direction among remembered locations. The results showed that direction judgments following auditory learning were performed most accurately at a particular orientation in the same way as were those following visual learning, indicating that auditorily encoded spatial memory is orientation dependent. In combination with previous findings that spatial memories derived from haptic and proprioceptive experiences are also orientation dependent, the present finding suggests that orientation dependence is a general functional property of human spatial memory independent of learning modality.

In everyday environments, spatial information about object locations is encoded in memory through multiple modalities: We can see where objects are located in space (vision), localize sounds (audition), and register direction and distance between objects through manual exploration (haptics) and ambulatory movements (proprioception, broadly defined; see Yamamoto & Shelton, 2005). Therefore, in order to attain a comprehensive understanding of human spatial cognition, it is critical to examine similarities and differences among spatial memories learned through different modalities. The primary objective of the present study was to investigate how a spatial layout is encoded and represented in memory through auditory experiences of object locations and, more specifically, to explore whether spatial memory derived from audition shows orientation dependence.

Spatial memory is considered to be orientation dependent when memory of an environment is more accessible from one perspective than from other perspectives, as indicated by faster and/or more accurate retrieval at the preferred orientation. Orientation dependence has been a focus of spatial cognition research because it provides important clues as to specific memory codes with which spatial information is represented in the brain. That is, if spatial memory is dependent on a single preferred orientation, this suggests that the environment is represented explicitly in only one particular orientation, whereas other orientations must be inferred by mentally transforming the spatial information represented in the preferred orientation. Therefore, it is of theoretical importance to investigate whether and how particular types of spatial memory show orientation dependence.

Numerous studies have investigated how different ways of learning a spatial layout affect the selection of a par-

ticular preferred orientation (for a review, see McNamara, 2003). Converging evidence has suggested that spatial memory generally is dependent on the most salient orientation. The salience of an orientation can be determined by a variety of factors, including the alignment of an observer with environmental axes, such as walls; the geometric structure of the layout; and the presence of conspicuous distal landmarks. However, when these cues are absent during spatial learning, an egocentric perspective from which the spatial layout is learned provides the most salient orientation (Shelton & McNamara, 2001). Although much of the research concerning orientation dependence has employed visual spatial learning, recent evidence has suggested that spatial memory acquired solely through haptic learning (Newell, Woods, Mernagh, & Bühlhoff, 2005) or proprioceptive learning (Yamamoto & Shelton, 2005) also exhibits orientation dependence.

These converging results from visual, haptic, and proprioceptive spatial learning lead to the functional equivalence hypothesis, which claims that memory representations of space function equivalently, regardless of the modality through which spatial information is encoded (Avraamides, Loomis, Klatzky, & Golledge, 2004; Loomis, Lippa, Klatzky, & Golledge, 2002). In the context of orientation dependence research, a more specific hypothesis derived from this functional equivalence hypothesis is that the property of orientation dependence is common to all spatial memories, irrespective of learning modalities. However, this hypothesis lacks one last piece of evidence: It is yet unknown whether spatial memory learned from audition alone is also orientation dependent. Previous work on auditory spatial learning has focused on the encoding of a self-to-object relation between an observer

N. Yamamoto, nyama@gwu.edu

and a single sound source, leaving auditory memory for spatial layout largely untouched. For example, Loomis, Klatzky, Philbeck, and Golledge (1998) and Klatzky, Lippa, Loomis, and Golledge (2002) presented multiple auditory targets sequentially in different locations to stationary observers and had them indicate egocentric distances (Loomis et al., 1998) or directions (Klatzky et al., 2002) of sound sources individually for each target. Thus, the previous studies did not investigate how spatial relations among sound locations are encoded and represented in memory. Considering that just one counterexample (i.e., auditorily encoded spatial memory showing orientation independence) could falsify the hypothesis about the generality of orientation dependence across modalities, it is important to complete such an investigation by examining explicitly whether spatial memory learned through audition is also orientation dependent.

Therefore, the present study was designed to explore whether auditorily acquired memory for spatial layout would show orientation dependence. To achieve this goal, identical pure tones originating from different locations in a room were presented sequentially to stationary participants, who were asked to learn the layout of sound locations through auditory localization of each sound source. Subsequently, orientation dependence of their spatial memories was assessed by having the participants make judgments of relative direction among remembered sound locations. In addition, they learned another layout via stationary viewing, to allow for a direct comparison with the orientation dependence observed for vision. The functional equivalence hypothesis predicts that spatial memory from audition is also orientation dependent.

METHOD

Participants

Sixteen participants (8 males, 8 females, 18–22 years of age) from the Johns Hopkins community volunteered in return for extra credit in psychology courses. All participants reported normal or corrected-to-normal vision and normal hearing.

Materials and Stimuli

Two spatial layouts were constructed from two different sets of six objects (see Figure 1 for an example). The objects were common, were visually distinct, were similar in size (approximately 15 cm in length and width, 10 cm in height), had monosyllabic names, and shared no primary semantic associations. For visual learning, these objects were placed on the floor of an approximately 3 × 3.7 m room. For auditory learning, each object was replaced with an identical pair of box-shaped speakers that were aligned with the 0°–180° axis in Figure 1, with their sound-emitting surfaces facing toward the participant's position in the room. The pair of speakers was about the size of the objects: 9 cm long, 17 cm wide, and 10 cm high. In both learning conditions, the floor was covered with a plain white sheet. During visual learning, a curtain made of a white canvas created a uniform texture around the walls of the room.

Auditory stimuli were composed of 440-Hz pure tones, each of which lasted for 500 msec, and object labels spoken by a female voice. The duration of each object label ranged from 333 to 680 msec. From each object location, 10 successive pure tones and a corresponding object label inserted between the 5th and 6th tones were presented through speakers. Each pure tone was separated from other tones or object labels by 500 msec (see Figure 2). All of the pairs of speakers were connected to a computer simul-

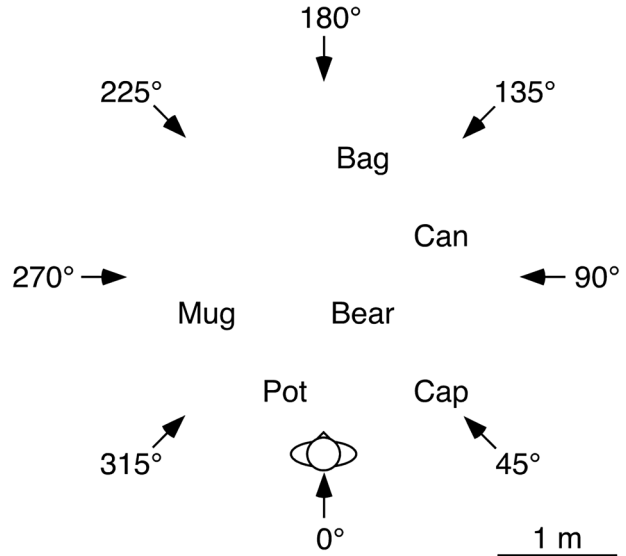


Figure 1. One of the spatial layouts used in the experiments. Real objects were used for visual learning. For auditory learning, a pair of speakers aligned with the 0°–180° axis was placed at each object location, with sound-emitting surfaces facing toward the participant. The arrow labeled as 0° indicates the participant's position and orientation during learning. The arrows labeled 0°, 90°, 180°, and 270° indicate locations from which reference sounds were presented in the auditory learning condition.

taneously via a switch that allowed switching among the speaker pairs so that the auditory stimuli for a given object were presented through the appropriate speakers. All pure tones had a sound level of 90 dB(A) measured at the speakers. The sound level of the object labels was 85 dB(A) on average, measured in the same manner. The baseline sound level gauged in the center of the room was approximately 50 dB(A). This was caused by constant ambient noise, and no localizable sounds were present except experimental stimuli during auditory learning.

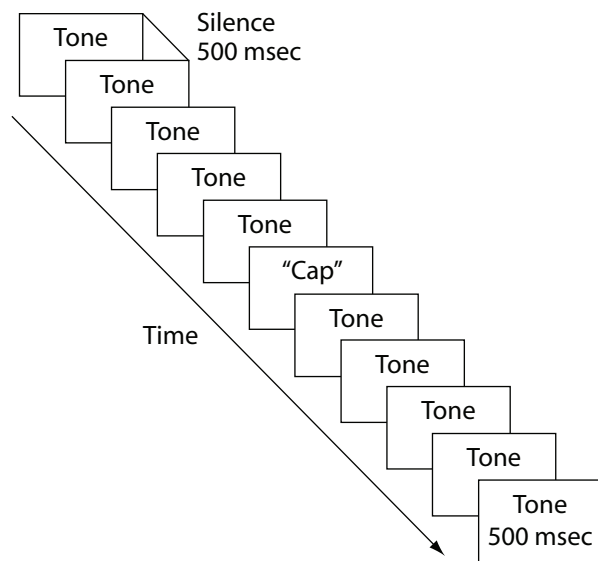


Figure 2. Timeline of the auditory stimulus presentation for a given object. Total presentation time for each object ranged from 10.33 to 10.68 sec.

Design and Procedure

Each participant learned both layouts, one through vision and one through audition. Each layout was used equally often in visual and auditory learning conditions across participants. Half the participants (4 males, 4 females) did visual learning first, and the other half did auditory learning first.

Learning phase. Before entering the room in which the layout was presented, the participant was shown the actual objects that made up the layout and was told that all objects would be placed on the floor. Then the participant sat in a wheelchair, put on a blindfold, and wore white-noise headphones that blocked auditory cues. An experimenter wheeled the participant to a position from which he/she learned the layout. (In Figure 1, the arrow labeled as 0° indicates this position.) The participant was stationary at this position throughout the learning phase. While being taken to this position, the participant was spun gently in the chair and disoriented. This disorienting procedure was included to ensure that the participant used only spatial information acquired in the room for learning the spatial layout. The participant was then asked to get up from the chair.

In the visual learning condition, the participant was first asked to remove the blindfold, and then objects were presented sequentially in a random order. Each object was presented alone for 10 sec. Only eye and head movements were allowed during viewing. To control for viewing time, the participant was asked to close his/her eyes while the experimenter replaced the just-viewed object on the floor with a new object in a new location. After viewing the last object, the participant was blindfolded and was asked to indicate both direction and distance to each of the objects by pointing and naming them in an arbitrary order. This study–test sequence was repeated until the participant fluently pointed to the correct object locations twice in a row. (Accuracy and fluency were determined by visual inspection of pointing performance.) On average, this criterion was met after the participant had viewed the objects two to three times. Upon completion of the learning phase, the participant was asked to sit in the chair again and was disoriented while being wheeled out of the room.

In the auditory learning condition, the participant remained blindfolded throughout the learning phase. After all pairs of speakers used to present sound stimuli were placed on the floor, the participant was asked to remove the white-noise headphones. This kept the participant from hearing any sounds made when the speakers were moved, which ensured that no adventitious auditory cues were available for localizing stimulus locations. Prior to the learning phase, the participant was informed that all pure tones originating from various locations in the room would be identical in their identity (i.e., they were in the same frequency) and intensity at the source. The participant was also instructed to focus on the pure tones, not object labels, for the purpose of localization. In the beginning of the learning phase, the sound stimuli were presented one at a time from four reference locations in the room, and the participant was verbally informed of the specific locations: at the participant's feet, the right wall, the far end of the room, and the left wall. The participant was told in advance that speakers emanating these reference sounds were placed on the floor. It was also made explicit to the participant that these reference locations defined a rectangular area within which all objects would be presented. In Figure 1, the locations of these speakers are indicated by arrows labeled as 0°, 90°, 180°, and 270°, respectively. These reference sounds were presented to provide the participant with information about the spatial extent of the room. In addition, the one presented at the participant's feet specified the sound intensity at the source. The sound stimuli were then presented from each object location sequentially in a random order, one at a time. The participant was encouraged to move his/her head while listening to the sound in order to enhance auditory localization, but no body movement was allowed. Following the presentation of the last sound stimulus, the participant was asked to point to and name the objects in the same manner as in the visual learning condition. This study–test sequence was repeated until the participant met a criterion that was the same as that described above, which took place after the sounds had been heard five to six times on average. Then

the participant was asked to wear the white-noise headphones and was disoriented while being wheeled out of the room.

Test phase. After learning each layout, the participant performed judgments of relative direction (JRDs) in another room. Overall, the JRD task was designed and implemented in the same way as in previous studies (e.g., Shelton & McNamara, 2001). In each trial, the participant was asked to imagine standing at one object and facing another object and then to point to a third (e.g., "Imagine you are at the cap facing the pot. Point to the mug."). The first two objects constituted an imagined heading. The third was a target to which the participant was to point. Each layout yielded 16 imagined headings, two instances of each of the eight orientations differing by 45°. These orientations were labeled counterclockwise from 0° to 315° in 45° steps, with 0° corresponding to the orientation experienced by the participant (see Figure 1). Target objects were chosen so that their directions were varied systematically; the egocentric space defined by each imagined heading was divided into four 90°-wide regions, each centered on the front–back and left–right axes, and the imagined headings at each orientation had approximately equal instances of target directions in each of these four regions.

Trials were presented on a computer screen. After receiving instructions about the task and about how to use the computer program, the participant performed three practice trials involving familiar buildings on campus. In each trial, sentences giving an imagined heading and a target were displayed with a circle and a movable line. The participant positioned the line by using a mouse so that it pointed to the target with respect to the imagined heading, which was represented by the 12 o'clock position of the circle. An example of the display is available in Fields and Shelton (2006, Figure 3). A total of 64 trials were presented in a random order. Note that not all possible permutations of imagined location, headings, and targets were used, because target directions were counterbalanced, as described above.

The principal dependent variable was accuracy in JRDs, measured by absolute angular distance between the pointed and target directions (i.e., absolute angular error). Trials were not time limited, and instructions stressed accuracy. Response latencies were also measured, in order to check for speed–accuracy trade-offs.

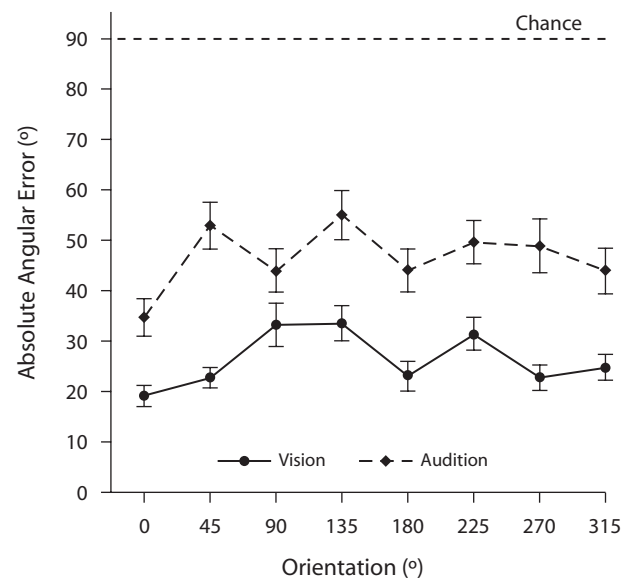


Figure 3. Mean absolute angular errors in judgments of relative direction (JRDs) as a function of learning modality and orientation. Error bars represent ± 1 standard errors of the means. Because absolute errors can range from 0° to 180° in the JRD task, the mean error expected by chance is 90° (indicated by the horizontal dashed line).

RESULTS

Two data points from 2 participants were excluded from the following analyses because very short response latencies (59 and 179 msec) indicated that JRDs were not performed in these trials. First, correlation coefficients between absolute angular errors and response latencies were computed for each participant. They ranged from $-.066$ to $.20$, indicating that none of the participants showed significant speed-accuracy trade-offs. Mean absolute angular errors were then calculated for each participant and for each condition and were subjected to a split-plot factorial ANOVA with order of learning (visual learning first, auditory learning first) and gender (male, female) as between-subjects factors and learning modality (vision, audition), orientation (from 0° to 315° with 45° intervals) and target direction (front, right, left, back) as within-subjects factors. All statistical tests reported below were corrected for nonsphericity when appropriate.

Figure 3 shows mean absolute angular errors in JRDs collapsed across participants as a function of learning modality and orientation. As is shown in this figure, JRDs were most accurate when the imagined headings were aligned with the learned orientation (0°), indicating that spatial memories acquired from vision and audition were orientation dependent. This observation was supported statistically by the significant main effect of orientation [$F(7,84) = 3.88, p < .008$] and planned contrasts comparing the 0° orientation with all others (45° – 315°) within each modality condition [visual learning, $F(1,15) = 13.32, p < .002$; auditory learning, $F(1,15) = 16.30, p < .002$]. Moreover, the interaction between learning modality and orientation was not significant [$F(7,84) = 1.67, p > .17$], suggesting that the effect of orientation was similar in each modality of spatial learning. Another finding from the present experiment was that JRDs were more accurately performed after visual learning than after auditory learning [$F(1,12) = 25.25, p < .001$].

The ANOVA also revealed that the main effect of target direction was significant [$F(3,36) = 16.86, p < .001$]. Mean absolute angular errors and corresponding standard errors of the means for each target direction were $M = 29.08^\circ$, $SEM = 1.77^\circ$ (front); $M = 35.71^\circ$, $SEM = 1.80^\circ$ (right); $M = 33.55^\circ$, $SEM = 1.91^\circ$ (left); and $M = 47.88^\circ$, $SEM = 2.41^\circ$ (back). As in previous studies (e.g., Shelton & McNamara, 2001), participants were more accurate in pointing to targets in front than in pointing to those on sides (both right and left), which were pointed to more accurately than those in back. Post hoc contrasts (by Scheffé's method) corresponding to these comparisons were significant [$F(1,15) = 8.62, p < .05$, front vs. sides; $F(1,15) = 17.80, p < .003$, sides vs. back]. The effect of target direction did not alter the conclusions regarding the effects of learning modality and orientation stated in the previous paragraph. Any main effects and interactions not mentioned above did not reach statistical significance ($\alpha = .05$).

DISCUSSION

The present experiment demonstrated clearly that memory for a spatial layout derived from auditory localization

of multiple sound sources is dependent on a learned orientation in the same manner as is spatial memory acquired from stationary viewing of the environment. Together with previous findings that spatial memory is orientation dependent when it is learned from visual (Shelton & McNamara, 2001), haptic (Newell et al., 2005), and proprioceptive (Yamamoto & Shelton, 2005) experiences, the present finding suggests that orientation dependence is a general property of human spatial memory that is independent of learning modality.

It is worth noting that the present finding provides additional support for the functional equivalence hypothesis (Avraamides et al., 2004; Loomis et al., 2002). To date, functional equivalence has been found for spatial representations constructed from direct sensory experiences in haptics (Newell et al., 2005) and proprioception (Yamamoto & Shelton, 2005) as well as indirect nonsensory sources of spatial information, including maps (Richardson, Montello, & Hegarty, 1999) and spatial language (Avraamides et al., 2004; Loomis et al., 2002). It has also been demonstrated that memories of a single spatial location encoded through vision and audition elicit comparable performance on distance and direction judgment tasks (Klatzky et al., 2002; Loomis et al., 1998). The present study reveals that visual and auditory spatial memories also mediate functionally equivalent performance on the JRD task, extending the notion of functional equivalence to auditorily acquired memory for spatial layout.

The present experiment also shows a clear difference between visual and auditory learning: Auditory learning yielded less accurate JRD performance overall than did visual learning. Given that auditory space perception, especially distance perception, tends to be relatively imprecise in sound localization (Zahorik, Brungart, & Bronkhorst, 2005), one possibility suggested by the decreased accuracy in JRDs is that spatial layouts were not learned well enough through auditory learning. That is, even though participants achieved the same accuracy in pointing to each object in the visual and auditory learning conditions, it is possible that this criterion was not stringent enough and that object locations were still represented less precisely in memory following auditory learning than following visual learning. In fact, when using auditory learning, participants required more exposure to the object locations in order to learn the layout to criterion, which suggests that even more exposure might have been needed for them to attain the same level of memory precision as in visually derived spatial memory. Alternatively, it is also possible that although object locations were registered in memory with the same accuracy and precision in both conditions at the time of initial encoding, the less precise nature of auditory distance perception might manifest itself in the course of memory consolidation processes. These possibilities aside, an important point here is that the finding of orientation dependence of auditorily acquired spatial memory is not compromised by the overall accuracy difference. On the contrary, if anything, it could even indicate the robustness of the orientation dependence. Furthermore, as has been pointed out by Avraamides et al. (2004), the functional equivalence

hypothesis does not postulate that spatial representations of comparable quality should be formed from any modality with equal ease. As such, the lower accuracy in JRDs following auditory learning did not make any alterations to the conclusion of the present study. Scrutinizing the differences in JRD accuracy between visual and auditory learning is a challenge for future studies, which will lead to better understanding of auditory spatial learning and memory.¹

In conclusion, the present study has shown that auditorily derived memory for spatial layout is orientation dependent in the same manner as are spatial memories constructed from visual, haptic, or proprioceptive experiences. This finding provides one missing piece of the puzzle to the literature, corroborating the claim that orientation dependence is a fundamental property universal to all spatial memories acquired through any modality.

AUTHOR NOTE

Portions of this work were presented at the 113th annual convention of the American Psychological Association in Washington, DC (August 2005). N.Y. was supported by a graduate fellowship from the Nakajima Foundation. We thank Eric Fortune, Chen Fu, and Nisha Jambulingam for their help in stimulus development, and Steven Marchette and anonymous reviewers for comments on earlier versions of the manuscript. Correspondence should be addressed to N. Yamamoto, Department of Psychology, George Washington University, 2125 G Street NW, Washington, DC 20052 (e-mail: nyama@gwu.edu).

REFERENCES

- AVRAAMIDES, M. N., LOOMIS, J. M., KLATZKY, R. L., & GOLLEDGE, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: Evidence from allocentric judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 801-814. doi:10.1037/0278-7393.30.4.804
- FIELDS, A. W., & SHELTON, A. L. (2006). Individual skill differences and large-scale environmental learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 506-515. doi:10.1037/0278-7393.32.3.506
- KLATZKY, R. L., LIPPA, Y., LOOMIS, J. M., & GOLLEDGE, R. G. (2002). Learning directions of objects specified by vision, spatial audition, or auditory spatial language. *Learning & Memory*, **9**, 364-367. doi:10.1101/lm.51702
- LOOMIS, J. M., KLATZKY, R. L., PHILBECK, J. W., & GOLLEDGE, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, **60**, 966-980.
- LOOMIS, J. M., LIPPA, Y., KLATZKY, R. L., & GOLLEDGE, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 335-345. doi:10.1037/0278-7393.28.2.335
- MCNAMARA, T. P. (2003). How are the locations of objects in the environment represented in memory? In C. Freksa, W. Brauer, C. Habel, & K. Wender (Eds.), *Spatial cognition: III. Routes and navigation, human memory and learning, spatial representation and spatial reasoning* (pp. 174-191). Berlin: Springer.
- NEWELL, F. N., WOODS, A. T., MERNAGH, M., & BÜLTHOFF, H. H. (2005). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research*, **161**, 233-242. doi:10.1007/s00221-004-2067-y
- RICHARDSON, A. E., MONTELLO, D. R., & HEGARTY, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, **27**, 741-750.
- SHELTON, A. L., & MCNAMARA, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, **43**, 274-310. doi:10.1006/cogp.2001.0758
- YAMAMOTO, N., & SHELTON, A. L. (2005). Visual and proprioceptive representations in spatial memory. *Memory & Cognition*, **33**, 140-150.
- ZAHORIK, P., BRUNGART, D. S., & BRONKHORST, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica United With Acustica*, **91**, 409-420.

NOTE

1. To this end, we conducted another experiment in which pure tones were replaced with bursts of broadband noise. Results from this experiment were identical to those reported in this article, suggesting that the present findings were not due to specific acoustic properties of pure tone stimuli.

(Manuscript received June 7, 2008;
revision accepted for publication October 7, 2008.)