Neuronal representations of distance in human auditory cortex

Norbert Kopčo[a,b,c,1], Samantha Huang[a], John W. Belliveau[a,d], Tommi Raij[a], Chinmayi Tengshe[a], and Jyrki Ahveninen[a]

[a]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Harvard Medical School/Massachusetts General Hospital, Charlestown, MA 02129; [b]Institute of Computer Science, P. J. Šafárik University, Košice, 04001, Slovakia; [c]Hearing Research Center, Boston University, Boston, MA 02215; and [d]Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139

Neuronal mechanisms of auditory distance perception are poorly understood, largely because contributions of intensity and distance processing are difficult to differentiate. Typically, the received intensity increases when sound sources approach us. However, we can also distinguish between soft-but-nearby and loud-but-distant sounds, indicating that distance processing can also be based on intensity-independent cues. Here, we combined behavioral experiments, fMRI measurements, and computational analyses to identify the neural representation of distance independent of intensity. In a virtual reverberant environment, we simulated sound sources at varying distances (15–100 cm) along the right-side interaural axis. Our acoustic analysis suggested that, of the individual intensity-independent depth cues available for these stimuli, direct-to-reverberant ratio (D/R) is more reliable and robust than interaural level difference (ILD). However, on the basis of our behavioral results, subjects' discrimination performance was more consistent with complex intensity-independent distance representations, combining both available cues, than with representations on the basis of either D/R or ILD individually. fMRI activations to sounds varying in distance (containing all cues, including intensity), compared with activations to sounds varying in intensity only, were significantly increased in the planum temporale and posterior superior temporal gyrus contralateral to the direction of stimulation. This fMRI result suggests that neurons in posterior nonprimary auditory cortices, in or near the areas processing other auditory spatial features, are sensitive to intensity-independent sound properties relevant for auditory distance perception.

computational modeling | psychophysics | spatial hearing | what and where pathways | neuronal adaptation

Determining the distance of objects is of key value in many everyday situations. For objects that fall outside the field of vision, hearing is the only sense that provides such information. For example, consider a person reaching for a ringing phone (1–3) or a listener using distance differences to help focus on one talker in a chattering crowd (4, 5). However, whereas cortical representations of features such as spectral or amplitude modulations have been intensively examined (6, 7), the neural mechanisms of auditory distance perception are poorly understood.

In comparison with the detailed mapping of human visual cortex, knowledge of the subsystems of human auditory cortex is less complete. Only relatively broad anatomical divisions, such as that between the anterior "what" vs. posterior "where" (8–13) pathways, have been shown in human neuroimaging studies. The posterior auditory "where" pathway, which encompasses nonprimary auditory cortex areas including the planum temporale (PT) and posterior superior temporal gyrus (STG), is strongly activated by horizontal sound direction changes (12, 14–16) and movement (17, 18). However, although human (19, 20) (and nonhuman primate, refs. 21, 22) auditory systems have been shown to have neurons preferring sound sources approaching the listener (i.e., "looming"), populations responsible for auditory distance processing have not been previously identified. In fact,

auditory distance perception is also relatively poorly understood from the psychophysical and computational points of view (1).

In many situations, the dominant acoustic cue for distance is the overall received stimulus intensity. However, the overall received stimulus intensity is an ambiguous distance cue (e.g., if the emitted stimulus intensity varies independently of the source distance), which can complicate studies on intensity vs. distance processing (23). Acoustically, for sources off the midline, it is impossible to fix the received stimulus intensity at both ears as distance varies, because the rate of change in intensity with distance is different at the two ears. Physiologically, when the stimulus intensity varies, detectors sensitive to a range of stimulus features can get activated. In such cases, neurons tend to respond in a nonspecific way for loud stimuli, even if they are sensitive to a specific feature at low stimulus intensities. These nonspecific responses are most commonly observed when the sounds are broadband noise (24, 25), as typically used in studies of audiospatial perception. However, recent results show that robust intensity-independent distance perception is possible for nearby sources (up to 100 cm from the listener) in simulated reverberant environments. For such sources, two major stimulus-independent and intensity-independent distance cues exist, the interaural level difference (ILD) (26, 27) and direct-to-reverberant energy ratio (D/R) (28–30). Although it is currently not fully known which of the two cues the listeners use and/or how they combine them (31), the cues' availability gives an opportunity to study intensity-independent distance representations using human neuroimaging.

In humans, tuning properties of neuronal populations can be noninvasively studied by examining neuronal adaptation, i.e., suppression of responses to a given stimulus as a function of its similarity and temporal proximity to preceding stimuli (12, 32–34). Adaptation studies can use invasive neurophysiological measurements (33) or noninvasive methods such as electroencephalography (EEG) and magnetoencephalography (MEG) (12, 32, 34), or fMRI (35). Adaptation fMRI compares signals elicited by two sound-sequence conditions, one consisting of identical (i.e., constant) stimuli and the other consisting of sounds varying along the feature dimension of interest. Adaptation fMRI presumably differentiates the tuning properties of neurons within each voxel. Specifically, in a voxel populated by neurons tuned loosely to the feature dimension of interest, such as sound direction (16), fMRI signals get equally adapted during constant and varying stimulation. In contrast, in a voxel with neurons tuned sharply to the feature dimension of interest, a release of adaptation and increased fMRI signal is observed during varied vs. constant stimulus blocks.

NEUROSCIENCE
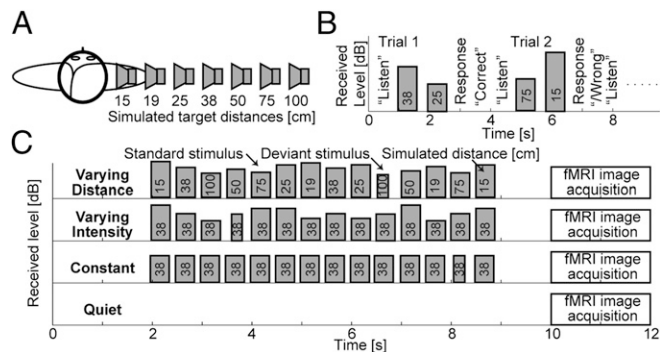
PSYCHOLOGICAL AND COGNITIVE SCIENCES

Fig. 1. Experimental design. (*A*) Simulated source locations. (*B*) Timing of events during trials in the behavioral experiment: The instruction "listen" appeared on the screen, followed by presentation of two stimuli from different distances. Listeners responded by indicating whether the second stimulus sounded more or less distant than the first stimulus. On-screen feedback was provided. Presentation intensity was randomly roved for each stimulus so that received intensity could not be used as a cue in the distance discrimination task. (*C*) Timing of stimuli and of image acquisition during one imaging trial in the fMRI experiment, shown separately for the four stimulus conditions used in the experiment. Height of the stimulus bars corresponds to the *received* stimulus intensity at the listeners' ears. In the varying distance condition, the stimulus distance changed randomly, whereas the stimulus presentation intensity was fixed (thus, both the perceived distance and intensity varied). In the varying intensity condition, the stimulus distance was fixed, whereas the received intensity varied over the same range as in the varying distance condition. In the constant condition, the 38-cm stimulus was presented repeatedly at constant intensity. In these active stimulation conditions, the listener's task was to detect deviant stimuli that were shorter than the standard stimuli. No feedback was provided. In the fixation condition, no auditory stimuli were presented, and the subjects were instructed to look at a point on the screen.

Here, we examined auditory distance perception and its neural correlates. Nearby sources located at various distances were simulated along the interaural axis (to the right of the listener) in a reverberant environment (Fig. 1*A*). Two different adaptation fMRI experiments were conducted to localize brain areas sensitive to intensity-independent auditory distance cues. A behavioral experiment was performed to validate the virtual acoustics used during fMRI, and to examine the perceptual cues and psychophysical principles determining listeners' sensitivity to changes in distance.

Results

Behavioral Experiment. We used virtual acoustics to measure perceptual sensitivity to intensity-independent stimulus distance cues (Fig. 1*A*). The main goal was to confirm that the listeners could use the intensity-independent cues to judge source distance in our simulated auditory environment. Subjects were instructed to indicate whether the second of two noise bursts, simulated from different distances, originated closer or farther than the first burst. These discriminations had to be made irrespectively of perceived sound intensity, which varied randomly for each burst to eliminate its contribution as a distance cue (Fig. 1*B*). Overall, subjects accurately judged sound source distances despite the confounding information provided by intensity cues. As expected, the listeners' accuracy improved with increasing distance difference between the two simulated sound sources (Fig. 2 and for further details, *SI Results* and Fig. S1).

In addition to confirming that our virtual auditory environment was robust, these results demonstrate that distance discrimination sensitivity is largely independent of the baseline distance of the simulated sound sources, as long as the ratio of compared distances is fixed. Specifically, in Fig. 2, bars are grouped such that each group shows the performance for a fixed number of unit log-distance intervals (i.e., the ratio of source distances). Within each
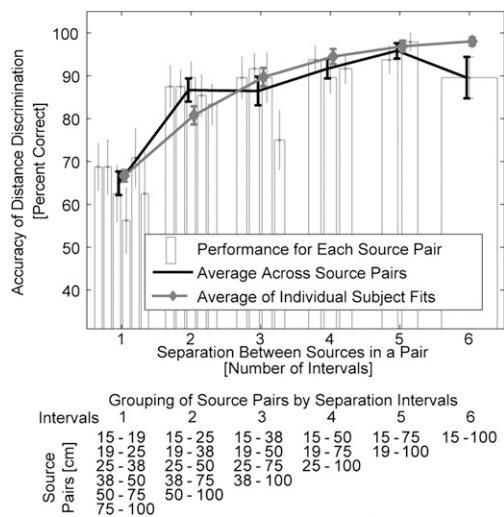

Fig. 2. Behavioral distance discrimination responses. The black line shows across-subject average accuracy collapsed across simulated source pairs separated by the same number of unit log-distance intervals (Fig. 1*A* and table *Below*). The accuracy improved as the simulated source separation (number of intervals) increased, reflecting the robustness of our virtual 3D stimuli. Gray line represents the averages of individual-subjects' accuracy predictions on the basis of subjects' individual estimates of distance sensitivity (*d'*) (*SI Results*). Bars show across-subject average performance separately for each source-distance pair. Bars are grouped on the basis of the number of intervals between sources within the pair (table *Below* lists in each column the source pairs that are separated by the same number intervals; separation ranges from 1 interval (all possible pairs of adjacent sources) to 6 intervals (the closest vs. farthest source). In each group, bars are ordered from left to right on the basis of absolute distances (listed from top to bottom in the table, respectively). No systematic upward or downward trend is visible in performance across source pairs within each group, showing that Weber's law holds for intensity-independent near-head sound source distances. Error bars represent SEM.

group, bars are ordered from the nearest to the farthest pair. No systematic upward or downward trend is visible in performance across simulated source-distance pairs within each group. To confirm this lack of trend, a linear approximation was made to each subject's performance as a function of the distance pair, separately for each group and each subject. Across subjects, none of the slopes was significantly different from zero (two-tailed Student *t* test with Bonferroni correction performed on the fitted slopes separately for each source-pair group; $P > 0.1$), consistent with the predictions of Weber's law (for more details, see the two sections, *Acoustic Analysis of Stimuli* and *Predictions of Distance Discrimination* in *SI Results*, and Fig. S2). Finally, our acoustical analyses described in *SI Results* and Fig. S2 show that, of the individual intensity-independent distance cues, D/R is more robust and reliable than ILD.

fMRI Experiment. To localize intensity-independent distance representations, we compared auditory cortex areas activated during distance vs. intensity changes vs. constant repetitive sound stimulation (Fig. 1*C*). To control for fluctuations in attention and alertness, subjects were asked to detect occasional changes in sound duration that occurred independently of distance or intensity (across-subject mean hit rate = 93%; mean reaction time = 650 ms, data of one subject failing to conduct this task were excluded from the analyses; see *SI Results*).

As expected, constant auditory stimulation significantly activated primary (the medial 2/3 of Heschl's gyrus, HG) and nonprimary (STG, PT, and planum polare, PP) auditory cortex areas (Fig. 3*A*). Introducing changes to auditory stimulation increased auditory cortex activations, as shown by the contrasts between
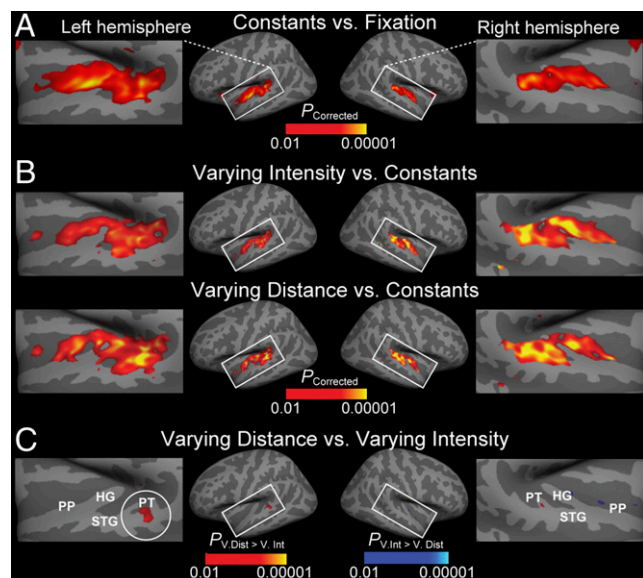
Fig. 3. fMRI-adaptation data from the fMRI experiment presented on inflated cortical surface. (*A*) Compared with the silent fixation condition, constant condition stimulation activates primary (HG) and surrounding nonprimary auditory cortex areas (STG, PT, and PP). (*B*) Compared with constant condition, the varying intensity and varying distance conditions increase activations particularly in nonprimary areas. (*C*) An area specifically sensitive to auditory distance cues, independent of intensity, is revealed in the contrast between varying distance vs. varying intensity conditions, in the PT/posterior STG (circled), in or near the putative posterior auditory-cortex "where" pathway.

Fig. 4. Hypothesis-based region-of-interest (ROI) analysis of posterior nonprimary auditory cortex activations during auditory distance processing. A significant increase of left posterior auditory cortex ROI activity was observed during varying distance vs. varying intensity conditions, suggesting that posterior nonprimary auditory cortices include neurons with intensity-independent distance representations. Data also show increased activities in both hemispheres during varying vs. constant stimulation (**P = 0.01, ***P < 0.001, SEM error bars).

varying intensity or varying distance vs. constant conditions (Fig. 3*B*). Note that these contrasts did not reach significance at the crest of HG, consistent with the view that primary auditory cortices are adapted less prominently by high-rate stimulation than the surrounding nonprimary areas (36). In Fig. 3*B*, the areas activated in the varying distance contrast overlap with the areas activated in the varying intensity contrast, as expected given that the stimuli in varying distance condition also varied in intensity. Finally, Fig. 3*C* shows the areas that represented more prominent release from adaptation during varying distance than during varying intensity condition. This contrast suggests that the posterior nonprimary auditory cortex areas (posterior STG and PT) are specifically sensitive to intensity-independent auditory distance cues.

In an additional hypothesis-based fMRI analysis (Fig. 4), a region of interest (ROI) was defined in each hemisphere by combining two anatomical FreeSurfer standard-space labels (PT and posterior aspect of STG) that were a priori (8–12) conjectured to encompass areas activated by audiospatial features. Consistent with the whole-brain mapping results (Fig. 3), blood oxygen level dependent (BOLD) percentage signal changes were significantly stronger ($F_{1,10}$ = 9.3, P = 0.01) during varying distance than varying intensity conditions in the left-hemisphere posterior auditory cortex ROI, but not in the right hemisphere. These data also show the significant increases of auditory activity during varying vs. constant stimulation in the left ($F_{1,10}$ = 33.4, P < 0.001) and right ($F_{1,10}$ = 83.6, P < 0.001) hemisphere ROIs.

Additionally, we examined the structures activated when the presentation intensity was normalized such that the overall energy received at the ear closer to the source was constant, independent of the source distance (*SI Results* and Fig. S3). This experiment revealed more widespread activations than the varying distance vs. varying intensity contrast in the main fMRI experiment (Fig. 3*C*), most likely because the near-ear intensity normalization did not result in exclusive activation of the distance-sensitive areas.
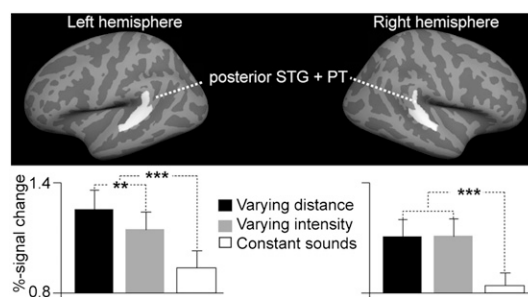
Finally, to enhance comparability to fMRI studies in volume space, we conducted a whole-brain analysis in a 3D standard brain, using a statistical approach similar to the surface-space analysis (Fig. 5). In the varying distance vs. varying intensity contrast, a significant activation cluster was identified. The strongest voxel was located in the left PT (Montreal Neurological Institute Talairach coordinates $\{x, y, z\}$ = $\{-60, -35, 15\}$, Talairach $\{x, y, z\}$ = $\{-59, -33, 16\}$, cluster volume 912 mm^3) with the area extending also to the posterior STG.

Discussion

We studied perceptual and neuronal mechanisms of auditory distance processing by combining computational acoustic analyses with behavioral and adaptation fMRI measurements in a virtual auditory environment. Our results provide evidence for auditory areas specifically sensitive to acoustic distance cues that are independent of sound intensity changes. These activations centered in the posterior STG and PT, near the so-called auditory "where" pathway (10). Although a number of recent studies have associated these areas with perception of sound direction changes (12, 16) and movement (17, 37, 38), our study to our knowledge is unique in revealing neuron populations sensitive to sound-source distance variations.

Auditory-cortex neurons have a strong tendency to stimulus-specific adaptation (33) that suppresses their responsiveness to repetitive sounds. Differential release from this adaptation, as a function of increasing dissimilarity of successive stimuli, is believed to reveal units that are specifically sensitive to the varying feature dimension (12, 34). However, this approach can be confounded by the fact that stimulus intensity changes can increase activations of various feature detectors. In particular,
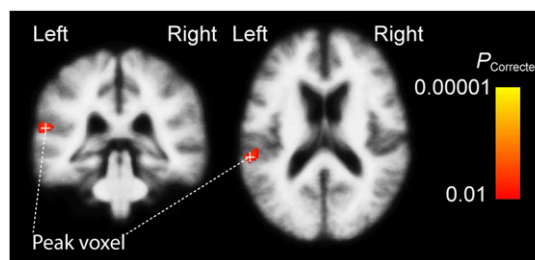

Fig. 5. Volume-based fMRI analysis of activations during varying distance vs. intensity. Significant activation cluster extends from the left PT to the left posterior STG.

when the baseline stimulus is broadband noise containing all possible features, stimulus intensity is known to interact with feature specificity of auditory neurons (24, 25). A supporting fMRI experiment (*SI Results* and Fig. S3) was performed, in which the *received* intensity of varying-distance stimuli was normalized at the near ear to minimize the intensity variation. This experiment showed relatively widespread and nonspecific activation patterns, thus confirming the expectation that various feature detectors can be released from adaptation when intensity varies. However, our main analyses were based on the contrast between conditions with varying distance and received intensity in the listener's ears (equally loud sounds simulated from different distances) vs. varying only the received intensity (sounds of different intensity from one simulated distance; Fig. 1*C*). The significant increase in auditory-cortex activation in PT and posterior STG in this contrast presumably revealed neurons sensitive to sound-distance cues, independent of intensity.

The present study concentrated in the near-head range where robust distance cues were presumed to be available. Although larger distances could be represented differently (39), it is possible that the posterior nonprimary auditory cortex areas, near those activated during intensity-independent nearby-source distance processing, also represent distances of sounds originating from farther away. As mentioned above, these areas have been previously shown to be associated with other aspects of spatial hearing (12, 16, 17, 37, 38). For example, a study (16) using an adaptation fMRI paradigm analogous to our approach found bilateral activations to sound-direction changes in PT, near the present distance-dependent activations. It is also noteworthy that in the present study, the strongest activations to intensity-independent sound-distance changes occurred in the left PT/posterior STG, contralateral to the virtual origins of the stimuli. However, given that only one simulated direction was studied, the present data cannot offer conclusive evidence that the distance of sounds originating on one side of the head is always represented by the contralateral hemisphere or that sound-source distance (or direction) is represented in a topographical fashion analogous to other sensory modalities (note also evidence supporting an alternative two-channel model, refs. 40–42).

An important theoretical question is whether the present PT/posterior STG activations reflect neurons encoding auditory distance, per se, or whether they correspond to distinct sub-populations for D/R and/or ILD distance cues. If the latter were true, the present fMRI results could also be interpreted as a partial "byproduct" of ILD neurons that are primarily related to direction discrimination. However, basic binaural cues such as ILD are extracted and extensively processed already at the subcortical level (for a review, see ref. 42). If we understand the cerebral cortex as a self-organizing network, it is more likely that at higher processing levels, beyond the subcortical pathways and primary sensory areas, representations develop for combinations of features that frequently occur together. Previous neurophysiological (8, 11) and human neuroimaging studies (43, 44) generally suggest that nonprimary regions, including PT/posterior STG, process progressively more complex sounds than the primary "core" regions of auditory cortex. These observations are in line with a previous MEG study (45) that systematically compared activations to individual and combined 3D sound features, and showed that ILD alone is not sufficient for producing direction-specific activations at the cortical level. It is also worth noting that in humans, no consistent evidence on ILD-specific fMRI activations in PT/posterior STG exists. The few published ILD–fMRI studies have, instead, compared transient vs. sustained BOLD response differences in more primary regions (46) or showed direction-specific ILD effects only in areas beyond auditory cortices (47). Nevertheless, some MEG (48) and EEG (49) studies have reported nonspecific differences in cortical response patterns to ILD and interaural time differences (ITD).

There is also evidence from animal studies suggesting ILD neurons in primary auditory cortices, for example, in the cat (50) (note, however, that the link between animal physiology and human fMRI is relatively weak in the spatial auditory domain) (51). Further, our auditory systems are highly adaptive, and different cues may govern distance perception in different acoustic environments. For example, a recent psychophysical study suggested that listeners do not use the ILD cue for distance perception in natural reverberation (31), even though it is an important distance cue in artificial anechoic spaces where other naturally occurring cues are unavailable (52). Further studies are therefore needed to investigate how individual spatial features are processed across the hierarchy of auditory pathways in humans.

Our behavioral data show that sensitivity to auditory distance changes is constant for sources separated by a constant distance ratio, following Weber's law. This result holds for nearby distances varying along the interaural axis, as examined here, which could be relevant for general models of auditory distance perception. Specifically, both the rate of distance-dependent changes in the auditory distance cues (ILD and D/R) (*SI Results*, *Acoustic Analysis of Stimuli*, and Fig. S2) and/or the perceptual sensitivity to these cue changes have been suggested to vary as a function of sources' absolute distance (27, 30, 53). One might therefore expect that distance sensitivity would vary with absolute distance if performance were based on either of the two cues alone. Here, we did not find any systematic trends suggesting that distance-discrimination performance depends on the absolute source distance. For example, subjects discriminated distance differences between the smallest relative intervals on average at a 67–71% accuracy, with the best nominal performance observed for the distance pair 50 vs. 75 cm from the head (Fig. 2). On the basis of the previous findings concerning discriminability of individual cues (as discussed in *SI Results*, *Predictions of Distance Discrimination*), the D/R-only–based performance would be expected to be poorest for the smallest absolute distances (in the case of the smallest interval, 15 vs. 19 cm) and best for the largest distance (75 vs. 100 cm). Because no such D/R-only (or ILD-only) dependence was observed, our results suggest that distance judgments were not based on a simple mapping of a single distance cue (in contrast to a previous interpretation) (31). However, as shown in *SI Results*, it may still be most parsimonious to assume that the main cue used by the listeners was the monaural near-ear D/R.

In summary, our results reveal an area in the posterior nonprimary auditory cortex that is sensitive to intensity-independent auditory distance cues.

Materials and Methods

Subjects. The same set of subjects (*n* = 12, ages 20–33 y, five females) with self-reported normal hearing ability participated in all three experiments. Human subjects' approval was obtained and voluntary consents were signed before measurements.

Stimuli and Setup. All stimuli used in this study were generated in virtual auditory space (54, 55), using binaural room impulse responses (BRIR) that included realistic room reverberation. A single set of nonindividualized BRIRs was used, measured on a listener that did not participate in this study. Unless specified otherwise, all details of the measurement procedures, including the microphone, speaker, and the BRIR measurement technique used were identical to our previous study (27).

The BRIRs were measured in a small classroom (3.4 m × 3.6 m × 2.9 m height) using the Bose FreeSpace 3 Series II surface-mount cube speaker (Bose). The room was carpeted, with hard walls and acoustic tiles covering the ceiling. The room reverberation times, T_{60}, in octave bands centered at 500, 1,000, 2,000, and 4,000 Hz ranged from 480 to 610 ms. The BRIRs were measured by placing miniature microphones (FG-3329c; Knowles Electronics) at the blocked entrances of the listener's ear canals. The loudspeaker was set to face the listener at various distances (15, 19, 25, 38, 50, 75, or 100 cm) from the center of the listener's head, directly to the right of the listener along the interaural axis, at the level of the listener's ears (Fig. 1*A*).

The target stimulus was a 300-ms-long sample of white noise, bandpass-filtered at 100–8,000 Hz. A set of 50 independent noise burst tokens was generated and each token was convolved with each of the BRIRs to create standard stimuli for each source distance (Fig. 1A). An identical set of deviant stimuli was also generated, differing from the standard stimuli only in that the deviant duration was 150 ms. For each experimental trial, either two (behavioral experiment, Fig. 1B) or 14 (imaging experiments, Fig. 1C) noise bursts were randomly selected, scaled depending on the normalization scheme used, and placed in series with stimulus onset asynchrony, to create the stimulus sequence. Finally, the stimuli were filtered to compensate for the headphone transfer functions.

The stimuli used in the experiments mainly differed by the overall stimulus intensity normalization used. For the behavioral experiment, each noise burst was normalized so that its overall intensity received at the right ear (which was close to the simulated source) was fixed and then randomly roved over a 12 dB range so that the monaural overall intensity distance cue was eliminated at both the left and right ears (Fig. 1B).

The imaging experiment stimulus sequences always consisted of 14 noise bursts (Fig. 1C). Each stimulus sequence in trials with varying distance contained two noise bursts for each of the seven distances, ordered pseudorandomly such that each distance was present at least once before the second occurrence of any of the distances. The varying distance stimuli were normalized such that the presentation intensity was fixed (i.e., the overall intensity distance cue was present in these stimuli because the received intensity for the near sources was higher than the received intensity for the far sources). The varying intensity stimuli were simulated from the fixed distance of 38 cm and their presentation intensity was varied such that the received intensity at the near ear varied across the same range as for the varying distance stimuli. The constant stimuli were presented at a fixed distance of 38 cm and at a fixed presentation intensity, equal to the varying distance stimuli.

The stimulus sequence files, generated at a sampling rate of 44.1 kHz, were stored on the hard disk of the control computer (IBM PC-compatible), running Presentation (Neurobehavioral Systems). On each trial, one of the sequences was selected and presented through the Fireface 800 sound processor (RME), Kramer 900XL amplifier (Kramer Electronics), and Sensimetrics S14 (Sensimetrics) MRI-compatible headphones. Average received level was 65 dB(A), measured at the near ear of a KEMAR (Manikin for Acoustic Research; Knowles Electronics) manikin equipped with the DB-100 Zwislocki Coupler and the Etymotic Research ER-11 microphones. Responses were collected via MRI-compatible five-key universal serial bus (USB) keyboard. A video projector was attached to the control computer and projected the instructions to the subject in the scanner.

Experimental Procedure and Data Acquisition. All experiments were performed during a single 2-h long experimental session using the same experimental setup, equipment, and stimuli. After the initial preparation and practice runs, which were performed outside the scanner, the experiments were performed inside the scanner. First, the imaging experiments were run while the listener was instructed to perform a sound duration deviant detection task. Then, the behavioral experiment was performed. This ordering (behavioral experiment after imaging experiments) was chosen to prevent listeners from focusing on the stimulus distance during the imaging experiments, which could alter the listeners' concentration level during the different imaging stimulus conditions.

Behavioral experiment. The behavioral experiment was a distance discrimination experiment, consisting of four runs, each containing 21 randomly ordered trials (1 trial for each combination of two of the seven distances). Two different random noise burst tokens were selected for each trial, one for each distance. A trial started by the word "listen" appearing on the computer screen (Fig. 1B), followed after 200 ms by the two auditory stimuli presented with a 1,000 ms stimulus onset asynchrony. The subject responded by indicating whether the second sound source was closer or farther away than the first source by pressing one of two keys on the keyboard. The experiment was self-paced and the total duration of one trial was, on average, ~5 s.

Imaging experiment. The imaging experiment consisted of two runs, each containing 24 different random trials for each of the four stimulus sequence

types (Fig. 1C). The order of stimulus trial types was randomized to maximize the sensitivity of the experiment (56). Each trial consisted of 2 s of silence, followed by 7 s of auditory stimulus presentation, 1 s of silence, and 2 s of fMRI image acquisition (Fig. 1C). On 50% of trials, one randomly chosen burst in the stimulus sequence was replaced by a shorter, 150-ms deviant burst. Subjects were instructed to press a key on the keyboard whenever this deviant was detected (i.e., the listeners were instructed to focus on the stimulus duration, even though they were aware that the stimulus distance varied).

Whole-head fMRI was acquired at 3T using a 32-channel coil (Siemens; TimTrio). To circumvent response contamination by scanner noise, we used a sparse-sampling gradient-echo BOLD sequence (TR/TE = 12,000/30 ms, 9.82-s silent period between acquisitions, flip angle = 90°, FOV 192 mm) with 36 axial slices aligned along the anterior–posterior commissure line (3-mm slices, 0.75-mm gap, 3×3 mm² in-plane resolution), with the coolant pump switched off. T1-weighted anatomical images were obtained for combining anatomical and functional data using a multiecho MPRAGE pulse sequence (TR = 2,510 ms; 4 echoes with TEs = 1.64 ms, 3.5 ms, 5.36 ms, 7.22 ms; 176 sagittal slices with $1 \times 1 \times 1$ mm³ voxels, 256×256 mm² matrix; flip angle = 7°).

Data Analysis. *Behavioral data.* In the behavioral experiment, the percentage of correct responses was analyzed for each distance pair, repeat, and subject. Across-subject means and SEMs were computed. To obtain an overall measure of individual subject performance, a simple decision theory model was proposed and fitted to the percentage of correct data (*SI Results*).

In the imaging experiment, the subjects' task was to detect a deviant sound in the stimulus sequence. The response was accepted as a correct deviant detection if it occurred within 2.5 s after the deviant onset. Hit rates and reaction times for the correct detections were analyzed (*SI Results*). Repeated-measures ANOVA was used to test whether these measures varied significantly for different types of stimulus.

fMRI Data. Cortical surface reconstructions and standard-space coregistrations of the individual anatomical data, as well as functional data analyses, were conducted using Freesurfer 5.0. Individual functional volumes were motion corrected, coregistered with each subject's structural MRI, intensity normalized, resampled into standard cortical surface space (57, 58), smoothed using a 2-dimensional Gaussian kernel with an FWHM of 5 mm, and entered into a general-linear model (GLM) with the task conditions as explanatory variables. A random-effects GLM was then conducted at the group level. To control for multiple comparisons, a cluster analysis (Monte Carlo simulations with 10,000 iterations; $P < 0.01$) was used, with specific contrasts between the conditions (varying intensity, varying distance, and constant stimulation) constrained to areas where the main effect (auditory stimulation vs. baseline) was significant (corrected $P < 0.01$). In addition, to enhance comparability of our main results, data were also analyzed in a 3D standard volume. The statistical procedures were identical to the main analysis.

In the hypothesis-based fMRI analysis, a ROI was defined in each hemisphere by combining two anatomical FreeSurfer standard-space labels (PT and posterior aspect of STG) that were a priori (8–12) conjectured to encompass areas activated by audiospatial features. These ROIs were then resampled through the spherical standard space onto each individual subject's brain representations. Within each subject, within-ROI voxels showing significant activations ($P < 0.01$) were used to determine percentage of signal changes vs. the baseline rest condition during different task conditions. Results were entered into a repeated-measures ANOVA with a priori difference contrasts.

1. Zahorik P, Brungart DS, Bronkhorst AW (2005) Auditory distance perception in humans: A summary of past and present research. *Acta Acust. United Ac* 91:409–420.
2. Maier JX, Neuhoff JG, Logothetis NK, Ghazanfar AA (2004) Multisensory integration of looming signals by rhesus monkeys. *Neuron* 43:177–181.
3. Neuhoff JG (1998) Perceptual bias for rising tones. *Nature* 395:123–124.
4. Brungart DS, Simpson BD (2002) The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal. *J Acoust Soc Am* 112:664–676.
5. Shinn-Cunningham BG, Schickler J, Kopčo N, Litovsky RY (2001) Spatial unmasking of nearby speech sources in a simulated anechoic environment. *J Acoust Soc Am* 110:1118–1129.
6. Adriani M, et al. (2003) Sound recognition and localization in man: Specialized cortical networks and effects of acute circumscribed lesions. *Exp Brain Res* 153:591–604.
7. Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14:474–480.
8. Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.

NEUROSCIENCE

PSYCHOLOGICAL AND COGNITIVE SCIENCES

9. Rauschecker JP (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol Suppl* 532:34–38.
10. Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci USA* 97:11800–11806.
11. Rauschecker JP (1998) Cortical processing of complex sounds. *Curr Opin Neurobiol* 8: 516–521.
12. Ahveninen J, et al. (2006) Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci USA* 103:14608–14613.
13. Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293.
14. Brunetti M, et al. (2005) Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. *Hum Brain Mapp* 26:251–261.
15. Tata MS, Ward LM (2005) Early phase of spatial mismatch negativity is localized to a posterior "where" auditory pathway. *Exp Brain Res* 167:481–486.
16. Deouell LY, Heller AS, Malach R, D'Esposito M, Knight RT (2007) Cerebral responses to change in spatial location of unattended sounds. *Neuron* 55:985–996.
17. Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. *Neuron* 34:139–148.
18. Krumbholz K, et al. (2005) Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. *Cereb Cortex* 15:317–324.
19. Hall DA, Moore DR (2003) Auditory neuroscience: The salience of looming sounds. *Curr Biol* 13:R91–R93.
20. Seifritz E, et al. (2002) Neural processing of auditory looming in the human brain. *Curr Biol* 12:2147–2151.
21. Ghazanfar AA, Neuhoff JG, Logothetis NK (2002) Auditory looming perception in rhesus monkeys. *Proc Natl Acad Sci USA* 99:15755–15757.
22. Maier JX, Ghazanfar AA (2007) Looming biases in monkey auditory cortex. *J Neurosci* 27:4093–4100.
23. Zahorik P, Wightman FL (2001) Loudness constancy with varying sound source distance. *Nat Neurosci* 4:78–83.
24. Sigalovsky IS, Melcher JR (2006) Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers. *Hear Res* 215:67–76.
25. Ernst SM, Verhey JL, Uppenkamp S (2008) Spatial dissociation of changes of level and signal-to-noise ratio in auditory cortex for tones in noise. *Neuroimage* 43:321–328.
26. Brungart DS (1999) Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am* 106:3589–3602.
27. Shinn-Cunningham BG, Kopčo N, Martin TJ (2005) Localizing nearby sound sources in a classroom: Binaural room impulse responses. *J Acoust Soc Am* 117:3100–3115.
28. Mershon DH, King LE (1975) Intensity and reverberation as factors in the auditory perception of egocentric distance. *Percept Psychophys* 18:409–415.
29. Hartmann WM (1983) Localization of sound in rooms. *J Acoust Soc Am* 74:1380–1391.
30. Zahorik P (2002) Direct-to-reverberant energy ratio sensitivity. *J Acoust Soc Am* 112: 2110–2117.
31. Kopčo N, Shinn-Cunningham BG (2011) Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am* 130:1530–1541.
32. Näätänen R, et al. (1988) Frequency and location specificity of the human vertex N1 wave. *Electroencephalogr Clin Neurophysiol* 69:523–531.
33. Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
34. Jääskeläinen IP, Ahveninen J, Belliveau JW, Raij T, Sams M (2007) Short-term plasticity in auditory cognition. *Trends Neurosci* 30:653–661.

35. Grill-Spector K, Malach R (2001) fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107:293–321.
36. Lü ZL, Williamson SJ, Kaufman L (1992) Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Res* 572:236–241.
37. Baumgart F, Gaschler-Markefski B, Woldorff MG, Heinze HJ, Scheich H (1999) A movement-sensitive area in auditory cortex. *Nature* 400:724–726.
38. Warren JD, Griffiths TD (2003) Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J Neurosci* 23:5799–5804.
39. Graziano MS, Reiss LA, Gross CG (1999) A neuronal representation of the location of nearby sounds. *Nature* 397:428–430.
40. McAlpine D, Jiang D, Palmer AR (2001) A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4:396–401.
41. Salminen NH, May PJ, Alku P, Tiitinen H (2009) A population rate code of auditory space in the human cortex. *PLoS ONE* 4:e7600.
42. Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. *Physiol Rev* 90:983–1012.
43. Hall DA, et al. (2002) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 12:140–149.
44. Wessinger CM, et al. (2001) Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *J Cogn Neurosci* 13:1–7.
45. Palomäki KJ, Tiitinen H, Mäkinen V, May PJ, Alku P (2005) Spatial processing in human auditory cortex: The effects of 3D, ITD, and ILD stimulation techniques. *Brain Res Cogn Brain Res* 24:364–379.
46. Lehmann C, et al. (2007) Dissociated lateralization of transient and sustained blood oxygen level-dependent signal components in human primary auditory cortex. *Neuroimage* 34:1637–1642.
47. Zimmer U, Lewald J, Erb M, Karnath HO (2006) Processing of auditory spatial cues in human cortex: An fMRI study. *Neuropsychologia* 44:454–461.
48. Johnson BW, Hautus MJ (2010) Processing of binaural spatial information in human auditory cortex: Neuromagnetic responses to interaural timing and level differences. *Neuropsychologia* 48:2610–2619.
49. Tardif E, Murray MM, Meylan R, Spierer L, Clarke S (2006) The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. *Brain Res* 1092:161–176.
50. Imig TJ, Irons WA, Samson FR (1990) Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *J Neurophysiol* 63:1448–1466.
51. Werner-Reiss U, Groh JM (2008) A rate code for sound azimuth in monkey auditory cortex: Implications for human neuroimaging studies. *J Neurosci* 28:3747–3758.
52. Brungart DS, Durlach NI, Rabinowitz WM (1999) Auditory localization of nearby sources. II. Localization of a broadband source. *J Acoust Soc Am* 106:1956–1968.
53. Larsen E, Iyer N, Lansing CR, Feng AS (2008) On the minimum audible difference in direct-to-reverberant energy ratio. *J Acoust Soc Am* 124:450–461.
54. Carlile S (1996) *Virtual Auditory Space: Generation and Applications* (RG Landes, New York).
55. Zahorik P (2002) Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111:1832–1846.
56. Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114.
57. Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9:195–207.
58. Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8:272–284.

Supporting Information

Kopčo et al. 10.1073/pnas.1119496109

SI Results

Additional Analysis of the Behavioral Experiment Data. The main goal of the behavioral experiment was to confirm that the listeners could use the intensity-independent cues to judge source distance in our simulated auditory environment. The virtual space simulation had limitations due to technical constrains of the fMRI-compatible stimulus presentation equipment and due to the fact that nonindividualized binaural room impulse responses (BRIRs) were used in the simulation (1). Fig. 2 shows that, on average, the subjects could judge distance well, despite these limitations. To evaluate individual subjects' performance, the sensitivity index d' was estimated for each subject on the basis of his/her discrimination performance. The estimation was based on a psychophysical decision theory model (2) that assumed that (*i*) the listener's internal percept evoked in response to a stimulus presented from distance s can be described by a Gaussian-distributed random variable X; (*ii*) the mean value of X grows logarithmically with the actual source distance s; (*iii*) the variance in the internal representation, corresponding to internal noise, is fixed for each listener, independent of the actual stimulus distance, presentation intensity, or any other stimulus parameter; and (*iv*) each listeners' discrimination responses are unbiased optimal decisions based solely on the currently observed two values of X, corresponding to the two source distances presented in a given trial. On the basis of these assumptions, the following equation defines the percenage of correct performance, P_C, as a function of d' and the locations of the stimulus sources:

$$P_c = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{d'}{2}} e^{\frac{-t^2}{2}} \mathrm{dt},$$

where $d' = |\ln s_1 - \ln s_2|/\sigma$. The constant σ [with the units of ln (cm)] was fitted for each subject to minimize the mean squared error between the subject's average performance for each source pair and the predicted performance based on the subject's estimate of σ. On the basis of these estimates, individual subject performance prediction was computed for each distance interval from Fig. 2. Gray line in Fig. 2 plots the across-subject averages based on this performance estimate. There is a good agreement between the gray and black lines in Fig. 2, confirming that the assumptions of the model and the individual estimates of σ provide a good basis for characterizing the across-subject average percentage of correct performance.

Fig. S1 shows the individual subjects' estimates of internal noise σ. The values are very similar across most subjects (all except for subjects S4 and S12), confirming that they could reliably discriminate distance in our simulated auditory environment. The relatively worse performance of subject S12 (and, to a lesser extent subject S4) could be due to the technical limitations of our simulations or due to other factors, e.g., subjects' inability to follow the instructions.

A behavioral measurement was also performed during the fMRI experiments. During the experiments, the subject's vigilance and focus needed to be constant across the different stimulus conditions to prevent vigilance-related changes in brain activation from confounding the stimulus-related changes. To establish that different stimulus conditions were approximately equally demanding, 50% of trials during imaging experiments contained deviants, noise bursts of duration that was half of the standard bursts (Fig. 1*C*). The listeners were instructed to respond with a button press whenever they detected the deviant. Analysis of hit rates and reaction times was performed. One subject (S12) did not perform the task correctly (hit rate 0–20%). For the remaining subjects, the task difficulty was similar across the three different stimulus types (across-subject average hit rate of 93.5, 93.2, and 91% and reaction times of 630, 660, and 660 ms for the varying distance, varying intensity, and constant stimuli, respectively), suggesting that any fMRI activation differences across conditions cannot be attributed to differences in task difficulty. Subject S12 failed to follow the instructions for the behavioral task during the imaging experiments and also had the worst performance in the behavioral experiment. Therefore, this subject's imaging data were excluded.

Acoustic Analysis of Stimuli. Our behavioral results suggested that subjects used a combination of available cues, instead of direct to reverberant ratio (D/R) or interaural level difference (ILD) alone, to discriminate sound-source distances. However, an additional acoustic analysis of the present stimuli was conducted to examine whether either one of the individual cues, considered separately, provided information consistent with the observed distance discrimination. This analysis was also necessary because, unlike D/R, ILD also contributes to direction discrimination, thus suggesting an alternative hypothesis that our main results reflect a partial byproduct of ILD-specific direction neurons, instead of populations encoding distance, per se.

This acoustic analysis of D/R and ILD is presented in Fig. S2. To generate the graphs, the BRIRs used in the simulations were first divided into their direct-sound part, corresponding to the sound that a listener would hear in an anechoic environment, and their reverberant part, corresponding to the reflections heard in the room. Then, the complete BRIRs as well as the separate impulse responses were convolved with random noise tokens corresponding to the stimuli. The noises were either broadband stimuli, as used in the experiments, or one-third octave-filtered narrowband stimuli, with the filters centered at 450 Hz or 6,500 Hz, marked in Fig. S2 as Broad, Low, and High, respectively. Root-mean-square energy was computed in the convolved stimuli and differences in this energy were plotted as D/Rs or ILDs. D/Rs were computed separately for the near ear (ipsilateral to the simulated target location) and far ear (contralateral to the target). ILDs were computed either for the total stimulus (including direct and reverberant part) or for the direct part only.

Fig. S2, *Left* shows that, as expected, the D/Rs tend to decrease as the stimulus moves away from the listener. The decrease is much steeper at the near ear than at the far ear (compare the thick solid and dotted lines). This difference in steepness confirms that the near-ear D/R is a more informative cue than the far-ear D/R. At the near ear, this decrease is approximately linear, independent of the stimulus frequency/bandwidth (compare the three red solid lines in the D/R panel). Furthermore, although the absolute D/R differs across frequencies, the slopes of three D/R graphs for the near ear are very similar to each other, changing linearly and spanning ~20 dB as the stimulus moves from 15 to 100 cm. That is, a particular relative change in the sound-source distance (say, 50 vs. 100 cm) is represented by a consistent relative D/R change at all frequencies.

The right-hand panel in Fig. S2 shows the results for the ILDs. The pattern of results is more complex than for the D/Rs. First, as with D/Rs, the ILDs decrease as the stimulus moves away from the listener. However, unlike with D/R, this decrease is nonlinear as

a function of stimulus distance (steeper for near than far sources) and it is steeper at high frequencies (spanning ∼25 dB, symbols "Δ") than at low frequencies (spanning only about 15 dB, symbols "∇"). Similarly, the absolute values of the ILDs are larger at high frequencies than at low frequencies. Another important aspect is revealed by the comparison of the total and direct broadband ILDs (compare the solid and dotted lines). Whereas reverberation has a very small effect on ILD for very near sources, it results in a larger reduction of ILDs for distant sources.

In summary, these results show that, if distance perception were based on a single cue, the near-ear D/R would be a more parsimonious cue in reverberation than ILD. This is the case because there is a direct linear relationship between D/R and distance, and the slope of this linear relationship is similar across frequencies. Thus, one can always perform a relative distance judgment on the basis of D/R, independent of the stimulus frequency content or the distances considered. This result is not surprising, as the reverberant field is approximately constant for a stimulus presented at a fixed level, whereas the direct sound level is simply inversely proportional to the stimulus distance. For example, the D/R panel in Fig. S2 shows that a decrease in D/R by ∼20 dB always corresponds to quadrupling of the stimulus distance. On the other hand, to perform a relative distance judgment on the basis of ILD, different ILD-to-distance relationships apply depending on stimulus frequency (ILD changes more at high frequencies), the reference distance (ILD changes more for nearby sources), and on the level of reverberation in a given room (the more reverberant the room, the more the ILD changes with distance). The dependence of ILDs on frequency, distance, and reverberation level is particularly important from the neurophysiological perspective, as it contradicts the scenario that at the final processing stages sound distances are estimated by a fixed set of neurons tuned specifically to ILD only. If ILD were the dominating cue, there would have to be a complex arrangement of frequency-specific, distance-specific, and reverberation-level–specific neurons tuned to a variety of ILDs. In other words, although the average ILD of a broadband sound (thick black line in Fig. S2) decreases monotonically, the lack of consistency across frequencies, distance, and reverberation level makes it unlikely that such a cue could be used, consistently and reliably, for distance discriminations, particularly because much better cues related to the D/R are concurrently available in natural listening situations. Finally, as also noted in the main results, the discrimination sensitivity for ILD, on the basis of Weber's law, would differ from the present observations.

Predictions of Distance Discrimination. The distance discrimination analysis of the behavioral experiment found that performance was independent of the absolute distance of sources as long as the relative distance was fixed (e.g., the listener's ability to discriminate 25- vs. 50-cm sources was the same as the ability to discriminate 50- vs. 100-cm sources; Fig. 2). That is, subjects' ability to discriminate simulated distances was predicted by Weber's law, a principle that has been well documented to apply to many aspects of human perceptual sensitivity. This would be consistent with an idea that the combination of available cues, constituting the building blocks of a sound-distance percept, was used by the subjects. The purpose of this supporting discussion and analysis is to discuss predictions of how subjects would have performed, as predicted by Weber's law, if discriminations had been based on either D/R or ILD alone.

Kopčo and Shinn-Cunningham (3) performed a computational acoustic analysis of a behavioral experiment in which absolute distance responses were reported and concluded that all of the results of that study could be explained by a model that assumed that the listeners only used the D/R cue for their judgments. For

the current data this model would predict that the sensitivity in distance discrimination is dependent on only (*i*) the rate with which the D/R varies as a function of the source distance and (*ii*) the perceptual sensitivity (just noticeable difference, JND) to variations in D/R. The dependence of D/R on distance is linear (as shown in Fig. S2). The perceptual sensitivity to changes in D/R has only been studied by Larsen et al. (4) who found that it is fairly nonlinear for broadband stimuli as used here. Specifically, Larsen et al. measured D/R JNDs at baseline D/Rs of −10, 0, 10, and 20 dB, and found JNDs of ∼3 dB at 0 and 10 dB and JNDs of 6–8 dB at −10 and 20 dB . Given that the rate of change in D/R with distance is constant and considerably larger than 10 dB, except for the most distant sources (Fig. S2), this pattern of JNDs suggests that the listeners should be more sensitive to changes in distance at the largest distances examined here (in particular when the low-frequency channels are considered). However, no such difference in distance discrimination performance was observed here.

A similar model can be used to predict discrimination performance if subjects only used ILDs. Fig. S2 shows that the rate of change in ILDs with distance is highest for nearby sources. Because ILD JNDs are approximately constant at 1 dB (5), the best distance discrimination performance based on ILDs would be expected at the nearest distances, which is not consistent with our main results (Fig. 2).

Taken together, these results suggest that, whereas it may still be the most parsimonious assumption that performance is based on D/Rs, the process by which the brain computes distance estimates is more complex than a simple D/R-to-distance mapping. For example, D/R information might be combined across multiple channels or ILD information might also be used. Future studies will have to be performed to describe the mechanism of distance processing in more detail.

Supporting fMRI Experiment. The monaural D/R at the ear closer to the source (near ear) has been suggested to be the main intensity-independent distance cue used for nearby sources in reverberation (3). Therefore, a supporting experiment was performed in which presentation intensity was normalized such that the overall energy received at the near ear was constant, independent of the source distance. It was expected that, after the normalization, the intensity-independent D/R cue would dominate distance processing whereas activations related to variations in intensity would be minimized. The supporting experiment was identical to the main fMRI experiment, except that only two types of stimuli were used: constant vs. a near-ear normalized version of the varying distance stimulus (Fig. 1*C*). When D/R and ILD were varied randomly and intensity at the near ear was held constant, we observed a more widespread activation of auditory cortices (red-yellow in Fig. S3*A*) than in the main experiment. The observed activations are consistent with our interpretation that, compared with the main experiment, near-ear intensity normalization does not eliminate all intensity variations (such as the far-ear overall intensity variation, near-ear/far-ear direct-sound intensity variation, and reverberant-sound intensity variations). More specifically, as predicted by our hypothesis, the strongest "extra" activations were observed in the right hemisphere, that is, contralateral to the far ear for which the overall intensity was not normalized (Fig. S3*A*, *Right*). Nevertheless, in the hemisphere contralateral to the stimuli, the voxel with strongest activation was located in planum temporale (PT), in the posterior nonprimary auditory cortex (surface Talairach {*x, y, z*} = {−50, −30, 8}), consistent with our main result and the predictions of the dual pathway model of auditory cortex (6–10) (Fig. S3*B*).

1. Zahorik P (2002) Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111:1832–1846.
2. Durlach NI, Braida LD (1969) Intensity perception. I. Preliminary theory of intensity resolution. *J Acoust Soc Am* 46:372–383.
3. Kopčo N, Shinn-Cunningham BG (2011) Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am* 130:1530–1541.
4. Larsen E, Iyer N, Lansing CR, Feng AS (2008) On the minimum audible difference in direct-to-reverberant energy ratio. *J Acoust Soc Am* 124:450–461.
5. Hartmann WM, Constan ZA (2002) Interaural level differences and the level-meter model. *J Acoust Soc Am* 112:1037–1045.
6. Rauschecker JP, Tian B, Hauser M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 268:111–114.
7. Rauschecker JP (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol Suppl* 532(Supplement):34–38.
8. Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci USA* 97:11800–11806.
9. Rauschecker JP (1998) Cortical processing of complex sounds. *Curr Opin Neurobiol* 8:516–521.
10. Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293.

Fig. S1. Individual listeners' sensitivity to distance changes. Each bar represents one subject's estimate of the noise (i.e., variance) in his/her internal representation of sound-source distance on the basis of a decision-theory model fit to that subject's individual percentage of correct data from the behavioral experiment. Performance of subjects S4 and S12 was considerably worse than for the remaining subjects. As an example, on the basis of these sensitivity estimates, subject S12 is expected to correctly discriminate the two most separated speakers only in less than 90% of trials, whereas the well-performing subjects would achieve accuracy of about 99%.

Fig. S2. Mean D/R (*Left*) and ILD (*Right*) as a function of distance for the stimuli used in this study. Cue values were computed either for the whole broadband stimulus (Broad), or in one-third octave bands centered at 450 Hz (Low) and 6,500 Hz (High). D/Rs are shown for the ear ipsilateral to the stimulus (Near Ear) or contralateral to the stimulus (Far Ear). ILDs were computed either for the whole stimulus (Total; including both the direct sound and reverberation) or for the direct sound only (Direct; corresponding to the ILD that would be obtained in an anechoic environment). Error bars represent SDs across random noise tokens used as stimulus.

Fig. S3. (*A*) Adaptation-fMRI data from the follow-up fMRI experiment presented on inflated cortical surface representations. Contrast shows activation for sounds varying in distance with intensity normalized at the near ear vs. nonvarying standard sounds. (*B*) Location of the strongest-activated voxel in the hemisphere contralateral to the stimulation direction. The voxel is very near to the significant activation cluster from varying distance vs. varying intensity contrast of the main experiment, presumably revealing intensity-independent distance-cue representations in the posterior nonprimary auditory cortices.