Auditory Distance Coding in Rabbit Midbrain Neurons and Human Perception: Monaural Amplitude Modulation Depth as a Cue

Duck O. Kim,¹ Pavel Zahorik,² Laurel H. Carney,³ Brian B. Bishop,¹ and Shigeyuki Kuwada¹

¹Department of Neuroscience, University of Connecticut Health Center, Farmington, Connecticut 06030, ²Division of Communicative Disorders, Department of Surgery, University of Louisville School of Medicine, Louisville, Kentucky 40292 and, ³University of Rochester, Rochester, New York 14642

Mechanisms underlying sound source distance localization are not well understood. Here we tested the hypothesis that a novel mechanism can create monaural distance sensitivity: a combination of auditory midbrain neurons' sensitivity to amplitude modulation (AM) depth and distance-dependent loss of AM in reverberation. We used virtual auditory space (VAS) methods for sounds at various distances in anechoic and reverberant environments. Stimulus level was constant across distance. With increasing modulation depth, some rabbit inferior colliculus neurons increased firing rates whereas others decreased. These neurons exhibited monotonic relationships between firing rates and distance for monaurally presented noise when two conditions were met: (1) the sound had AM, and (2) the environment was reverberant. The firing rates as a function of distance remained approximately constant without AM in either environment and, in an anechoic condition, even with AM. We corroborated this finding by reproducing the distance sensitivity using a neural model. We also conducted a human psychophysical study using similar methods. Normal-hearing listeners reported perceived distance in response to monaural 1 octave 4 kHz noise source sounds presented at distances of 35–200 cm. We found parallels between the rabbit neural and human responses. In both, sound distance could be discriminated only if the monaural sound in reverberation had AM. These observations support the hypothesis. When other cues are available (e.g., in binaural hearing), how much the auditory system actually uses the AM as a distance cue remains to be determined.

Key words: amplitude; localization; looming; modulation; personal space; sound

Introduction

Localization of sound source has been extensively studied both neurally in animals and psychophysically in humans (for review, see Middlebrooks and Green, 1991; Grothe et al., 2010). However, localization of sound source distance has been much less investigated than azimuth and elevation (for review, see Zahorik et al., 2005; Ahveninen et al., 2014). Localizing the distance of a sound source is important because avoiding a danger or understanding ongoing environmental events often requires recognizing sound source distances. A sound in a near-field (<1 or 2 m), peripersonal space (Holmes and Spence, 2004; Dramas et al., 2008; Guipponi et al., 2013), such as a looming sound (Maier et al., 2008), is particularly salient because the event may require the listener's immediate response (e.g., in a fight-or-flight situation) (Cannon, 1915). Furthermore, speech communication often

takes place in such a nearby space in reverberant environments and benefits from a listener's ability to localize distances of sound sources (Shinn-Cunningham et al., 2005).

Humans (Brungart et al., 1999; Zahorik, 2002a; Shinn-Cunningham et al., 2005; Kopčo and Shinn-Cunningham, 2011; Kopčo et al., 2012) and animals (Naguib and Wiley, 2001; Kuwada et al., 2015) can localize sound distance with varying accuracy in different conditions. Particularly, localization of sound distance is better in reverberant than in anechoic conditions (Mershon et al., 1989; Nielsen, 1993; Kolarik et al., 2013). When the distance between a sound source and a listener is varied in a reverberant environment, the direct signal energy changes, whereas the reverberant energy remains approximately constant. This fact led to the suggestion that direct-to-reverberant (D/R) energy ratio may mediate the ability of the auditory system to localize sound distance (Mershon and Bowers, 1979; Hartmann, 1983; Bronkhorst and Houtgast, 1999; Zahorik, 2002b; Kim et al., 2008; Larsen et al., 2008; Kolarik et al., 2013). However, how a varying D/R ratio may be converted into a distance-representing neural signal is not known. Here we propose a novel monaural mechanism whereby such neural coding of sound distance may be achieved.

Reverberation produces loss of amplitude modulation (AM) depth (Houtgast and Steeneken, 1985; Zahorik et al., 2012; Kim et al., 2013; Kuwada et al., 2014), and the modulation loss in-

Received Sept. 10, 2014; revised Jan. 21, 2015; accepted Feb. 22, 2015.

Author contributions: D.O.K., P.Z., L.H.C., B.B.B., and S.K. designed research; D.O.K., P.Z., L.H.C., B.B.B., and S.K. performed research; D.O.K., P.Z., L.H.C., B.B.B., and S.K. contributed unpublished reagents/analytic tools; D.O.K., P.Z., L.H.C., B.B.B., and S.K. analyzed data; D.O.K., P.Z., L.H.C., B.B.B., and S.K. wrote the paper.

The neural portion of this study was supported by NIH R01 DC002178, the modeling portion NIH R01 NIH DC010813, and the human psychophysics portion by NIH R21EY023767.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Duck O. Kim, Department of Neuroscience, University of Connecticut Health Center, 263 Farmington Avenue, Farmington, CT 06030-3401. E-mail: kim@neuron.uconn.edu.

DOI:10.1523/JNEUROSCI.3798-14.2015

Copyright © 2015 the authors 0270-6474/15/355360-13\$15.00/0

creases with distance as the D/R ratio decreases. This is one part of the proposed mechanism. The other part is that many neurons in the inferior colliculus (IC) in the midbrain are sensitive to AM depth; some neurons increase firing rates with AM depth, whereas others decrease rates (Krishna and Semple, 2000; Joris et al., 2004; Nelson and Carney, 2007). We hypothesize that a combination of these two parts constitutes a novel mechanism for IC neurons' ability to code the distance of sound sources. We tested this hypothesis both neurally in the unanesthetized rabbit IC and psychophysically in humans. We address here monaural processing of sound distance because understanding monaural processing is a good initial step toward an eventual understanding of the more complex binaural processing of sound distance. We present neural and psychophysical evidence that supports the hypothesis.

Materials and Methods

The neural portion of this study was approved by the University of Connecticut Health Center Animal Care Committee and was conducted according to the National Institutes of Health guidelines. The human psychophysics portion of this study was approved by the University of Louisville Institutional Review Board.

Virtual auditory space and neural recording methods. The procedures for neural recording in the present study were the same as those reported by Kuwada et al. (2014). Briefly, extracellular action potentials were recorded in the right IC of unanesthetized Dutch-Belted rabbits with custom-made tungsten-in-glass microelectrodes. We rejected neural responses if the observed response contained very short interspike intervals (<0.95 ms) constituting >1% of the total sample (similar to Slee and Young, 2013).

Head-related impulse responses (HRIRs) were measured with the unanesthetized rabbit positioned in an anechoic chamber that met the specifications of ISO 3745 (1977) and had the inner walls lined with fiberglass wedges designed to be anechoic for low frequencies down to 110 Hz. The chamber's free inner dimensions were $9 \times 4 \times 4$ m. Binaural room impulse responses (BRIRs) were measured with the rabbit positioned in a highly reverberant chamber that had hard walls and inner dimensions $6.5 \times 5.7 \times 5$ m. This chamber's reverberation time (T60), averaged for 1 octave bands centered between 0.25 and 10 kHz and spaced at 1 octave intervals, was 2.2 s. The procedures for measuring HRIRs were described previously (Kim et al., 2010). The procedures for measuring BRIRs and HRIRs shared common elements, such as a point sound source, and a blocked-meatus method of recording signals with miniature microphones (Knowles FG-3329) embedded in ear-mold tips placed deep inside the rabbit's ear canals. The main difference is that the length of a logarithmically swept (0.05–49 kHz) chirp source sound was considerably longer, 3.6 s, for BRIRs than for HRIRs, 0.67 s. The reason for this difference is that the impulse response of a reverberant room decays much more slowly than the anechoic counterpart. The impulse responses were measured at 9 distances with an equal spacing on a logarithmic scale (10, 14, 20, 28, ..., 160 cm), and at 21 azimuths ($\pm 150^\circ$ in 15° steps) all at 0° elevation.

As source signals, we used two types of 1 octave band noise: unmodulated noise and AM noise. The carrier center frequency (CCF) was set at the neuron's best frequency. Different noise tokens were used in testing different neurons. The modulation envelope was either a sinusoid or a "raised-sine" (Bernstein and Trahiotis, 2010). We express AM depth in decibels (dB) as follows:

$$AM_depth_dB = 20 * \log_{10}((AM_depth_)/100). \quad (1)$$

At the source, AM depth was 0 dB (i.e., 100%). The modulation frequency was varied logarithmically from 2 to 512 Hz in octave steps. We kept the modulation frequency lower than 12% of the center frequency of the carrier noise band. The 12% corresponds to the difference between the upper cutoff and center frequency of a third octave band filter that approximates auditory filtering (Moore, 1997). Each source signal consisted of a 1000 ms noise burst (4 ms rise/fall, raised-cosine gate) and an 800 ms silence. VAS stimuli were created by convolving the source signals

Figure 1. Schematic drawing of a monaural auditory-brainstem model. AN, Auditory nerve fiber; IC BE, inferior colliculus band-enhanced neuron; IC BS, inferior colliculus band-suppressed neuron; IN1, interneuron 1; IN2, interneuron 2. The AN activities were simulated with the model of Zilany et al. (2014). The present model was adapted from Nelson and Carney (2004) and L.H. Carney (personal communication).

Table 1. Parameter values of the monaural auditory brainstem model described in Figure 1^a

Synapse no.	Strength	τ (ms)	Delay (ms)	Polarity	Target	Origin
1	1.6	1.6	NA	Excitatory	CN	AN
2	0.9	2.0	1.0	Inhibitory	CN	IN1
3	3	1.0	NA	Excitatory	IC BE	CN
4	3.3	3.0	2.0	Inhibitory	IC BE	IN2
5	3	1.0	NA	Excitatory	IC BS	CN
6	5	3.0	2.0	Inhibitory	IC BS	IC BE

^aThe auditory-nerve (AN) inputs were high-spontaneous-rate fibers simulated using the model of Zilany et al. (2014). The best frequency was specified for the AN fiber and inherited by higher-level cells. The IC band-enhanced (BE) cell model was based on the Same-Frequency Inhibitory-Excitatory model of Nelson and Carney (2004) with parameters based on cell "C" in their Figure 8. The present model is an extension of the Nelson–Carney model in that it includes the band-suppressed cell model. Each synaptic activation was represented by an α function with a time constant, τ . The "delay" value represents the relative delay between excitatory and inhibitory synaptic signals converging on a common target. BS, Band-suppressed; IN1 and IN2, interneuron 1 and 2; NA, not applicable.

with the rabbit's HRIRs and BRIRs for each sound-source location and each acoustic environment. These VAS stimuli were delivered monaurally to the left ear (contralateral to the recorded IC) through a custom cone enclosure that housed a Beyer DT-770 earphone coupled to a sound tube embedded in custom-fitted ear mold to form a closed system. The sounds were presented to the rabbit using TDT System 3 hardware and custom software written in MATLAB (MathWorks). From the impulse responses, we also derived the acoustic modulation transfer function (MTF) of the system that consists of the rabbit (pinna, head, and body), the sound source at a specific location, and the environment. The acoustic MTF represents modulation loss as a function of modulation frequency.

Neural firing rate versus distance functions were measured typically at the neuron's best azimuth and at additional azimuths, time permitting. We converted the absolute firing rate in spikes/s into a normalized firing rate such that 0% and 100% corresponded to the minimum and maximum firing rates across all conditions tested for the neuron. Thus, a 100% change of normalized firing rate tended to correspond to a change between the responses to an unmodulated stimulus and to a fully modulated (AM depth = 100%) stimulus at a close distance, averaged for the anechoic and reverberant environments. We normalized firing rate because it facilitates comparing different neurons, and also comparing the models with neurons.

Neural modeling. To facilitate the understanding of how the observed neural distance sensitivity may arise, we used a model for IC neurons that was adapted from Nelson and Carney (2004) and L.H. Carney et al.

(personal communication). A schematic neural circuit for the model is shown in Figure 1. The auditory nerve responses were simulated with the model of Zilany et al. (2014). The Nelson and Carney (2004) model achieves band-enhanced rate MTFs by combining the dynamics of fast excitatory input with a relatively large, slower and delayed inhibitory input. Inputs with these dynamics and amplitudes are convolved with the time-varying rates of the cochlear nucleus (CN) inputs to the model IC neurons. Each excitatory and inhibitory synaptic input was described by an α function (Jack et al., 1975), with a relative strength (corresponding to the area of the α function), delay, and time constant (Table 1). The present model and the L.H. Carney et al. model contain a band-suppressed IC neuron in addition to the band-enhanced IC neuron. The band-suppressed model cell received a large, slow, and delayed inhibitory input from the band-enhanced model cell and fast excitatory input from the CN model cell (Fig. 1). The values of all model parameters are provided in Table 1.

Human psychophysical methods. Six subjects (4 female, 2 male) participated in the experiment. All had audiometrically verified normal hearing (pure tone, air-conductive thresholds ≤ 25 dB HL from 125 to 8000 Hz). Subject age ranged from 18.1 to 22.5 years (median age, 20.7 years). Subjects received course credit for participation in the experiment.

Listeners were presented with sounds at different simulated source distances and asked to estimate the egocentric distance of each sound, using methods broadly similar to those described by Zahorik (2002a). Virtual auditory space techniques were used to represent sound field listening to sources at distances ranging from 0.35 to 2.0 m. The BRIRs were simulated using techniques described previously (Zahorik, 2009) and used nonindividualized HRIRs measured from a fixed distance of 1.4 m in anechoic space. Two types of sound fields were simulated: anechoic and reverberant (room volume: 500 m³, approximate broadband T60 = 3 s). The source incidence angle was 90° to the listener's right, at ear level. The source carrier signal was a 1 octave band of noise with a CCF of 4 kHz, 2 s in duration (500 ms rise/fall raised-cosine gate). In certain conditions, sinusoidal AM was imposed on this signal (100% AM depth) at a frequency of 32 Hz. With the selected CCF and modulation frequency, loss of AM depth across distance was substantial and systematic for the VAS stimuli used for the human study (data not shown).

To limit listener's use of level cues in performing the distance estimation tasks, two types of level controls were implemented. First, sound pressure level was equalized for distance by adjusting the gain of the simulated source to compensate for the 6 dB loss per distance doubling observed in anechoic space. Additionally, sound pressure level was randomly varied (roved) over ± 6 dB from trial to trial. Listeners were also explicitly instructed to ignore any loudness differences between trials. The sound was presented monaurally to the ear facing the sound source. Listeners provided 10 estimates for each target distance (presentation order randomized).

Results

We evaluated the hypothesis that the following novel mechanism creates a neural sensitivity to auditory distance: a combination of a neural sensitivity to AM depth of a sound and distance-dependent loss of AM depth in reverberation. This concept is illustrated in Figure 2 with responses of an example IC neuron, designated as Neuron 1. The solid line of Figure 2A displays normalized firing rate of Neuron 1, under anechoic conditions, versus modulation frequency (rate MTF). The firing rate of this neuron was enhanced when the sound was modulated (Fig. 2A, solid line) compared with the unmodulated stimulus condition (Fig. 2A, "◇" on the y-axis). The enhancement was strong in a band of modulation frequencies wherein the rate MTF showed a peak. Accordingly, we refer to this class of neurons as band-enhanced neurons. At the best modulation frequency, Neuron 1's firing rate increased monotonically with AM depth (Fig. 2B, solid line).

When the sound-source distance increased in the presence of reverberation, the firing rate to the modulated sound decreased (Fig. 2C, solid magenta), whereas that to the unmodulated sound remained nearly constant at low rates (Fig. 2C, solid blue). To

Figure 2. *A*, Normalized firing rate of an example band-enhanced IC neuron (Neuron 1, BF = 2.5 kHz) versus modulation frequency (solid line), and the corresponding response of the band-enhanced model (dash-dot line) to the same stimuli. SAM modulation depth, 100%. Data symbols (◇ and ○) on the y-axis represent the neural and model responses to unmodulated noise, respectively. For Neuron 1, minimum and maximum rates, 43 and 114 spikes/s; for model, 35 and 130 spikes/s. *B*, Normalized firing rates of the neuron and the model versus modulation depth in dB relative to 100%. Modulation frequency, 64 Hz. *C*, *D*, Normalized firing rates of the neuron and the model versus distance in the reverberant and anechoic environments. Magenta curves indicate responses to SAM noise, 64 Hz, 100% depth. Blue curves indicate responses to unmodulated noise.

reveal neural distance sensitivity that is independent of stimulus level, we kept the stimulus level essentially constant across distance. In the anechoic condition, the firing rate remained nearly constant across distance regardless of whether the sound was modulated (Fig. 2D, solid lines). These results demonstrate that the neuron's sensitivity to sound distance required both reverberation and AM of the sound.

To ascertain whether a plausible neural circuit may account for the observed neural sensitivity to sound distance, we used a model for the AM tuning of IC neurons (see Materials and Methods; Fig. 1). We set the best frequency of the band-enhanced model equal to that of Neuron 1 (2.5 kHz) and applied the same VAS stimuli as those used for Neuron 1. Normalized firing rates of the model are shown in Figure 2 as dash-dot lines. The model exhibited similar response features to those of Neuron 1. The similarity between the model and the neuron's responses supports the following: (1) neural distance sensitivity can arise from a combination of distance-dependent loss of AM depth and the neural sensitivity to AM depth; and (2) the synaptic mechanisms (i.e., dynamic interactions between fast excitatory and slow inhibitory synaptic potentials) assumed by the model of Figure 1 are a viable hypothesis, as are other models for IC neurons that exhibit modulation-depth sensitivity (Davis et al., 2010).

The effect of reverberation on a sound depends on the distance between the sound source and the listener's ear. One can

Figure 3. Comparison of predicted (blue) and actual (magenta) responses of Neuron 1 (described in Fig. 2C) versus distance. The prediction was based on the acoustic loss of AM depth (green) and the rate-AM depth function shown in Figure 2B.

Figure 4. Linear regression (green) of the normalized firing rate of Neuron 1 versus log-distance superposed on the actual response. The regression was applied to a region of distances where the actual response stayed within 10% from the regression line. From this result, we derived response range, distance range, slope, and correlation coefficient.

predict the neuron's distance sensitivity from the following: (1) the distance-dependent acoustic loss of AM depth and (2) neural sensitivity to AM depth. Such a prediction is provided in Figure 3 for the response of Neuron 1. Figure 3 (green curve) describes increasing loss of AM depth in the acoustic stimulus across distance. The AM envelope of the stimulus was extracted by performing the Hilbert transform of the acoustic stimulus signals. Figure 3 (blue and magenta curves) shows the predicted and actual responses of Neuron 1, respectively. The prediction reproduced the salient feature of the response (i.e., a decrease of response with distance).

Our approach to quantifying a neuron's distance sensitivity is illustrated in Figure 4 using the responses of the example neuron in Figures 2 and 3 (Neuron 1). We performed a linear regression of the normalized firing rate versus log-distance for each neu-

Figure 5. Normalized firing rates of four additional example band-enhanced IC neurons (Neurons 2–5) versus distance in the reverberant and anechoic environments. AM depth, 100%; modulation frequency, 32 Hz for Neuron 2 and 64 Hz for the rest. Modulation type was raised-sine with exponent 8 for Neuron 3 and sinusoidal for the rest.

ron's response to modulated sounds in reverberation (Fig. 4, green curve). The regression was applied to a region of distances where the actual response differed by <10% from the regression line. We chose this procedure because neural responses often varied systematically over a limited range of distances with saturation outside the range. From this fitted regression line, we derived response range, distance range, slope, and correlation coefficient between the response and distance (Fig. 4). A good distance-coding neuron should have a large response range, a large distance range, and a high correlation between the response and distance. The example Neuron 1 exhibited these properties (Fig. 4).

There was variability among band-enhanced IC neurons regarding their rate-distance functions as illustrated by four additional band-enhanced neurons in Figure 5. The four neurons, with BFs ranging from 1.6 to 10.1 kHz, all showed normalized firing rates that decreased with distance when the stimulus sound had AM in reverberation (left column), whereas the responses remained approximately constant across distance in the anechoic

Table 2. Measures of rate-distance function for the five example band-enhanced IC neurons described in Figures 2 and 5

	BF (kHz)	Response range (%)	Distance range (doublings)	Slope (%/doubling)	Correlation
Neuron 2	1.6	64	4.0	−17	−0.995
Neuron 1	2.5	67	3.2	−23	−0.991
Neuron 3	3.2	55	3.7	−16	−0.988
Neuron 4	4.0	80	4.0	−18	−0.975
Neuron 5	10.1	23	3.1	−9	−0.969

condition regardless of whether the sound was amplitude modulated (right column). The response and distance ranges, slope, and correlation coefficient of these four neurons along with those of Neuron 1 are listed in Table 2. The absolute values of the correlation coefficients of all five neurons were high (>0.96), indicating that these neurons were able to represent sound distance. Neuron 5 with the highest BF (10.1 kHz) showed a smaller response range, a shallower slope, and a lower correlation coefficient than those of the other neurons. The four neurons with BFs of 1.6 to 4.0 kHz showed large response ranges ($\geq 55\%$), large distance ranges (≥ 3.2 doublings), and considerable slopes ($\leq -16\%$ /doubling).

To visualize how AM depth was degraded in reverberation at different distances and with different CCFs of 1 octave noise, we display in Figure 6 the acoustic waveforms that led to the responses of Neuron 1 to Neuron 5. The columns are ordered such that CCF increased from left to right. The AM envelope was clearly visible at 10 cm (Fig. 6, top row) for all five CCFs. However, the AM envelope became increasingly more obscure as distance increased beyond 10 cm. The modulation loss tended to start at closer distances for lower CCF such that the AM envelope was degraded at 80 cm for CCFs of 1.6 and 2.5 kHz but clear even at 160 cm for 10.1 kHz. To help discern the dependence of acoustic AM depth in reverberation on sound distance and CCF, we show a color contour plot describing this relationship in Figure 7. The loss of AM depth was greater at farther distances and lower CCFs with a nonmonotonic pattern in a region surrounding 160 cm and ~ 1 kHz showing the greatest loss of AM depth. These acoustic properties predict that neurons with high BFs (>6.3 kHz) would show lower distance sensitivities than those with lower BFs. A shallow slope and a small response range of Neuron 5 (BF = 10.1 kHz) is consistent with this prediction. In the anechoic condition, AM depth was close to 0 dB (i.e., 100%) for all distances and CCFs examined (data not shown).

Another class of IC neurons has responses to amplitude modulated sounds that are opposite to those of band-enhanced neurons. Responses of such a neuron (designated as Neuron 6) are shown in Figure 8 using the same format as Figure 2. Figure 8A, B shows firing rate versus modulation frequency and AM depth, respectively. This neuron's rate is suppressed by amplitude modulated sounds as indicated by the fact that the neuron's response to the unmodulated noise (Fig. 8A, B, " $\diamond$ " on the y-axis) is higher than the rates to the modulated sounds. The suppression was strong in a band of modulation frequencies wherein the rate MTF showed a trough. Accordingly, we refer to this class of neurons as band-suppressed neurons. The neuron's firing rate decreased monotonically with AM depth (Fig. 8B, solid line). When the sound-source distance increased in reverberation, the firing rate to the modulated sound increased (Fig. 8C, solid magenta), whereas that to the unmodulated sound remained nearly con-

stant at low rates (Fig. 8C, solid blue). In the anechoic condition, the firing rate remained nearly constant across distance regardless of whether the sound was modulated (Fig. 8D). As in the band-enhanced neurons described above, the band-suppressed neuron in Figure 8 also required reverberation and AM of the sound for distance sensitivity.

Analogous to the band-enhanced model (Fig. 2), we used a model for band-suppressed IC neurons for the purpose of ascertaining whether a plausible neural circuit may account for the band-suppressed neuron's sensitivity to sound. We set the best frequency of the band-suppressed model equal to that of Neuron 6 (3.2 kHz) and applied the same VAS stimuli as those used for Neuron 6. The responses of this model are shown in Figure 8 in dash-dot curves. The model responses exhibit remarkably similar features to those of Neuron 6. As before, the close similarity between the model and the neuron's responses supports the following: (1) neural distance sensitivity arises from a combination of distance-dependent loss of AM depth and the neural sensitivity to AM depth; and (2) the synaptic mechanisms assumed by the model of Figure 1 are a viable hypothesis.

As shown in Figure 3 above, one can also predict a band-suppressed neuron's distance sensitivity from the following: (1) the distance-dependent acoustic modulation loss and (2) neural sensitivity to AM depth. Such a prediction is provided for Neuron 6 in Figure 9, which shows loss of AM depth in the acoustic stimulus across distance (green), the predicted (blue), and actual (magenta) responses of Neuron 6. The prediction reproduced the salient feature of the response (i.e., an increase of response with distance).

There was variability among rate-distance functions of band-suppressed IC neurons, analogous to that of the responses of band-enhanced neurons. This variability is illustrated with responses of four additional neurons in Figure 10. These four neurons, with BFs ranging from 2.0 to 8.0 kHz, all showed responses that increased with distance when the stimuli were amplitude-modulated sounds in reverberation (left column) whereas the responses remained approximately constant across distance in the anechoic condition, regardless of whether the sound was modulated (right column). Measures of rate-distance functions of the five band-suppressed neurons of Figures 8 and 10 are listed in Table 3. The correlation coefficients of all five were high (>0.98). All five neurons showed large response ranges ($\geq 61\%$), distance ranges ≥ 2.1 doublings, and slopes $\geq 15\%$ /doubling. Even with a high BF (8.0 kHz), Neuron 10 showed distance-coding properties that were comparable with the neurons with lower BFs.

How diverse are the distance-coding properties of IC neurons? This question is addressed using distributions of the four measures derived from rate-distance functions of our sample of 54 IC neurons (Fig. 11). The neurons were divided into two groups: low and mid BFs (0.3–6.2 kHz) and high BFs (6.3–16 kHz). For this purpose, we inverted the responses of the band-enhanced neurons and combined them with those of the band-suppressed neurons. The mean values of each measure are indicated by vertical arrows in each panel. The response range of the low-mid-BF group (Fig. 11A) was higher ($71.0 \pm 15.5\%$, mean $\pm$ SD) than that of the high BF group (mean = $50.1 \pm 14.1\%$; Fig. 11E). This difference was statistically significant ($t_{(52)} = 5.10$, $p < 0.001$; $\chi^2(8) = 20.5$, $p < 0.01$). Distributions of the distance range of the two BF groups (Fig. 11B, F) were irregular, and their difference was not significant ($p > 0.05$). Likewise, distributions of slope of the two BF groups (Fig. 11C, G) were also irregular, and their difference was not significant ($p > 0.05$). The

Figure 6. VAS acoustic stimulus waveforms in a rabbit's ear with the sound source at various distances in the reverberant environment. The stimuli were amplitude modulated (AM) 1 octave band noise signals at 5 different CCFs. These stimuli were used to obtain the responses of Neurons 1–5 described in Figures 2 and 5. AM depth, 100%; modulation frequency, 64 Hz, except for CCF of 1.6 kHz (32 Hz). Modulation type was sinusoidal, except for CCF of 3.2 kHz (raised-sine with exponent 8).

Figure 7. Color contour plot of acoustic AM depth in dB (with the color scale shown on the right) versus distance and CCF. Modulation frequency 64 Hz, azimuth -60° , modulation type sinusoidal. In addition to distance and CCF, AM depth also depends on modulation frequency, modulation type, azimuth, and reverberant environment.

correlation coefficient between the normalized rate and log-distance of the low-mid BF group (Fig. 11D) was higher (0.986 ± 0.009) than that of the high BF group (0.978 ± 0.014 ; Fig. 11H). This difference was significant in the t test ($t_{(52)} = 2.35$, $p = 0.02$) but not significant in the χ^2 test ($p > 0.05$). Overall, the results of Figure 11 support the view that band-enhanced and -suppressed IC neurons have varying degrees of ability to represent sound distance and that distance sensitivities of low-mid BF neurons are high than those of high BF neurons.

Strong reverberation is associated with a long reverberation time (T60) (Allen and Berkley, 1979; Beranek, 1986). T60 of an

acoustic environment depends, in general, on sound frequency (Nielsen, 1993; Zahorik, 2002a). The T60 of the reverberant chamber of the present study, averaged over the low-mid frequency band and the high-frequency band, were 2.5 s and 1.3 s, respectively. Furthermore, our acoustic analysis showed that loss of AM depth was small for high CCFs (Fig. 7). Thus, our finding that response ranges and correlation coefficients of the high-BF neurons were smaller than the low-mid BF counterparts (see Fig. 13) is consistent with the acoustics.

How well can IC neurons represent sound source distance individually and collectively, and what distances can they discriminate? To address these questions, we determined mean and SD of each neuron's normalized firing rate across distance and derived d' (Green and Swets, 1974) versus distance as follows:

$$d'(x) = (\mu(x) - \mu(x_0)) / (0.5 * (\sigma^2(x) + \sigma^2(x_0)))^{0.5}, \quad (2)$$

where x = sound source distance, x_0 = reference sound source distance, μ = mean of normalized firing rate, and σ = SD of normalized firing rate.

For this purpose, each neuron's response to two repetitions of 1000 ms noise bursts were divided into six 300 ms epochs by removing the initial 100 ms portion (to remove onset discharges) and dividing the remaining 900 ms portion into 300 ms epochs. Thus, mean and SD of a neuron's firing rate at each distance were derived from the six samples of firing rates.

Additionally, the firing rates of band-enhanced neurons were inverted and pooled with those of band-suppressed neurons so that the firing rates of both types of neurons would change in the same direction across distance. Among our total sample of 54 neurons, 35 neurons showed a correlation coefficient between d' and $\log(\text{distance ratio}) > 0.9$ for distance ratios of 1.0, 1.4, and 2.0 with a reference distance of 14 cm, indicating that their responses varied consistently with $\log(\text{distance ratio})$. The functions of d'

Figure 8. *A*, Normalized firing rate of an example band-suppressed IC neuron (“Neuron 6,” BF = 3.2 kHz) versus modulation frequency (solid line), and the corresponding response of the band-suppressed model (dash-dot line) to the same stimuli in the same format as Figure 2. SAM modulation depth, 100%. For Neuron 6, minimum and maximum rates, 28 and 86 spikes/s; for model, 60 and 108 spikes/s. *B*, Normalized firing rates of the neuron and the model versus modulation depth in dB relative to 100%. Modulation frequency, 128 Hz. *C, D*, Normalized firing rates of the neuron and the model versus distance in the reverberant and anechoic environments. Magenta curves indicate responses to SAM noise, 128 Hz, 100% depth. Blue curves indicate responses to unmodulated noise. Green curve indicates a linear regression line.

Figure 9. Comparison of predicted (blue) and actual (magenta) responses of Neuron 6 (described in Fig. 8C) versus distance. The prediction was based on the acoustic loss of AM depth (green) and the rate-AM depth function shown in Figure 8B.

Figure 10. Firing rates of four additional example band-suppressed IC neurons (Neurons 7–10) versus distance in the reverberant and anechoic environments. AM depth, 100%; modulation frequency, 64 Hz for Neuron 10 and 128 Hz for the rest. Modulation type was sinusoidal for Neuron 9 and raised-sine with exponent 8 for the rest.

Table 3. Measures of rate-distance function for the five example band-suppressed IC neurons described in Figures 8 and 10

	BF (kHz)	Response range (%)	Distance range (doublings)	Slope (%/doubling)	Correlation
Neuron 7	2.0	94	4.0	25	0.990
Neuron 8	2.5	61	4.0	15	0.986
Neuron 6	3.2	98	2.2	40	0.994
Neuron 9	4.0	70	4.0	21	0.985
Neuron 10	8.0	69	2.1	36	0.982

versus $\log(\text{distance ratio})$ of the 35 neurons were then fit with linear regression lines, and we rank-ordered them based on the slope of the regression line. As the slope and threshold distance ratio for a constant d' are inversely related, this rank ordering is the same as that based on threshold.

The procedure of deriving d' is illustrated in Figure 12 for two neurons that exhibited the highest and median slopes of d' versus $\log(\text{distance ratio})$. The top row shows mean and SD of normalized firing rates as functions of $\log(\text{distance ratio})$, in magenta for

Figure 11. Distributions of rate-distance measures. For this purpose, rate-distance functions of band-enhanced neurons were inverted and combined with those of band-suppressed neurons. Vertical arrows indicate the mean values. **A–D**, Low and mid (0.3–6.2 kHz) BF group: 32 neurons consisting of 18 band-enhanced and 14 band-suppressed. **E–H**, High BF (6.3–16 kHz) group: 22 neurons consisting of 10 band-enhanced and 12 band-suppressed.

Figure 12. Normalized firing rate ($\pm$ SD), top row, and d' , bottom row, versus log(sound distance ratio) for 2 neurons that showed the highest (left) and median (right) slopes of d' versus log(distance ratio) among a group of 35 neurons. Bottom row, Straight line corresponds to a linear regression of the data. Inset values indicate threshold distance ratio corresponding to a d' of 1.09 (dotted lines). Reference distance: 14 cm.

AM noise and in blue for unmodulated noise. The d' measure (bottom row) increased with distance ratio when the sound had AM but remained near zero when the sound was unmodulated. Discrimination threshold at 71% correct performance in a one-interval psychophysical test corresponds to d' of 1.09 (Macmillan and Creelman, 2005). The neurons' threshold distance ratios, defined this way, were 1.18 and 1.49 for the AM noise. These

Figure 13. Distribution of threshold distance ratio (corresponding to $d' = 1.09$) among 35 neurons. Sound: amplitude modulated 1 octave noise; reference distance: 14 cm.

results indicate that, when the AM noise is presented monaurally in the reverberant environment, individual IC neurons could discriminate sound distances with a separation greater than the thresholds. In contrast, when the noise was unmodulated, d' values remained negligible (Fig. 12 *B, D*, blue), indicating the inability of the neurons to discriminate sound distances. Thus, AM is essential for AM-depth-sensitive IC neurons to discriminate distances of a monaurally presented 1 octave noise in reverberation.

How diverse are IC neurons regarding their abilities to discriminate sound distances? This information is provided in Figure 13 in terms of distribution of threshold distance ratios among the 35 neurons. The most frequently observed threshold distance ratios were clustered ~ 1.4 with a median threshold distance ratio of 1.49.

Figure 14. *A, C*, Optimal d' versus log(sound distance ratio) for 1 (*A*) and 32 (*C*) neurons that were rank-ordered according to the slope of d' versus log(distance ratio). *B, D*, Zoom-in plots of each curve near the origin that describe the threshold distance ratio corresponding to $d' = 1.09$ (dotted lines). Sound: amplitude modulated 1 octave noise; reference distance: 14 cm.

How may the auditory system combine distance-conveying information provided by multiple neurons? Specifically, how may the auditory system combine d' measures of multiple neurons? The optimal decision theory (Siebert, 1970; Colburn et al., 2003) assumes independence among the neurons and predicts the following:

$$d'_{\text{opt}} = \left(\sum (d'_i)^2 \right)^{0.5} \quad (3)$$

where d'_{opt} and d'_i correspond to optimally combined d' and an individual neuron's d' , respectively. We used this prediction in obtaining d'_{opt} based on a varying number of neurons. Figure 14 (left column) describes d'_{opt} versus log(sound distance ratio) based on the highest-ranking 1, 4, and 32 neurons of our sample. Figure 14 (right column) displays the same information in the region near the crossing of the d'_{opt} function across the threshold line (d' of 1.09). The results indicate that the discrimination performance improved with increasing number of neurons with the optimal threshold distance ratio decreasing from 1.18 to 1.06 for the three cases examined. Further results with more cases of number of neurons displayed in Figure 15 indicate that the optimal threshold decreased more noticeably when the number of neurons were few (i.e., 1, 2, and 4) and the threshold decreased more slowly with log(number of neurons) when the number of neurons were greater (i.e., 16, and 32).

Are the rabbit IC neurons' sensitivities to sound source distance consistent with human listeners' abilities to localize sound source distance? To address this question, we measured human listeners' perception of sound source distance using VAS implementation of similar stimuli as used for the rabbit. Human listeners' responses are illustrated with those of one listener in Figure 16*A, B* in the form of perceived distance versus sound distance on log-log axes. The perceived distance in response to modulated sound tended to increase with sound distance (Fig. 16*A*, average response as magenta line), whereas the response to unmodulated sound tended to be independent of sound distance (Fig. 16*B*). We characterized the perceived distance versus sound distance with a linear regression fit in a log-log plot of the perceived and sound distances (Fig. 16*A, B*, green lines). This yielded the following:

Figure 15. Threshold sound distance ratio achieved by an optimal combination of individual neurons' d' across a varying number of high-ranking neurons that were ordered according to the slope of d' versus log(distance ratio). Sound: amplitude modulated 1 octave noise; reference distance: 14 cm.

Figure 16. *A*, Log-log plot of distance perceived by one human listener versus sound distance for amplitude modulated sound. Small symbols represent individual trials; magenta curve indicates mean values; blue curve indicates a fitting function, $y = k * (x/2 \text{ m})^a$, with $k = 2.0 \text{ m}$, $a = 0.23$ (dimensionless), where x and y correspond to sound-source and perceived distances in meters, and $(x/2 \text{ m})$ a normalized distance relative to 2 m. Stimulus: 1 octave noise centered at 4 kHz; sinusoidal amplitude modulation frequency 32 Hz, 100% depth. *B*, Same as *A*, except that the noise was unmodulated. *C*, The value of the "a" parameter of modulated sound versus unmodulated sound for 6 listeners. Star represents average of the listeners. *D*, Same as *C*, except for k in meters on log-log axes.

$$\log_2(y) = \log_2(k) + a \cdot \log_2(x/2 \text{ m}) \quad (4)$$

and

$$y = k * (x/2 \text{ m})^a, \quad (5)$$

where x and y corresponded to sound and perceived distances in meters, respectively, and $(x/2 \text{ m})$, a normalized distance relative

to 2 m. A similar analysis method has been used in several past studies of distance perception (e.g., Zahorik, 2002a).

The dimensionless parameter “*a*” corresponds to the slope, providing a measure of relative sensitivity of distance perception independently of an overall scale factor (the vertical position of the curve in the log-log plot). The intercept, $\log_2(k)$, corresponds to an overall scale factor of the perceived distance with the parameter “*k*” having units of meters. The “*a*” and “*k*” parameters derived from the responses of six listeners are shown in Figure 16C, D, respectively, for the modulated condition versus the unmodulated condition. In Figure 16C, D, the average value is shown with the “star” symbol. The number next to each data point is an identifier. Listener 1 corresponds to the one represented in Figure 16A, B. In Figure 16C, except for one listener, the data points were above the diagonal line, indicating that the value of “*a*” was higher (i.e., higher distance sensitivity) in the modulated case than in the unmodulated case. This difference in “*a*” was significant in a paired *t* test ($t_{(5)} = 2.91$, $p = 0.033$). In contrast, the “*k*” values fell close the diagonal line. The difference in $\log_2(k)$ was not significant in a paired *t* test ($t_{(5)} = 0.25$, $p = 0.8$).

One may wonder how the variability seen among the present group of 6 listeners compares with what is reported in other studies of human distance perception. For this purpose, we compared the present study with a larger study that tested perceived distances of 62 listeners in response to binaurally presented wide-band VAS sounds (Anderson and Zahorik, 2014). The latter found SD of “*a*” to be 0.30, whereas the present study’s counterpart was smaller, 0.10. When we express the variability of the intercept of the two studies in the same way as the present study, SD of $\log_2(k)$ in the study of Anderson and Zahorik (2014) and present study were 1.1 and 0.82, respectively. Thus, the variabilities in both the slope and intercept measures of the present study were smaller than those of the larger study.

To verify that the observed pattern was significant, we performed a two-way repeated-measures ANOVA on the present human responses with factors of log-transformed sound distance (6 distances) and stimulus condition (modulated/unmodulated). The dependent variable in this analysis was log-transformed perceived distance. A statistically significant interaction between distance and stimulus condition was found ($F_{(5,25)} = 2.71$, $p = 0.043$). Follow-up simple-effects testing confirmed that the interaction resulted from a significant distance effect for the modulated signal ($F_{(5,25)} = 7.094$, $p < 0.001$), but not for the unmodulated signal ($F_{(5,25)} = 0.782$, $p = 0.573$). This finding supports the view that humans can perceive different distances of the sound monaurally if the sound has AM but not if the sound has no AM.

We also analyzed the human listeners’ responses in terms of *d'* analogous to the *d'* analysis of the neural responses shown above. Figure 17A shows *d'* versus sound distance for Listener 1 with a reference distance of 35 cm for the modulated sound. Figure 17B shows average *d'* of the 6 listeners ($\pm$ SEM) versus sound distance for the modulated sound. The values of *d'* (Fig. 17A, B) generally increased with distance. In contrast, when the sound was unmodulated, *d'* remained close to zero across sound distance (Fig. 17C, D). A *d'* of 1.09 was obtained at a distance ratio of 5.3, a much higher threshold than the rabbit optimal neural threshold distance ratio for 32 neurons, 1.06. The optimal neural threshold represents a model prediction of a lower bound of threshold that can be achieved when the sampled neurons are independent (Colburn et al., 2003). Because IC neurons are components of interconnected neural networks, their responses properties may

Figure 17. *A*, One human listener’s *d'* versus sound-source distance for modulated noise with a reference distance of 35 cm. The *d'* was derived from the data in Figure 16A. *B*, Same as *A*, except that *d'* was averaged for 6 listeners. Error bar = SEM across 6 listeners. *C*, *D*, Same as *A* and *B*, respectively, except that the noise was unmodulated.

not be independent. To that extent, the actual threshold achieved by the sampled neural population would be higher than the predicted optimal threshold. Furthermore, the optimal decision theory assumes no loss of information. Actual behavioral thresholds should always be worse than the optimal neural predictions based on IC neurons because more processing is needed between IC neurons and behavior and because further processing comes with loss of information (Blahut, 1987; Sinanovic and Johnson, 2000). Additional possible explanations for the difference between the neural and human thresholds include the following: (1) a species difference (see further comments in Discussion), (2) a distance estimation task is more difficult than a distance discrimination task, and (3) only in the human testing, sound pressure level was randomly varied (roved) over ± 6 dB from trial to trial. Despite this difference, the human listeners’ responses exhibit parallels to those of rabbit neurons (Fig. 12). That is, in both, sound source distance of a monaurally presented 1 octave noise in reverberation can be discriminated only if the sound is amplitude modulated.

Discussion

There were three main findings in this study: (1) Band-enhanced and band-suppressed IC neurons of the rabbit exhibited monotonic relationships between firing rates and distance of a monaurally presented 1 octave band noise when two conditions were met: (a) the sound had AM, and (b) the environment was reverberant. (2) A model comprising excitatory and inhibitory synapses in a neural circuit of the monaural auditory pathway between the auditory nerve and IC was able to reproduce the AM and distance coding properties of the two types of IC neurons. (3) Human distance localization performance to monaural 1 octave 4 kHz sounds showed parallels to the rabbit IC neurons (i.e., in both) sound distance could be discriminated only if the monaural sound in reverberation had AM. These observations provide evidence in support of a novel mechanism for monaural distance coding (i.e., a combination of neural sensitivity to AM depth and distance-dependent loss of AM depth in reverberation). When other cues are available (e.g., in binaural hearing), how much the

auditory system actually uses the AM as a distance cue remains to be determined.

In order for the proposed mechanism to operate, sounds must have AM, as is the case for many natural sounds (Singh and Theunissen, 2003; McDermott and Simoncelli, 2011), including speech (Shannon et al., 1995), and the AM depth of the source sound must be known. These requirements would be met when the source sound is a natural sound familiar to a listener. Neural sensitivity to AM depth is the other requirement for the proposed mechanism. Previous studies found that rate MTFs of IC neurons include bandpass, low-pass, and band-reject types (Krishna and Semple, 2000; Joris et al., 2004; Nelson and Carney, 2007). The present band-enhanced neurons correspond to the bandpass type, and the band-suppressed neurons include the band-reject and low-pass types. The findings that the rates of band-enhanced and -suppressed neurons increase and decrease, respectively, with AM depth fulfills the requirement of the proposed mechanism. Krishna and Semple (2000), and we found that rate MTFs of some IC neurons have a mixture of enhancement and suppression in different regions of modulation frequency. We refer to this as a hybrid type. We found that the vast majority of IC neurons fall into one of the three MTF types: band-enhanced, band-suppressed, and hybrid. Thus, the proposed distance coding mechanism can operate in the bulk of neurons.

The literature on neural representation of sound distance is limited. In a ground-breaking study, Graziano et al. (1999) found that many neurons in the ventral premotor cortex of the macaque monkey represented nearby (<30 cm) auditory distance by means of level-independent cue(s). In a study of functional imaging of human brain, Kopčo et al. (2012) found that specific areas of the brain were activated by sound distance-representing cues, such as D/R ratio and interaural level difference, that were independent of level. Neural coding of a looming (i.e., approaching) auditory and multisensory stimulus is a related subject (Ghazanfar et al., 2002; Hall and Moore, 2003; Guipponi et al., 2013). Whether/how the above mentioned cortical regions are interconnected and how they acquire distance sensitivities remain to be determined.

In the brainstem, Jones et al. (2013) found that IC neurons could represent ILDs present in near-field sound distances. However, that study did not directly address whether IC neurons could represent sound distance. The present study is the first to describe a level-independent representation of sound source distance in subcortical neurons. How information about sound distance is transmitted from the brainstem auditory pathways to the distance-sensitive cortical regions remains to be investigated.

How can the auditory system make sense of the outputs of the band-enhanced and -suppressed neurons that change their firing rates in opposite directions across distance? This situation is analogous to two opposing types of neurons in various neural systems (e.g., on and off cells in the visual system) (Hurvich and Jameson, 1957; Schiller, 1992), and opponent processes in the motivation system (Solomon and Corbit, 1974). Like these systems, the combined activities of the opposing auditory neurons may lead to better representation of sound distance. For example, if the enhanced and suppressed midbrain neurons provide excitatory and inhibitory inputs onto a neuron in the medial geniculate body, such a thalamic neuron may show increased distance sensitivity. This type of circuit is compatible with the observed inhibitory and excitatory colliculo-geniculate projections (Winer et al., 1996; Oliver, 2005).

Although the model described in Figure 1 reproduced salient features of IC neurons' distance sensitivities, one aspect of the model response was different from the responses of band-

enhanced IC neurons: that is, the model response to the modulated stimulus in the anechoic condition (Fig. 2D, dash-dot magenta) was higher than the maximum response in reverberation (Fig. 2C, dash-dot magenta), whereas the two response measures in band-enhanced IC neurons were rather similar (Figs. 2C,D, solid magenta, and 5). We anticipate that this discrepancy between the model and IC neurons may be mitigated if a future version of the model incorporates dynamic-range adaptation (e.g., Dean et al., 2005; 2008) because such adaptation tends to reduce firing rates when a neuron is highly active at all of the tested conditions, multiple distances in the anechoic condition in this case.

Our group (Kuwada et al., 2015) tested the rabbit's binaural behavioral discrimination of the distance of a sound source using a one-interval psychophysical procedure and a sound-field stimulation. The source sound was 250 ms bursts (50% duty cycle) of unmodulated noise. The study found that rabbit's binaural discrimination threshold distance ratios were ~1.6. These thresholds were somewhat lower than human binaural discrimination thresholds (~2.3) inferred from distance estimate data for unmodulated noise from Zahorik (2002a), following the logic described by Zahorik (2002b). These findings suggest that rabbits may have similar or slightly higher distance sensitivities than do humans.

The findings of binaural distance sensitivity for unmodulated sounds are in stark contrast to the present findings of monaural distance insensitivity for unmodulated 1 octave band sounds. Even without AM, these findings suggest that binaural mechanism(s) can extract sound distance information, whereas monaural mechanism(s) cannot for 1 octave band sounds. The rabbit's monaural behavioral distance sensitivity is not available. Further studies are needed to provide this information.

How well can humans discriminate AM depth? It was found that, on average, a change of 1.3 dB could be discriminated for a reference depth of 0 to -13 dB (Ozimek and Sek, 1988; Wakefield and Viemeister, 1990; Ewert and Dau, 2004). When combined with the relationship between AM depth and distance, the AM threshold predicts a threshold distance ratio of 1.7 for a d' of 1.09. Although this is considerably more sensitive than the observed threshold distance ratio of 5.3, it nevertheless indicates there is sufficient change in the AM depth to convey different distances. The lower than predicted distance sensitivity is not surprising, however, considering that an estimation/classification task is inherently more difficult (i.e., involves greater internal noise) than a discrimination task and that level was randomly roved in this study.

Although D/R ratio has been frequently considered to underlie sound-distance localization as stated above, the auditory system may not be able to separate direct sound from reverberant sound (Bronkhorst and Houtgast, 1999) when the two components are mixed together as is the case in general. Therefore, several candidate cues that covary with D/R ratio have been suggested as cues for sound distance localization. Monaural candidate cues are the early-to-late power ratio (Bronkhorst and Houtgast, 1999), spectral centroid, and spectral variance (Larsen et al., 2008). Depth of AM that covaries with D/R ratio and sound distance, as addressed in the present study, is a novel example of such a cue. Our findings that rabbit IC neurons and human listeners could not discriminate the distance of monaurally presented unmodulated 1 octave band noise suggest that the monaural spectral centroid and spectral variance may have minimal contributions to distance processing of such sounds. Our human finding is consistent with Kopčo and Shinn-Cunningham

(2011) who also found poor distance sensitivity for a binaural 5 kHz narrow-band noise. When the stimulus is a monaural wide-band unmodulated noise, however, spectral centroid and spectral variance may provide effective distance information as Larsen et al. (2008) observed good sensitivity to D/R to such a sound. Because binaural cues, such as interaural coherence, were also shown to covary with D/R (Larsen et al., 2008; Kuwada et al., 2012), it is possible that the auditory system may use a combination of binaural and monaural cues to encode distance, as suggested by Bronkhorst (2002). Further study is needed to understand the dependence of all D/R-related distance cues, such as AM depth, on the acoustics of the listening environment. Along this line, a direct comparison of the performance of binaural versus monaural discrimination of distance of sounds of various bandwidths, with and without AM, both behaviorally and neurally, should help advance the understanding of the mechanisms underlying distance perception.

References

- Ahveninen J, Kopčo N, Jaaskelainen IP (2014) Psychophysics and neuronal bases of sound localization in humans. *Hearing Res* 307:86–97. [CrossRef Medline](#)
- Allen JB, Berkley DA (1979) Image method for efficiently simulating small-room acoustics. *J Acoust Soc Am* 65:943–950. [CrossRef](#)
- Anderson PW, Zahorik P (2014) Auditory/visual distance estimation: accuracy and variability. *Front Psychol* 5:1097. [CrossRef Medline](#)
- Beranek LL (1986) *Acoustics*. New York: American Institute of Physics for the Acoustical Society of America.
- Bernstein LR, Trahiotis C (2010) Accounting quantitatively for sensitivity to envelope-based interaural temporal disparities at high frequencies. *J Acoust Soc Am* 128:1224–1234. [CrossRef Medline](#)
- Blahut RE (1987) *Principles and practice of information theory*. Reading, MA: Addison-Wesley.
- Bronkhorst AW (2002) Modeling auditory distance perception in rooms. *EAA Forum Acusticum Sevilla, Sevilla, Spain*.
- Bronkhorst AW, Houtgast T (1999) Auditory distance perception in rooms. *Nature* 397:517–520. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Cannon WB (1915) *Bodily changes in pain, hunger, fear and rage*. New York: Appleton.
- Colburn HS, Carney LH, Heinz MG (2003) Quantifying the information in auditory-nerve responses for level discrimination. *J Assoc Res Otolaryngol* 4:294–311. [CrossRef Medline](#)
- Davis KA, Hancock KE, Delgutte B (2010) Computational model of inferior colliculus neurons. In: *Computational models of the auditory system* (Meddis R, Lopez-Poveda EA, Fay RR, Popper AN, eds). New York: Springer.
- Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 8:1684–1689. [CrossRef Medline](#)
- Dean I, Robinson BL, Harper NS, McAlpine D (2008) Rapid neural adaptation to sound level statistics. *J Neurosci* 28:6430–6438. [CrossRef Medline](#)
- Dramas F, Katz BF, Jouffrais C (2008) Auditory-guided reaching movements in the peripersonal frontal space. *J Acoust Soc Am* 123:3723. [CrossRef](#)
- Ewert SD, Dau T (2004) External and internal limitations in amplitude-modulation processing. *J Acoust Soc Am* 116:478–490. [CrossRef Medline](#)
- Ghazanfar AA, Neuheff JG, Logothetis NK (2002) Auditory looming perception in rhesus monkeys. *Proc Natl Acad Sci U S A* 99:15755–15757. [CrossRef Medline](#)
- Graziano MS, Reiss LA, Gross CG (1999) A neuronal representation of the location of nearby sounds. *Nature* 397:428–430. [CrossRef Medline](#)
- Green DM, Swets JA (1974) *Signal detection theory and psychophysics*. Huntington, NY: Robert E Krieger Publishing.
- Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. *Physiol Rev* 90:983–1012. [CrossRef Medline](#)
- Guipponi O, Wardak C, Ibarrola D, Comte JC, Sappey-Mariniere D, Pinède S, Ben Hamed S (2013) Multimodal convergence within the intraparietal sulcus of the macaque monkey. *J Neurosci* 33:4128–4139. [CrossRef Medline](#)
- Hall DA, Moore DR (2003) Auditory neuroscience: the salience of looming sounds. *Curr Biol* 13:R91–R93. [CrossRef Medline](#)
- Hartmann WM (1983) Localization of sound in rooms. *J Acoust Soc Am* 74:1380–1391. [CrossRef Medline](#)
- Holmes NP, Spence C (2004) The body schema and the multisensory representation(s) of peripersonal space. *Cogn Process* 5:94–105. [CrossRef Medline](#)
- Houtgast T, Steeneken HJM (1985) A review of the Mtf concept in room acoustics and its use for estimating speech-intelligibility in auditoria. *J Acoust Soc Am* 77:1069–1077. [CrossRef](#)
- Hurvich LM, Jameson D (1957) An opponent-process theory of color vision. *Psychol Rev* 64:384–404. [CrossRef Medline](#)
- Jack JJB, Noble D, Tsien RW (1975) *Electric current flow in excitable cells*. New York: Oxford UP.
- Jones HG, Koka K, Tollin DJ (2013) The sound source distance dependence of the acoustical cues to location and their encoding by neurons in the inferior colliculus: implications for the duplex theory. In: *Basic aspects of hearing* (Moore BCJ, ed), pp 273–282. New York: Springer.
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577. [CrossRef Medline](#)
- Kim DO, Moiseff A, Turner JB, Gull J (2008) Acoustic cues underlying auditory distance in barn owls. *Acta Otolaryngol* 128:382–387. [CrossRef Medline](#)
- Kim DO, Bishop B, Kuwada S (2010) Acoustic cues for sound source distance and azimuth in rabbits, a racquetball and a rigid spherical model. *J Assoc Res Otolaryngol* 11:541–557. [CrossRef Medline](#)
- Kim DO, Zahorik P, Bishop B, Kuwada S (2013) Effect of reverberation on acoustic measures relevant for localization (where) and recognition (what) of sounds located at various azimuths and distances: a study of humans and rabbits, p 514. Association for Research in Otolaryngology Annual Midwinter Meeting, Baltimore.
- Kolarik A, Cirstea S, Pardhan S (2013) Discrimination of virtual auditory distance using level and direct-to-reverberant ratio cues. *J Acoust Soc Am* 134:3395–3398. [CrossRef Medline](#)
- Kopčo N, Huang S, Belliveau JW, Raji T, Tengshe C, Ahveninen J (2012) Neuronal representations of distance in human auditory cortex. *Proc Natl Acad Sci U S A* 109:11019–11024. [CrossRef Medline](#)
- Kopčo N, Shinn-Cunningham BG (2011) Effect of stimulus spectrum on distance perception for nearby sources. *J Acoust Soc Am* 130:1530–1541. [CrossRef Medline](#)
- Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 84:255–273. [Medline](#)
- Kuwada S, Bishop B, Kim DO (2012) Approaches to the study of neural coding of sound source location and sound envelope in real environments. *Front Neural Circuits* 6:42. [CrossRef Medline](#)
- Kuwada S, Bishop B, Kim DO (2014) Azimuth and envelope coding in the inferior colliculus of the unanesthetized rabbit: effect of reverberation and distance. *J Neurophysiol* 112:1340–1355. [CrossRef Medline](#)
- Kuwada S, Kim D, Koch K, Abrams K, Idrobo F, Zahorik P, Carney L (2015) Near-field discrimination of sound source distance in the rabbit. *J Assoc Res Otolaryngol*. Advance online publication. Retrieved Jan. 17, 2015. doi: 10.1007/s10162-014-0505-5. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Macmillan NA, Creelman CD (2005) *Detection theory: a user's guide*. Mahwah, NJ: Erlbaum.
- Maier JX, Chandrasekaran C, Ghazanfar AA (2008) Integration of bimodal looming signals through neuronal coherence in the temporal lobe. *Curr Biol* 18:963–968. [CrossRef Medline](#)
- McDermott JH, Simoncelli EP (2011) Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron* 71:926–940. [CrossRef Medline](#)
- Mershon DH, Bowers JN (1979) Absolute and relative cues for the auditory perception of egocentric distance. *Perception* 8:311–322. [CrossRef Medline](#)
- Mershon DH, Ballenger WL, Little AD, McMurtry PL, Buchanan JL (1989) Effects of room reflectance and background noise on perceived auditory distance. *Perception* 18:403–416. [CrossRef Medline](#)

- Middlebrooks JC, Green DM (1991) Sound localization by human listeners. *Annu Rev Psychol* 42:135–159. [CrossRef Medline](#)
- Moore BCJ (1997) An introduction to the psychology of hearing. San Diego: Academic.
- Naguib M, Wiley RH (2001) Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication. *Anim Behav* 62:825–837. [CrossRef](#)
- Nelson PC, Carney LH (2004) A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J Acoust Soc Am* 116:2173–2186. [CrossRef Medline](#)
- Nelson PC, Carney LH (2007) Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *J Neurophysiol* 97:522–539. [CrossRef Medline](#)
- Nielsen SH (1993) Auditory distance perception in different rooms. *J Audio Eng Soc* 41:755–770.
- Oliver DL (2005) Neuronal organization in the inferior colliculus. In: *The inferior colliculus* (Winer JA, Schreiner CE, eds), pp 69–114. New York: Springer.
- Ozimek E, Sek A (1988) AM difference limens for noise bands. *Acustica* 66:154–160.
- Schiller PH (1992) The on and off channels of the visual-system. *Trends Neurosci* 15:86–92. [CrossRef Medline](#)
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304. [CrossRef Medline](#)
- Shinn-Cunningham BG, Kopčo N, Martin TJ (2005) Localizing nearby sound sources in a classroom: binaural room impulse response. *J Acoust Soc Am* 117:3100–3115. [CrossRef Medline](#)
- Siebert WM (1970) Frequency discrimination in the auditory system: place or periodicity mechanisms? *Proc IEEE* 58:723–730. [CrossRef](#)
- Sinanovic S, Johnson DH (2000) Toward a theory of information processing. In: *Information theory, 2000*, p 493. *Proceedings of IEEE International Symposium*.
- Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114:3394–3411. [CrossRef Medline](#)
- Slee SJ, Young ED (2013) Linear processing of interaural level difference underlies spatial tuning in the nucleus of the brachium of the inferior colliculus. *J Neurosci* 33:3891–3904. [CrossRef Medline](#)
- Solomon RL, Corbit JD (1974) An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychol Rev* 81:119–145. [CrossRef Medline](#)
- Wakefield GH, Viemeister NF (1990) Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. *J Acoust Soc Am* 88:1367–1373. [CrossRef Medline](#)
- Winer JA, Saint Marie RL, Larue DT, Oliver DL (1996) GABAergic feedforward projections from the inferior colliculus to the medial geniculate body. *Proc Natl Acad Sci U S A* 93:8005–8010. [CrossRef Medline](#)
- Zahorik P (2002a) Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111:1832–1846. [CrossRef Medline](#)
- Zahorik P (2002b) Direct-to-reverberant energy ratio sensitivity. *J Acoust Soc Am* 112:2110–2117. [CrossRef Medline](#)
- Zahorik P (2009) Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J Acoust Soc Am* 126:776–791. [CrossRef Medline](#)
- Zahorik P, Brungart DS, Bronkhorst AW (2005) Auditory distance perception in humans: a summary of past and present research. *Acta Acoust* 91:409–420.
- Zahorik P, Kim D, Kuwada S, Anderson P, Brandewie E, Collecchia R (2012) Amplitude modulation detection by human listeners in reverberant sound fields: carrier bandwidth effects and binaural versus monaural comparison. *J Acoust Soc Am* 15:050002. [CrossRef Medline](#)
- Zilany MS, Bruce IC, Carney LH (2014) Updated parameters and expanded simulation options for a model of the auditory periphery. *J Acoust Soc Am* 135:283–286. [CrossRef Medline](#)