Sample discrimination of frequency by hearing-impaired and normal-hearing listeners

Joshua M. Alexander[a)]
Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706

Robert A. Lutfi
Department of Communicative Disorders, University of Wisconsin, Madison, Wisconsin 53706

(Received 15 May 2006; revised 12 September 2007; accepted 25 October 2007)

In a multiple observation, sample discrimination experiment normal-hearing (NH) and hearing-impaired (HI) listeners heard two multitone complexes each consisting of six simultaneous tones with nominal frequencies spaced evenly on an ERB_N logarithmic scale between 257 and 6930 Hz. On every trial, the frequency of each tone was sampled from a normal distribution centered near its nominal frequency. In one interval of a 2IFC task, all tones were sampled from distributions lower in mean frequency and in the other interval from distributions higher in mean frequency. Listeners had to identify the latter interval. Decision weights were obtained from multiple regression analysis of the between- interval frequency differences for each tone and listeners' responses. Frequency difference limens (an index of sensorineural resolution) and decision weights for each tone were used to predict the sensitivity of different decision-theoretic models. Results indicate that low-frequency tones were given much greater perceptual weight than high-frequency tones by both groups of listeners. This tendency increased as hearing loss increased and as sensorineural resolution decreased, resulting in significantly less efficient weighting strategies for the HI listeners. Overall, results indicate that HI listeners integrated frequency information less optimally than NH listeners, even after accounting for differences in sensorineural resolution. © *2008 Acoustical Society of America.* [DOI: 10.1121/1.2816415]

PACS number(s): 43.66.Ba, 43.66.Fe, 43.66.Sr [JHG] Pages: 241–253

I. INTRODUCTION

The objective of this study is to understand the factors that influence a hearing-impaired (HI) listener's ability to discriminate between sounds that require an integration of information about small frequency differences. Two factors are commonly discussed in the literature with regard to multiple observation, sample discrimination tasks. The first is a decision strategy that assigns relative importance or perceptual weight to the various channels of information. The second is internal noise, which is often used as a collective term for everything that cannot be explained by an observer's decision strategy. Sources of internal noise include sensorineural or peripheral processes associated with the precision of transduction and neural coding as well as central processes associated with observer variability or inconsistency in the execution of a decision strategy. As explained below, when the size of the differences to be discriminated is the same magnitude as the difference limens (DLs) and near the sensorineural noise floor, a delineation of the sources of internal noise is necessary for a proper interpretation of sensitivity relative to an ideal model. The primary aim of this paper is to offer a theoretical framework for investigating the influence of channel-specific sensorineural noise on decision weights and internal noise.

Most of the studies examining the effects of sensorineural hearing loss (SNHL) on decision weights and internal noise have used intensity discrimination tasks (Doherty and Lutfi, 1996, 1999; Lentz and Leek, 2002, 2003). The results from these experiments generally indicate that weighting strategies are variable across normal-hearing (NH) listeners and even more so across HI listeners. They also indicate that HI listeners adopt slightly different weighting strategies than NH listeners do. For example, Doherty and Lutfi (1996) found that HI listeners were apt to put the most weight on a 4000 Hz tone that was in the sloping region of hearing loss in a discrimination task where the signal was the sum of intensity level increments on six fixed-frequency tones. In contrast, the NH listeners as a group distributed their attention across the spectrum. A similar pattern of weights emerged for HI listeners in Doherty and Lutfi (1999) when the signal was a reliable level increment on only the 4000 Hz tone so that HI listeners were actually at an advantage relative to NH listeners. No differences were significant when the signal was at 250 and 1000 Hz, where hearing thresholds were more similar for the two groups.

In a spectral shape discrimination or "profile analysis" task in which listeners detected increments in the level of a 920 Hz tone relative to pairs of flanking tones above and below the signal frequency, Lentz and Leek (2002) found limited evidence that HI listeners used less optimal decision strategies than NH listeners. HI listeners were less likely to weight the flanking components when they should have been attended to, but more likely to weight the flanking compo-

[a)]Author to whom correspondence should be addressed. Present address: Boys Town National Research Hospital, Omaha, Nebraska, 63131. Electronic mail: alexanderj@boystown.org

nents when they should have been ignored. In a tone-detection task, Alexander and Lutfi (2004) also found that HI listeners as a group tended to put less weight on a 2000 Hz signal tone and more weight on flanking tones than NH listeners did. Differences between NH and HI listeners were also noted by Lentz and Leek (2003) when the spectral shape of the signal was a level increment in the three lowest or the three highest frequency components and a decrement in the other half. HI listeners as a group tended to put the most weight on the lowest and the highest frequencies of the complex while the NH listeners tended to put the most weight on the two centermost frequencies.

The above studies indicate that SNHL alters the way HI listeners weight information compared to NH listeners, although the differences are task dependent and do not always influence overall performance. They also suggest that internal noise is greater for HI listeners compared to NH listeners. In all of their conditions, Doherty and Lutfi (1996, 1999) found that an index of internal noise was significantly greater for HI listeners than for NH listeners. Lentz and Leek (2003) also indicated that HI listeners might have greater amounts of internal noise as demonstrated by a greater amount of variability in their weighting strategies across conditions. For these experiments, it is assumed that peripheral or sensorineural processes were not responsible for the observed differences between groups because intensity resolution is minimally affected by HI (Florentine *et al.*, 1993; Buus *et al.*, 1995). This suggests that central processes were responsible for the observed differences, but we cannot be sure because current analytic techniques do not allow for a separation of peripheral and central processes.

Challenges to traditional analyses originating from signal-detection theory arise when studying sensitivity for frequency discrimination in listeners with SNHL because differences within and between listeners can be more than a magnitude for individual tones (e.g., Freyman and Nelson, 1991). Because frequency distortion is a condition of many HI listeners' everyday lives and because current hearing aid technology does little to correct for it, it is of interest to examine how differences in the spectral profile of sensorineural noise influence the way listeners weight information as a function of frequency. In a decision-theoretic framework, best possible discrimination performance will occur when frequency information is weighted inversely proportional to the frequency-specific noise. Frequency regions where sensorineural noise is lower should be weighted more than regions where it is higher. Thus, for each sensorineural noise profile there is a corresponding optimal weighting strategy. HI listeners might listen differently than NH-listeners, but the differences might be systematic and follow each listener's individual ideal. In this case, comparing both NH and HI listeners to the stimulus ideal, where weights are proportional to the information sent instead of the information received, erroneously implies that HI listeners weight information inefficiently. In order to obtain a fair comparison of HI and NH listeners in conditions where sensorineural noise is not procedurally regulated, in what follows we will describe the concept of a *peripherally limited ideal observer* (PLIO), an analytic solution to separate the effects of sensorineural noise on decision weights and internal noise.

II. METHODS

A. Listeners

Thirteen NH listeners 20–45 years old (median age of 23), including the first author (NH 01), and 13 HI listeners 55–84 years old (median age of 77) participated in the study.[1] On standardized audiometric tests (ANSI S3.6-1996) NH listeners had pure-tone thresholds of 10 dB HL or better, except a few who had one threshold equal to 15 dB HL. The right ears of the NH listeners served as the test ear in the experiments. HI listeners had mild to moderate SNHL in the test ear and minimal conductive hearing loss (i.e., the differences between air and bone conduction thresholds were 10 dB or less and tympanometric tests were clinically normal). When both ears fit the inclusion criteria for the HI group, the right ear was used as the test ear.

B. Stimuli

Stimuli were generated with a 44.1 kHz sampling rate using the MATLAB® programming language and a 16 bit sound card. A programmable attenuator was used to control the overall level. Stimuli were presented monaurally through a Beyerdynamic DT 990 earphone. Unlike frequency discrimination, intensity discrimination follows the "near-miss to Weber's Law," which makes it convenient to use the dB scale to perturb the information for each tone in roughly equal units of the jnd (just noticeable difference). Difference limens for frequency (DLFs) do not vary systematically as a function of frequency for reasons beyond the current discussion (e.g., Moore, 1997). For this reason, we manipulated and measured frequency using a scale based on the equivalent rectangular bandwidth for *normal-hearing* listeners, $\mathrm{ERB_N}$ (Glasberg and Moore, 1990; Moore, 1997). The $\mathrm{ERB_N}$ is a log-based psychophysical scale for frequency where the $\mathrm{ERB_N}$ No. $=21.4\log_{10}(4.37F+1)$ for frequency F in Hz. It should be noted that because no measure of frequency selectivity was obtained, it is unknown how the frequency manipulations mapped onto listeners' filter space, especially for the HI listeners. Furthermore, while many log-based scales could have been used, the important point is that the smallest and largest DLFs between 250 and 8000 Hz for NH listeners are usually within an order of magnitude in $\mathrm{ERB_N}$ (Sek and Moore, 1995). Tones were sampled from one of six different frequency regions whose center frequencies were logarithmically separated by five $\mathrm{ERB_N}$: 257, 603, 1197, 2212, 3951, and 6930 Hz. Tones were 120 ms in duration with 12 ms cosine-squared ramps and random starting phase ($0-2\pi$ radians). In the discrimination experiments, each tone was presented at 82 dB sound pressure level (SPL) for all listeners.

C. Procedure

Listeners ran the experiments individually while seated in a double-walled, sound-attenuated chamber. Trials consisted of a two-interval, forced-choice task (2IFC) with a 1 s inter-stimulus interval, an initial warning, and visual feedback. The intervals were specified on a computer monitor and listeners indicated which of the two intervals contained the signal by pressing a button. There were no time limits for making a response.

1. Tests of cochlear function: Adaptive staircase method

The purpose of these tests was to provide information about the degree of hearing loss and about the processing limits imposed by the auditory periphery. Detection thresholds in quiet (QTs) and DLFs were measured for each of the six center frequencies in isolation. Signal levels and frequency differences were adapted using a two-down, one-up decision rule which converges on the 70.7% point on the psychometric function (Levitt, 1971), or equivalently, the point where $d'_{(2\mathrm{IFC})}=0.77$ (Macmillan and Creelman, 2005). Each trial block consisted of 12 reversals in the adaptive track. The average values of the final eight reversals determined the QT or DLF for the block.

The initial step size of the adaptive tracks for QT was 4 dB and was reduced to 2 dB after the third reversal. Maximum presentation-level was limited to 85 dB SPL. QT for each tone was based on the mean across three trial blocks. QT blocks were completed before DLF blocks because DLFs were not tested for tones where QT$>$77 dB SPL (i.e., SL $<$5 dB).

For the DLFs, the first block of trials was used to give the listeners practice and to calibrate the starting frequency differences of the following adaptive tracks. Listeners had to identify the interval with the highest-frequency tone. For most listeners it was expected that pitch would be the most salient cue, although some HI listeners indicated that they sometimes used loudness for a cue especially when the frequencies were in a sloping region of hearing loss. The initial frequency difference, D, in the practice trial block was 0.2 $\mathrm{ERB_N}$ for the NH listeners and 0.6 $\mathrm{ERB_N}$ for the HI listeners. The lower-frequency tone in each interval was $D/2$ less than the nominal center frequency and the higher-frequency tone was $D/2$ greater so that regardless of the size of the frequency difference, the unbiased criterion was always the same. The starting frequency difference for subsequent blocks was 0.04 $\mathrm{ERB_N}$ above the running average of all previous blocks for that frequency. The initial step size of the adaptive tracks was 0.02 $\mathrm{ERB_N}$ and was reduced to 0.01 $\mathrm{ERB_N}$ after the third reversal. Before running the sample discrimination experiment, listeners completed two additional blocks for each frequency after the initial practice blocks. After running the sample discrimination experiment, two more blocks for DLF were run. The DLF for each frequency was based on the mean across four blocks.[2]

2. Sample discrimination of frequency

In a sample discrimination experiment the difference to be discriminated is nonadaptively varied from trial to trial by sampling the stimuli in each interval (indexed j) from one of two overlapping distributions with identical standard deviations, $\sigma_{S(j)}$, but different means (e.g., Lutfi, 1989, 1990, 1992; Lutfi *et al.*, 1996; Jesteadt *et al.*, 2003; Sorkin *et al.*, 1987; Neff and Odgaard, 2004). Stimuli in this sample discrimination task consisted of six simultaneous tones whose frequencies (X_i) were independently sampled from one of two distributions. As shown in Fig. 1, in the "noise" interval of a 2IFC task the frequencies of the six-tone complex were sampled from distributions lower in mean frequency and in the "signal" interval they were sampled from distributions higher in mean frequency. Listeners were told to consider all the tones in each interval and then to select the interval where the tones were "on average higher in pitch."

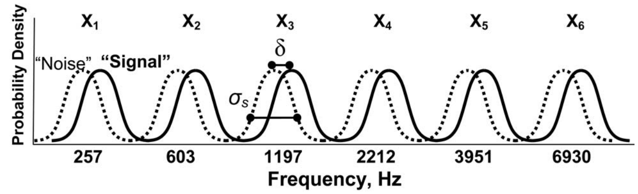

FIG. 1. Shown is a schematic of the two-interval, forced-choice sample discrimination experiment. The tones selected from the distributions higher in mean frequency are from the "signal" interval and those selected from the distributions lower in mean frequency are from the "noise" interval. The distributions for each interval were centered on six fixed frequencies that were separated by a logarithmic distance of five $\mathrm{ERB_N}$. The expected frequency difference between distribution means, δ, was constant for each of the six pairs of tones, but varied across ten different trial blocks. The standard deviation of the sampling distributions, σ_S, was a constant factor of δ.

The difference between the means of the signal and noise distributions, or the average difference to be discriminated is denoted δ_i, where i is the tone number. Across ten different 55 trial blocks, the expected frequency difference for each tone was fixed at $0.02, 0.04, \ldots,$ or 0.2 $\mathrm{ERB_N}$ for the NH listeners and at $0.06, 0.12, \ldots,$ or 0.6 $\mathrm{ERB_N}$ for the HI listeners. One block of trials was run for each condition (expected frequency difference). $\sigma_{S(ij)}$ was a constant factor of the expected frequency difference. For all listeners $\sigma_{S(ij)}=1.155\delta_i$ across all tones and conditions. For ease of discussion, the ten conditions are nominally identified by the value of δ_i. Larger frequency manipulations were used for the HI listeners because it was anticipated that greater amounts of sensorineural noise (greater DLFs) would prevent them from performing the experiment above chance. As with the DLF procedure, the expected mean frequencies of the signal and noise distributions for each condition were $\delta_i/2$ above and below the nominal center frequencies, respectively, so that regardless of the condition, the unbiased criterion was constant.

To minimize the effects of practice, the first five trials of each block were discarded and the ten conditions were run in random order during one session. Furthermore, before data collection began at least one block of practice trials was run with $\delta_i=0.2$ for the NH listeners and $\delta_i=0.6$ for the HI listeners, with $\sigma_{s(ij)}=0.5\delta_i$. Listeners began the experiment if the attained d' was at least 1.5 on the first or second practice block or at least 1.0 after the third practice block.

D. Theoretical framework

We model the listener's task when discriminating between two multitone patterns as a univariate decision variable, DV, equal to the weighted sum of the information at each frequency. For our experiments, the information sent is the difference in frequency between the pairs of tones across trial intervals. However, the actual information received is corrupted by frequency-specific sensorineural noise, which is modeled as random variables, ε_{ij}, with variance $\sigma^2_{P(ij)}$ and mean zero that add to each tone (X_{ij}). Variability in the use of received information is central noise and is modeled as a single random variable, e_C, with variance σ^2_C.

$$DV = e_c + \sum_{i=1}^{m} w_i[(X_{i2} + \varepsilon_{i2}) - (X_{i1} + \varepsilon_{i1})], \qquad (1)$$

where m is the number of tones in each interval (indexed i), and w_i are the weights. On every trial, the listeners are assumed to respond "interval 2" when the DV exceeds some fixed, internal decision criterion (ideally, 0 in a 2IFC task).

Internal noise limits the ability to make trial-by-trial predictions; therefore, model predictions were made for sensitivity, d'_{DV}. An analytic prediction for the Z transform of the hit and false-alarm (FA) rates is given by the expected values of the DV when the signal is in interval 2 and interval 1, respectively.

$$Z(hit) = -Z(FA) = \sum_{i=1}^{m} w_i\delta_i \Bigg/ \left[\sigma_c^2 + \sum_{i=1}^{m} w_i^2(2\sigma^2_{S(i)} + 2\sigma^2_{P(i)})\right]^{1/2}. \qquad (2)$$

Because the expected value of each sensorineural noise variable, ε_{i2} and ε_{i1}, is zero their expected difference is also zero. In addition, because the expected value of X_{i2} is $(\delta_i/2)$ and the expected value of X_{i1} is $-(\delta_i/2)$, the expected value of their difference is δ_i when the signal is interval 2 and $-\delta_i$ when the signal is interval 1. Therefore, $\Delta = 2\Sigma_{i=1}^{m} w_i\delta_i$, where Δ is the overall difference between the signal and noise distributions of the decision variable. For the unbiased observer, the hit and false-alarm rates are equal and $d'_{DV} = [Z(hit) - Z(FA)]_{2IFC}/\sqrt{2}$, (cf. Eqs. (7.2) and (7.11) in Macmillan and Creelman, 2005).[3] Therefore, the analytic prediction for sensitivity is

$$d'_{DV} = \sqrt{2} \times \sum_{i=1}^{m} w_i\delta_i \Bigg/ \left[\sigma_C^2 + \sum_{i=1}^{m} w_i^2(2\sigma^2_{S(i)} + 2\sigma^2_{P(i)})\right]^{1/2}. \qquad (3)$$

Within our theoretical framework, increased DLFs (reduced resolution for frequency differences) associated with SNHL are simply modeled as a higher sensorineural noise floor, σ_P^2. Empirically, σ_P^2 is estimated by the magnitude of hypothetical stimulus noise, σ_S^2, needed to describe the discriminability ($d'_{DL(i)}$) between two simple sounds free of external variability or experimental uncertainty: $d'_{DL(i)} = D_{DL(i)}/\sigma_P$ where $D_{DL(i)}$ is the difference limen or the smallest discriminable difference for the ith tone. For each frequency, the average DLF across four blocks, $D_{DL(i)}$, was converted to an estimate of sensorineural noise, $\sigma^2_{P(i)}$, using the expected sensitivity for the two-down/one-up, 2IFC task, $d'_{DL(i)} = 0.77$.

TABLE I. Differences between the various decision models arise from the way decision weights are derived and the source of internal noise.

Model	Weights	Internal noise
d'_{ideal}	$\delta_i/\sigma_{S(i)}$	None
d'_{PLIO}	$\delta_i/\sqrt{\sigma^2_{S(i)}+\sigma^2_{P(i)}}$	σ_P^2
d'_{wgt}	Regression	None
d'_{pred}	Regression	σ_P^2

The observed sensitivity of the listener, d'_{obs}, was computed from the difference between the Z-transformed hit and false-alarm rates for the 2IFC task (see above). For each listener and for each sample discrimination condition, four estimates of sensitivity, $d'_{\text{ideal}}, d'_{\text{PLIO}}, d'_{\text{wgt}}$, and d'_{pred} were analytically derived using Eq. (3). As shown in Table I, the four models differed in terms of the weights and sources of internal noise included in the model.

d'_{ideal} is the maximum expected sensitivity of a hypothetical listener free from internal noise, the stimulus ideal. Ideal sensitivity for a noiseless observer is achieved when weights are proportional to the information *sent* in each channel, $d'_i = \delta_i/\sigma_{S(i)}$. Because d'_i was kept constant (0.866) for all listeners and all sample discrimination conditions, ideal weights, w_{ideal}, for each tone were each equal to each other (1/6) and d'_{ideal} was equal to 2.12 (i.e., $d'_i\sqrt{m}$).

d'_{PLIO} is the maximum expected sensitivity given an individual listener's unique profile of sensorineural noise, what we term a "*peripherally limited ideal observer*" (PLIO). Sensitivity for a PLIO is achieved when weights are proportional to the information *received* in each channel. Information is less reliable in spectral regions where sensorineural noise is relatively large and should be weighted less in a listener's decision. Weights for the PLIO, w_{PLIO}, were proportional to the estimated $\delta_i/\sqrt{\sigma^2_{S(i)}+\sigma^2_{P(i)}}$ for each tone, where $\sigma^2_{P(i)}$ was estimated from the listener's DLFs. d'_{PLIO} each condition was derived by including $\sigma^2_{P(i)}$ in the denominator of Eq. (3) and by substituting w_{PLIO} for w_i.

d'_{pred} represents the predicted sensitivity after accounting for the listener's weighting strategy and peripheral limitations. Because d'_{pred} into account the stimulus statistics and all other listener variables, any remaining difference between d'_{pred} and d'_{obs} attributed to central noise *post hoc*. d'_{pred} was computed in the same way as d'_{PLIO} but using the listeners' regression weights, w_{obs}. This differs from traditional computations d'_{wgt} in that $\sigma^2_{P(i)}$ is included in the pooled variance of the expected decision variable. Observed weights (w_{obs}) across all ten conditions were computed by logistic regression. Listeners' interval 1 responses were coded as "0" and interval 2 responses as "1" and the probability of responding "interval 2" was regressed against the difference in ERB_N between each pair of tones in interval 2 and interval 1. Only statistically significant weights were used for the model estimates with nonsignificant weights set to zero in the equation for sensitivity [Eq. (3)]. So that the weights could be

TABLE II. Information for each of the six frequency regions tested in this study for the NH listeners (arranged from left to right in descending order of observed sensitivity). Section 1: mean quiet thresholds in dB SPL. Section 2: estimates of σ_P expressed as a proportion of $\mathrm{ERB_N}$. Section 3: The weights for the PLIO, w_{PLIO}. Section 4: Observed weights, w_{obs}, that are statistically different from 0 are in bold and those that are not are un-bolded and in parentheses.

		NH 08	NH 10	NH 07	NH 06	NH 12	NH 02	NH 11	NH 14	NH 04	NH 09	NH 01	NH 13	NH 03	Mean	(SE)
QT (dB SPL)	**257**	11.7	12.8	7.5	9.2	9.5	13.3	12.5	13.5	12.3	15.8	7.5	22.4	9.4	**12.1**	(1.15)
	603	3.0	13.0	5.8	7.8	2.3	4.4	11.6	6.0	7.0	9.4	5.0	12.3	7.3	**7.3**	(0.99)
	1197	7.1	18.9	2.1	13.7	3.7	0.8	7.7	0.2	−0.1	13.5	2.8	2.8	7.3	**6.2**	(1.72)
	2212	7.3	14.0	−4.2	12.8	10.5	1.6	4.5	−0.2	1.5	8.6	−1.6	1.5	14.8	**5.5**	(1.82)
	3951	−3.1	7.6	6.0	−1.0	−3.8	−5.8	−2.3	−0.8	1.1	1.1	8.8	−3.2	−1.7	**0.2**	(1.32)
	6930	7.1	15.1	9.1	6.1	2.2	−2.9	8.7	10.2	8.7	10.1	18.3	11.3	17.0	**9.3**	(1.65)
$\boldsymbol{\sigma_p}$ ($\mathbf{ERB_N}$)	**257**	0.07	0.12	0.14	0.03	0.09	0.06	0.04	0.04	0.05	0.12	0.03	0.05	0.08	**0.07**	(0.011)
	603	0.06	0.09	0.06	0.04	0.04	0.05	0.03	0.03	0.03	0.08	0.05	0.07	0.09	**0.05**	(0.006)
	1197	0.07	0.08	0.05	0.02	0.03	0.07	0.04	0.04	0.04	0.04	0.03	0.08	0.10	**0.05**	(0.006)
	2212	0.03	0.11	0.08	0.04	0.04	0.07	0.04	0.06	0.03	0.06	0.05	0.06	0.09	**0.06**	(0.007)
	3951	0.06	0.20	0.13	0.03	0.04	0.04	0.06	0.13	0.07	0.03	0.07	0.05	0.06	**0.08**	(0.014)
	6930	0.18	0.25	0.26	0.46	0.15	0.11	0.31	0.44	0.48	0.33	0.10	0.32	0.17	**0.27**	(0.038)
$\boldsymbol{w_{\mathrm{PLIO}}}$	**257**	0.17	0.17	0.16	0.19	0.13	0.16	0.18	0.21	0.19	0.16	0.21	0.18	0.17	**0.18**	(0.006)
	603	0.17	0.19	0.18	0.19	0.19	0.19	0.18	0.23	0.21	0.18	0.17	0.18	0.17	**0.19**	(0.005)
	1197	0.17	0.19	0.19	0.19	0.21	0.16	0.18	0.21	0.20	0.19	0.20	0.18	0.17	**0.19**	(0.005)
	2212	0.18	0.18	0.18	0.19	0.19	0.14	0.18	0.19	0.20	0.19	0.17	0.18	0.17	**0.18**	(0.004)
	3951	0.17	0.14	0.16	0.19	0.19	0.25	0.17	0.12	0.16	0.19	0.14	0.18	0.18	**0.17**	(0.010)
	6930	0.14	0.12	0.13	0.06	0.09	0.10	0.11	0.04	0.03	0.09	0.10	0.11	0.14	**0.10**	(0.010)
$\boldsymbol{w_{\mathrm{obs}}}$	**257**	**0.17**	**0.23**	**0.06**	**0.22**	**0.43**	**0.55**	**0.21**	**0.21**	**0.16**	**0.22**	**0.16**	**0.12**	**0.44**	0.24	(0.040)
	603	**0.33**	**0.19**	**0.30**	**0.31**	**0.16**	**0.19**	**0.31**	**0.45**	**0.36**	**0.18**	**0.17**	**0.11**	**0.20**	0.25	(0.029)
	1197	**0.21**	**0.32**	**0.31**	**0.33**	**0.13**	**0.14**	**0.34**	**0.33**	**0.26**	**0.49**	**0.19**	**0.67**	**0.37**	0.31	(0.042)
	2212	**0.15**	**0.08**	**0.18**	**0.08**	**0.18**	(0.05)	(0.04)	(0.04)	**0.09**	**0.11**	(0.01)	(0.01)	(0.09)	0.09	(0.017)
	3951	**0.14**	**0.12**	**0.15**	**0.07**	**0.10**	**0.12**	**0.14**	(−0.02)	**0.13**	(−0.01)	**0.15**	**0.10**	(0.08)	0.10	(0.016)
	6930	(0.05)	**0.07**	(−0.01)	(−0.01)	(0.00)	(0.00)	(0.02)	(0.01)	(−0.05)	(0.05)	**0.34**	(0.00)	(0.00)	0.04	(0.028)

interpreted as the proportion of attention devoted to each frequency region, they were normalized so that the sum of the absolute values of the significant weights was one (Berg, 1990; Lutfi, 1992; Doherty and Lutfi, 1996, 1999).

III. RESULTS

A. Sensorineural resolution and weights for the PLIO

A quantification of sensorineural resolution is necessary to determine the contribution of sensorineural noise to overall internal noise and its spectral profile is important for determining the relative weighting function for a PLIO (w_{PLIO}). The means of the DLF-transformed estimates of σ_P and standard errors of the mean (SE's) are provided in Tables II and III for the NH and HI listeners, respectively, and are plotted in Fig. 2 with 95% confidence intervals (CIs).

Sensorineural resolution estimates for HI 12 were excluded from the means and related statistical analyses because the estimates were unusually large and in some cases contributed to over 1/3 of the error variance. Lower values of σ_P indicate higher (better) sensorineural resolution. NH listeners clearly had better sensorineural resolution than the HI listeners did. On average, σ_P was about 3–10 times greater in HI listeners than in NH listeners. Across individual HI listeners, there was substantial variability in sensorineural resolution. The profile of σ_P ranged from normal to more than a magnitude greater than normal for a few listeners.

The absolute differences in σ_P had little influence on w_{PLIO} because the σ_S/σ_P ratios were about the same magnitude for the two groups of listeners. Plotted in the left panel of Fig. 3 are the mean PLIO weights for the NH and HI listeners (different symbols) with 95% CIs. Between group comparisons revealed that w_{PLIO} were significantly greater for the HI listeners at 257 and 603 Hz [$t(24)=2.4$, $p=0.03$; $t(24)=2.6$, $p=0.02$] but significantly greater for the NH listeners at 3951 Hz [$t(23)=3.9, p<0.001$]. These results demonstrate the advantage of incorporating sensorineural resolution into the model of the ideal listener. For best discrimination performance, HI listeners should listen differently than NH listeners. In particular, as a group, the observed weights should be slightly greater for the lowest frequency tones but less for the 3951 Hz tone. The important point is that individual listeners have unique ideal models based on information received and comparing them to the same ideal based on information sent instead of information received can penalize them for otherwise listening optimally.

B. Factors influencing observed decision weights

1. Hearing loss

The mean observed weights with 95% CIs for the NH and HI listeners are plotted in the right panel of Fig. 3. Observed weights for both groups of listeners clearly deviated from the weights for the PLIO. NH listeners relied primarily on the low-frequency tones (257–1197 Hz) for discrimina-

TABLE III. Hearing-impaired listeners (see Table II for description). Estimates σ_P for HI 12 were excluded from the means and SE's.

		HI 09	HI 01	HI 17	HI 04	HI 13	HI 11	HI 15	HI 10	HI 02	HI 18	HI 05	HI 12	HI 03	Mean	(SE)
QT	**257**	23.6	33.3	28.7	26.4	26.4	12.1	28.6	19.1	28.3	57.1	44.6	32.0	36.3	**30.5**	(3.24)
(dB SPL)	**603**	24.5	12.5	27.9	27.8	26.8	7.5	18.9	21.5	49.1	44.5	54.8	31.1	33.2	**29.2**	(3.96)
	1197	20.9	12.8	40.8	41.1	26.8	7.0	14.4	48.4	45.7	26.6	51.5	31.3	41.5	**31.4**	(4.22)
	2212	55.8	14.7	53.7	54.5	36.7	60.6	40.7	57.0	58.5	18.0	68.9	42.0	45.1	**46.6**	(4.66)
	3951	62.5	36.8	58.1	61.1	[a]	59.1	51.8	62.8	59.6	17.9	63.7	44.6	50.0	**52.3**	(4.11)
	6930	62.2	53.1	72.3	74.9	[a]	50.8	72.6	71.4	[a]	36.4	[a]	56.1	74.2	**62.4**	(4.32)
$\boldsymbol{\sigma_p}$	**257**	0.07	0.04	0.09	0.15	0.17	0.08	0.21	0.30	0.14	0.46	0.26	0.89	0.35	**0.19**	(0.039)
($\mathbf{ERB_N}$)	**603**	0.06	0.03	0.06	0.06	0.12	0.05	0.16	0.23	0.07	0.53	0.22	0.73	0.36	**0.16**	(0.046)
	1197	0.08	0.05	0.06	0.09	0.09	0.09	0.15	0.32	0.08	0.18	0.67	0.98	0.70	**0.21**	(0.070)
	2212	0.29	0.06	0.08	0.23	0.09	0.40	0.10	0.23	0.37	0.29	0.88	0.71	0.63	**0.30**	(0.074)
	3951	1.35	0.12	0.36	0.73	[a]	0.23	0.24	1.09	0.68	1.01	1.39	2.32	1.32	**0.77**	(0.153)
	6930	1.10	1.02	0.49	1.19	[a]	0.42	0.90	0.35	[a]	1.83	[a]	1.24	1.41	**0.97**	(0.173)
w_{PLIO}	**257**	0.20	0.20	0.18	0.20	0.18	0.19	0.18	0.18	0.21	0.19	0.28	0.18	0.26	**0.20**	(0.010)
	603	0.20	0.20	0.18	0.20	0.25	0.19	0.19	0.21	0.33	0.18	0.30	0.22	0.25	**0.22**	(0.014)
	1197	0.20	0.20	0.18	0.20	0.29	0.19	0.19	0.17	0.32	0.23	0.18	0.17	0.15	**0.21**	(0.014)
	2212	0.19	0.19	0.18	0.19	0.29	0.14	0.20	0.21	0.09	0.22	014	0.22	0.17	**0.19**	(0.014)
	3951	0.09	0.17	0.15	0.12	[a]	0.17	0.17	0.06	0.05	0.12	0.10	0.07	0.09	**0.11**	(0.013)
	6930	0.11	0.04	0.13	0.09	[a]	0.13	0.07	0.16	[a]	0.07	[a]	0.13	0.08	**0.10**	(0.013)
w_{obs}	**257**	**0.49**	**0.47**	**0.38**	**0.68**	**0.44**	**0.61**	**0.49**	**0.74**	**0.23**	**0.86**	**1.00**	**0.45**	**0.64**	**0.58**	(0.060)
	603	**0.26**	**0.30**	**0.37**	**0.22**	**0.27**	**0.10**	**0.24**	**0.26**	**0.69**	(−0.09)	(0.07)	**0.36**	**0.36**	**0.2**	(0.053)
	1197	**0.19**	**0.23**	**0.15**	**0.10**	**0.29**	**0.29**	**0.26**	(0.04)	**0.09**	**0.14**	(0.06)	**0.19**	**0.03**	**0.16**	(0.027)
	2212	(0.03)	(0.05)	(0.07)	(0.01)	(0.08)	(0.08)	(0.06)	(0.04)	(0.03)	(−0.04)	(−0.03)	(0.06)	(−0.14)	**0.02**	(0.018)
	3951	**−0.07**	(−0.01)	**0.10**	−(0.014)	[a]	(−0.07)	(0.02)	(0.04)	(−0.03)	(0.05)	(−0.07)	(−0.13)	(0.06)	**−0.01**	(0.020)
	6930	(0.30)	(−0.05)	(−0.04)	(0.06)	[a]	(0.00)	(−0.06)	(0.01)	[a]	(0.11)		(−0.12)	(−0.10)	**−0.02**	(0.024)

[a]Denotes tones where QT >85 dB SPL and where the DLF (σ_P) was subsequently not tested.

tion to the exclusion of the high-frequency tones (2212–6930 Hz). As indicated by the bolded entries in Table II, which indicate statistical significance, only two NH listeners (NH 01 and NH 10) put significant weight on the highest-frequency tone. One of these listeners was the first author, a highly trained listener on this task, who put the greatest weight on the 6930 Hz tone. The HI listeners also put very little weight on the high-frequency tones. As shown in Table III, only two HI listeners put significant weight on a high-frequency tone (3951 Hz for HI 09 and HI 17). Unlike the NH listeners who put about equal emphasis on the low-frequency tones, the HI listeners put substantially greater weight on the lowest-frequency tone at 257 Hz with progressively less weight on the 603 and 1197 Hz tones.

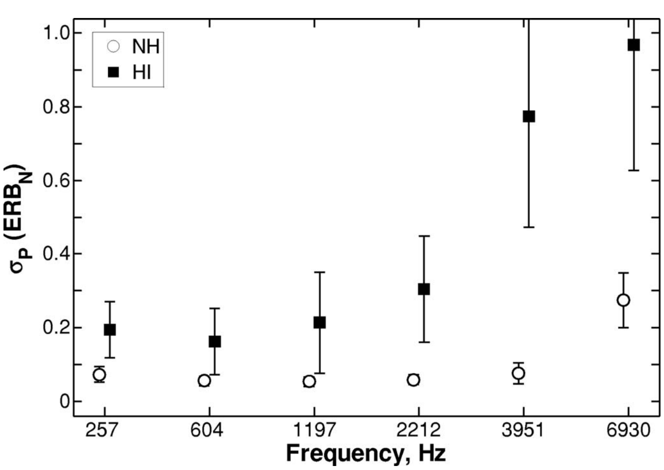

FIG. 2. Displayed are mean estimates of sensorineural resolution for each frequency region for the NH and HI listeners (open circles and filled squares, respectively) as obtained on an ERB$_N$ scale. Error bars represent the 95% confidence intervals. (Estimates σ_P for HI 12 are excluded from the means and CI's.)

The above observations were qualified by within-subjects analyses of variance (ANOVAs) conducted separately for each group of listeners. Weights for the NH listeners were significantly greater than zero for every tone except the 6930 Hz tone. None of the lower-frequency weights was significantly different from each other and none of the higher-frequency weights was significantly different from each other. Each paired comparison across the lower- and higher-frequency groups was significant. For the HI listeners only the lower-frequency weights were significantly different from zero, especially the 257 Hz weight, which was significantly greater than all the other weights ($p<0.001$). The 603 Hz weight was significantly greater than the 2212 Hz weights and above and the 1197 Hz weight was significantly greater than the 3951 and 6930 Hz weights.

To see if there was a systematic relationship between the observed weights and the amount of hearing loss, the weights for each frequency were correlated with the QTs across both groups of listeners. The 257 Hz weight had a significant positive relationship with QT $[R(24)=0.68, p<0.001]$ and the 1197 and 3951 Hz weights had significant negative relationships with QT $[R(24)=0.62$, $p<0.001$ and

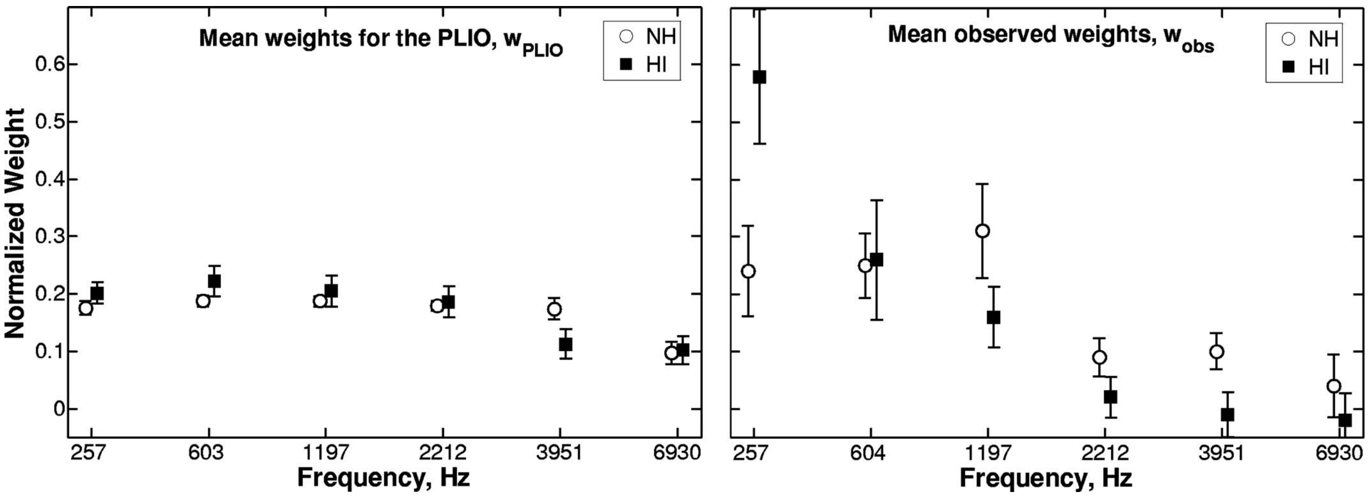

FIG. 3. Mean normalized weights with 95% confidence intervals are shown for each frequency region for the NH and HI listeners (open circles and filled squares, respectively). Weights for the peripherally limited ideal observer are plotted in the left panel and the observed weights in the right panel.

$R(23)=0.67$, $p<0.001$, respectively]. The 2212 Hz weight had a marginally significant negative relationship with QT [$R(24)=0.39, p=0.05$]. These results indicate that the low-frequency emphasis in the profile of observed weights increased as hearing loss increased.

2. Sensorineural resolution

The correlations between the observed weights and the weights for the PLIO for the NH listeners [$R(76)=0.44, p<0.001$] and for the HI listeners [$R(72)=0.50, p<0.001$] were moderate but also indicate that neither group adopted optimal decision rules derived from estimates of sensorineural resolution. w_{PLIO} are roughly equivalent to normalized values of σ_P for each listener. Perhaps weights were influenced by un-normalized values of σ_P, the degree of sensorineural distortion. To test this, the weights for each frequency were correlated with the logarithm of σ_P across both groups of listeners. There was a significant positive relationship between sensorineural distortion and the 257 Hz weights [$R(23)=0.61, p=0.001$], but significant negative relationships between sensorineural distortion and the 1197, 2212, and 3951 Hz weights [$R(23)=0.40$, $p<0.05$; $R(23)=0.61$, $p=0.001$; and $R(23)=0.51$, $p=0.01$, respectively]. The 603 Hz weight had a marginally significant negative relationship with sensorineural distortion [$R(23)=0.37, p=0.069$]. This is the same general pattern observed with QT. As overall sensory information becomes distorted, the 257 Hz weight takes on greater importance at the expense of the weights higher in frequency.

In summary, weighting functions for the NH and HI groups were quite variable across listeners. There was an overall tendency for the lower frequencies to be given the greatest weight by both groups. A clear pattern emerged with how HI listeners weighted the lowest-frequency tones centered around 257 Hz. As overall hearing loss increased and sensorineural resolution decreased, the 257 Hz weight increased at the expense of a decrease in the higher-frequency weights. A more formal assessment of whether NH and HI listeners differed in terms of how efficiently they weighted information will require that the weights and sensorineural noise estimates be integrated into the theoretical framework of the PLIO.

3. Sensitivity

The theoretical framework above introduced four decision models that progressively account for factors that influence listeners' sensitivity when discriminating stimuli in a multiple-observation, sample discrimination task. The stimulus ideal observer model, d'_{ideal}, accounts for the stimulus microstructure—the mean frequency differences across trial intervals for each pair of tones and the trial-by-trial perturbations in frequency. The PLIO model, d'_{PLIO}, accounts for sensorineural noise in both the derivation of the weights (w_{PLIO}) and in the variance of the expected decision variable. The models for d'_{wgt} and d'_{pred} account for listeners' unique weighting strategies, with the latter also accounting for the effects of limited sensorineural resolution in the standard deviation of the expected decision variable. Any remaining difference between d'_{pred} and observed sensitivity is attributed to central noise. The top panels in Fig. 4 display the means of these four estimates of sensitivity for the NH and HI listeners as a function of the expected frequency difference, δ_i. Predictions for each model are plotted as different line types and the observed sensitivities, d'_{obs}, are plotted as data points with 95% CIs.

From Fig. 4, it is clear that the observed sensitivities for both groups of listeners are substantially less than the stimulus ideal, d'_{ideal} (dotted line). The differences between these two reflect the combined influences of listeners' inefficient weighting strategies and internal noise. Compared to d'_{ideal}, the predictions for d'_{wgt} (dashed line) are much closer to d'_{obs}, which indicates that a majority of listeners' less-than-ideal sensitivity can be attributed to the inefficient weighting strategies discussed in the previous section. Still, estimates of d'_{wgt} are consistently greater than d'_{obs}, especially when the expected frequency differences are relatively small. This indicates that internal noise also contributes to the reduction in observed sensitivity. The differences in sensitivity between

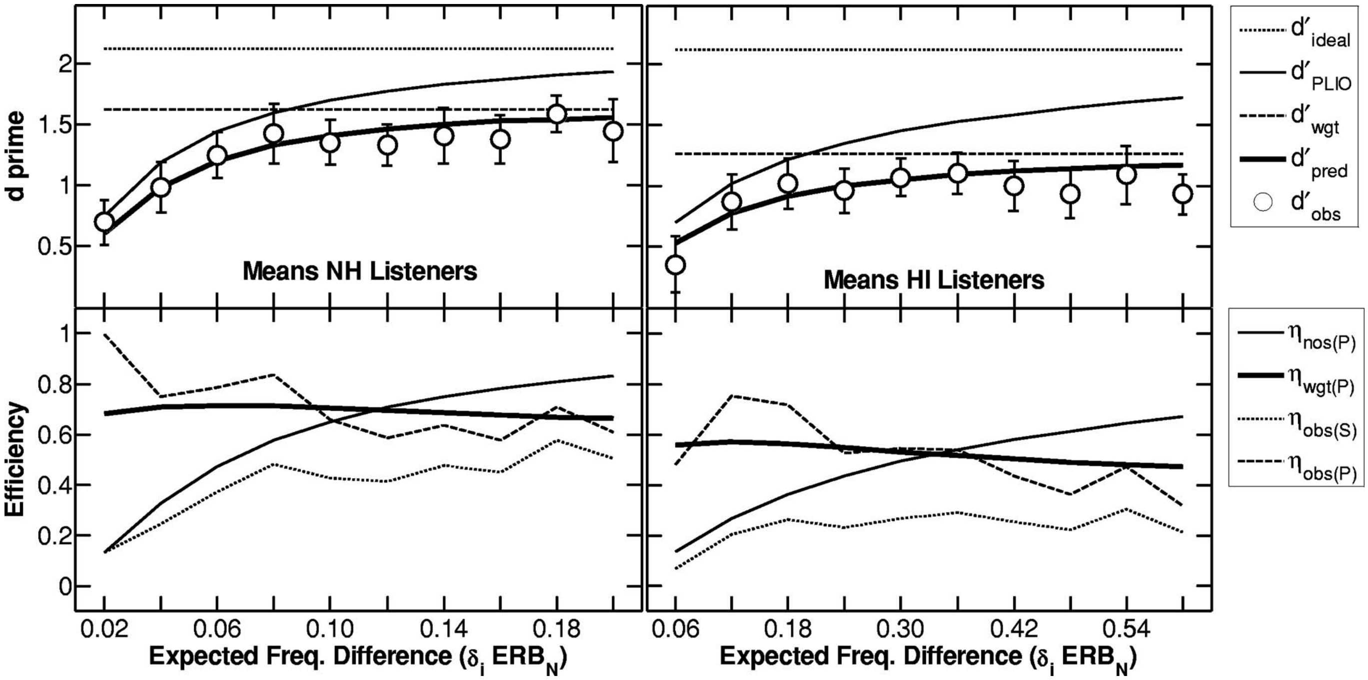

FIG. 4. For each decision model, the average estimates of sensitivity as a function of the expected frequency difference between intervals are displayed as different lines for the NH and HI listeners (left and right panels, respectively) in the top panels. The observed sensitivities are plotted as data points with 95% confidence intervals. Efficiencies, which compare different pairs of d-prime estimates from the top panels, are displayed in the bottom panels.

the stimulus ideal observer and the PLIO, d'_{PLIO} (thin solid line), reflects the hypothesized contribution of sensorineural noise to the observed sensitivity. As expected, the effects of limited frequency resolution are greatest when the frequency manipulations are smallest. As the size of the frequency manipulations increase relative to the sensorineural noise, the mean sensitivity for the PLIO increases and approaches sensitivity for the stimulus ideal. The predicted sensitivity for d'_{pred} (thick solid line) combines the contributions of listeners' inefficient weighting strategies with sensorineural noise. Estimates for d'_{pred} are lower than d'_{PLIO} which reflects the fact that even when compared to a PLIO most listeners employ inefficient weighting strategies. Estimates for d'_{pred} very closely approximate d'_{obs}. This indicates that for this experiment, sensorineural noise alone seems to account for most of the reduction in observed sensitivity that would otherwise be attributed to internal noise generally or to central noise specifically.

The left half of Table IV presents the means of the sensitivity estimates for the different decision models across the 10 trial blocks for each listener and for each group as a whole. Sensitivity for the stimulus ideal is not displayed because d'_{ideal} was a constant 2.12 for each listener. As displayed in the last row of the table, sensitivity estimates for each model (excluding d'_{ideal}) were significantly less for the HI listeners compared to the NH listeners. The significant difference in d'_{obs} indicates that HI listeners were less sensitive than NH listeners even though frequency manipulations were three times greater for the former. The significant differences between the groups in d'_{wgt}, d'_{pred}, and d'_{PLIO} indicate that both less efficient weighting strategies and greater sensorineural noise in the HI listeners were responsible for the differences in observed sensitivity.

C. Efficiency

1. Theoretical overview

To isolate the different factors that influence listeners' observed sensitivities, the above model estimates for sensitivity need to be compared. For this purpose, we adopt a normalizing statistic akin to R^2 with a long history in the literature. With origins in energy-detector analogs for observers, efficiency is defined as a ratio of variances and is computed by taking the ratio of squared d primes (Tanner and Birdsall, 1958). Within our theoretic treatment of the data, this property makes the efficiency measure more attractive than simply taking the difference between two d primes. For example, an efficiency of 0.5 could indicate that the pooled variance of the expected decision variable differs by a factor of 2 between the two models or that the expected values of the decision variables differ by a factor of $\sqrt{2}$. Most often, the expected decision variables will differ in both the numerator (mean) and denominator (variance). Efficiency, denoted η, is 1 when the two models under consideration have the same sensitivity.

Efficiency analysis as outlined by Berg (1990) provides a useful framework for partitioning the effects of decision weights (η_{wgt}) and internal noise (η_{nos}) on a listener's sensitivity relative to the stimulus ideal observer (η_{obs}). We expand on Berg's convention to (1) account for the fact that the ideal weighting strategy will differ for listeners and conditions as a function of information received, (2) attempt to partial out the sensorineural and central components of internal noise, and (3) isolate as much as possible the effects sensorineural noise on the observed efficiency.

Table V lists the different efficiencies used in our analysis along with the ratios (model comparisons) that define

TABLE IV. Displayed are the means of the sensitivity estimates (d') for the different decision models and efficiency measures (η) for the NH and HI listeners collapsed across the ten trial blocks, rank ordered for each group according to d'_{obs} (see text). Note that $d'_{\text{ideal}}=2.12$ for all listeners. The means and standard errors of the mean for each group are also provided.

	d'_{obs}	d'_{wgt}	d'_{PLIO}	d'_{pred}	$\eta_{\text{obs}(S)}$	$\eta_{\text{wgt}(P)}$	$\eta_{\text{nos}(P)}$	$\eta_{\text{nos}(C)}$	$\eta_{\text{obs}(P)}$
NH 08	1.59	1.84	1.66	1.50	0.56	0.82	0.61	1.12	0.92
NH 10	1.55	1.87	1.31	1.26	0.53	0.91	0.38	1.53	1.39
NH 07	1.42	1.76	1.44	1.38	0.45	0.91	0.46	1.07	0.98
NH 06	1.41	1.70	1.70	1.56	0.44	0.83	0.64	0.82	0.69
NH 12	1.37	1.67	1.72	1.26	0.42	0.54	0.66	1.18	0.63
NH 02	1.36	1.42	1.70	1.15	0.41	0.46	0.64	1.39	0.64
NH 11	1.35	1.65	1.68	1.46	0.41	0.76	0.63	0.86	0.65
NH 14	1.29	1.44	1.56	1.31	0.37	0.71	0.54	0.97	0.69
NH 04	1.26	1.74	1.64	1.57	0.35	0.91	0.60	0.65	0.59
NH 09	1.16	1.50	1.55	1.21	0.30	0.61	0.54	0.92	0.56
NH 01	1.00	1.82	1.76	1.37	0.22	0.61	0.69	0.53	0.32
NH 13	0.95	1.24	1.57	0.96	0.20	0.38	0.55	0.98	0.37
NH 03	0.95	1.44	1.51	1.06	0.20	0.49	0.51	0.81	0.40
Mean	**1.28**	**1.62**	**1.60**	**1.31**	**0.38**	**0.69**	**0.57**	**0.99**	**0.68**
(SE)	(0.06)	(0.06)	(0.04)	(0.05)	(0.03)	(0.05)	(0.03)	(0.08)	(0.08)
HI 09	1.19	1.27	1.46	1.06	0.31	0.53	0.47	1.25	0.66
HI 01	1.14	1.44	1.81	1.40	0.29	0.61	0.72	0.66	0.40
HI 17	1.14	1.54	1.76	1.41	0.29	0.64	0.69	0.65	0.42
HI 04	1.12	1.19	1.50	1.02	0.28	0.47	0.50	1.19	0.56
HI 13	1.06	1.46	1.54	1.25	0.25	0.66	0.53	0.72	0.48
HI 11	1.03	1.27	1.70	1.18	0.24	0.48	0.64	0.77	0.37
HI 15	1.01	1.42	1.60	1.14	0.23	0.51	0.57	0.79	0.40
HI 10	0.95	1.10	1.37	0.78	0.20	0.32	0.42	1.51	0.48
HI 02	0.95	1.19	1.49	1.11	0.20	0.56	0.49	0.74	0.41
HI 18	0.79	1.03	1.14	0.59	0.14	0.26	0.29	1.80	0.48
HI 05	0.76	0.87	1.01	0.63	0.13	0.40	0.23	1.44	0.57
HI 12	0.53	1.43	0.72	0.55	0.06	0.59	0.12	0.92	0.54
HI 03	0.44	1.20	0.98	0.78	0.04	0.64	0.21	0.31	0.20
Mean	**0.93**	**1.26**	**1.39**	**0.99**	**0.20**	**0.51**	**0.45**	**0.98**	**0.46**
(SE)	(0.07)	(0.06)	(0.10)	(0.09)	(0.03)	(0.04)	(0.06)	(0.12)	(0.03)
t	4.0[a]	4.7[a]	2.1[c]	3.3[b]	4.2[a]	2.8[b]	2.1[c]	0.1	2.6[c]

The last row displays the results of *t* tests with 24 degrees of freedom for each estimate between the NH and HI listeners
$p \leqslant 0.001$;
$p \leqslant 0.01$;
$p \leqslant 0.05$.

them and the theoretical constructs they isolate. Observer efficiency, $\eta_{\text{obs}(S)}$, is the same as Berg's original η_{obs} but is subscripted (S) to denote that the observed sensitivity is compared to the stimulus ideal. Differences between the two sensitivities are attributed to observed weights that deviate from the stimulus ideal (w_{ideal}), to sensorineural noise, and to central noise. Unlike Berg's original η_{wgt}, which compares the observed weighting strategy to one based on an index of information sent, we are interested in how efficiently listeners weight information compared to an ideal model that weights information proportional to the reliability of information received, w_{PLIO}. This comparison is particularly important for the present study because the reliability of information received varies considerably between listeners, conditions, and frequencies. We introduce a modified weight efficiency, $\eta_{\text{wgt}(P)}$, that is obtained by comparing two estimates of sensitivity that include the listener's unique profile of sensorineural noise. One of these estimates uses the listener's unique weighting strategy, d'_{pred}, and the other estimate uses the ideal weighing strategy based on information received, d'_{PLIO}.

TABLE V. Analytic definitions and constructs of interest for each of the efficiency measures.

Efficiency	Notation	Analytic definition	Constructs
Observer	$\eta_{\text{obs}}(S)$	$(d'_{\text{obs}}/d'_{\text{ideal}})^2$	w_{obs} *vs.* w_{ideal} σ_P^2 and σ_C^2
Weight	$\eta_{\text{wgt}}(P)$	$(d'_{\text{pred}}/d'_{\text{PLIO}})^2$	w_{obs} *vs.* w_{PLIO}
Sensorineural noise	$\eta_{\text{nos}}(P)$	$(d'_{\text{PLIO}}/d'_{\text{ideal}})^2$	w_{PLIO} *vs.* w_{ideal} σ_P^2
Central noise	$\eta_{\text{nos}}(C)$	$(d'_{\text{obs}}/d'_{\text{pred}})^2$	σ_C^2
Adjusted observer	$\eta_{\text{obs}}(P)$	$(d'_{\text{obs}}/d'_{\text{PLIO}})^2$	w_{obs} *vs.* w_{PLIO} σ_C^2

In an attempt to separate the effects of sensorineural and central noise, we define $\eta_{\text{nos}(P)}$ and $\eta_{\text{nos}(C)}$, respectively (cf. Appendix). By $\eta_{\text{nos}(P)}$ we mean the loss of efficiency due to sensorineural noise. Again, the effects of sensorineural noise are twofold. First, sensorineural noise alters how an ideal observer should weight information. Second, sensorineural noise adds to the variance of the expected decision variable. To index the cumulative effects of sensorineural noise, we compare the sensitivity of the PLIO (d'_{PLIO}) to that of the stimulus ideal (d'_{ideal}). By our definition, central noise is the residual variance in the expected decision variable after accounting for sensorineural noise and the listener's weighting strategy. Therefore, $\eta_{\mathrm{nos}(C)}$ is obtained by comparing the observed sensitivity of the listener (d'_{obs}) to the sensitivity of a hypothetical listener who shares the same weighting strategy and sensorineural limitations (d'_{pred}).

With the above analytic definitions in place, the efficiency of the observed sensitivity relative to the stimulus ideal, $\eta_{\mathrm{obs}(S)}$, can be partitioned as follows:

$$\eta_{\mathrm{obs}(S)} = \eta_{\mathrm{wgt}(P)} \times \eta_{\mathrm{nos}(P)} \times \eta_{\mathrm{nos}(C)}. \tag{4}$$

To isolate as much as possible the effects of sensorineural noise on observer efficiency for a fair comparison between NH and HI listeners, the listener's observed sensitivity (d'_{obs}) should be compared not to the stimulus ideal observer (d'_{ideal}), but to a peripherally limited observer who weights the information received at each frequency optimally (d'_{PLIO}). We call this comparison the adjusted observer efficiency, $\eta_{\mathrm{obs}(P)}$, which is obtained by dividing both sides of Eq. (4) by $\eta_{\mathrm{nos}(P)}$. The resultant equation, Eq. (5), is reminiscent of Berg's original efficiency analysis

$$\eta_{\mathrm{obs}(P)} = \eta_{\mathrm{wgt}(P)} \times \eta_{\mathrm{nos}(C)}. \tag{5}$$

2. Results

The relationships between the sensitivity estimates of the different decision models are shown as efficiencies in the bottom panels of Fig. 4. Observed sensitivity for both NH and HI listeners were limited by sensorineural noise, which diminished the reliability of information received when the frequency manipulations were relatively small. This is indicated by $\eta_{\mathrm{nos}(P)}$, which was about 0.2 at the smallest frequency manipulations for both groups. At the largest frequency manipulations, $\eta_{\mathrm{nos}(P)}$ was about 0.8 for the NH listeners and about 0.65 for the HI listeners. Because $\eta_{\mathrm{nos}(P)}$ was much less than 1 for the HI listeners, this indicates that there is still potential for improvement in observed sensitivity at larger frequency manipulations than what was tested. After accounting for the effects of sensorineural noise in the listeners' observed sensitivity, $\eta_{\mathrm{obs}(P)}$ was much higher compared to the traditional assessment of observed sensitivity, $\eta_{\mathrm{obs}(S)}$, especially when the expected frequency differences were relatively small compared to sensorineural resolution.

The right half of Table IV presents the efficiency measures derived from the means of the different sensitivity estimates for each listener and for each group as a whole. As displayed in the last row of the table, all efficiency measures were significantly less for the HI listeners compared to the NH listeners, except $\eta_{\mathrm{nos}(C)}$. Observer efficiency (re: the stimulus ideal), $\eta_{\mathrm{obs}(S)}$, was significantly lower for HI listeners compared to the NH listeners. A significant contributing factor to this difference was sensorineural noise efficiency. However, even when differences in sensorineural noise were accounted for, observer efficiency (re: the PLIO), $\eta_{\mathrm{obs}(P)}$, was still significantly lower for the HI listeners. This is because the exaggerated low-frequency emphasis in the weighting strategies of the HI listeners resulted in significantly lower weighting efficiencies compared to the NH listeners, even though differences in sensorineural resolution were accounted for in the analysis of $\eta_{\mathrm{wgt}(P)}$. $\eta_{\mathrm{nos}(C)}$ was not statistically different between the NH and HI listeners indicating that SNHL did not influence the variability or consistency with which the decision strategies were used across trials and conditions.[4]

IV. DISCUSSION

One of the primary objectives of this study was to adapt the theory of ideal observers to include the effects of frequency-specific sensorineural noise, what we term the peripherally limited ideal observer, or PLIO, decision model. This allows us to shift the research focus from information sent to information received when trying to understand differences in observed sensitivity between listeners. This is important because frequency resolution at the auditory periphery can differ substantially as a function of hearing loss and frequency, which has implications for how internal noise and ideal weights are conceptualized within the decision theoretic framework of the PLIO.

A. Sensorineural and central noise

Internal noise is often used as a general umbrella term that includes both the effects of imprecise coding of the physical properties of the stimulus at the auditory periphery and the effects of noisy decision processes arising from inconsistency and inattention on the part of the listener. We identify these two sources of internal noise as sensorineural and central noise, respectively. Using listener and frequency specific estimates of DLFs, the effects of sensorineural noise on observed sensitivity were modeled as additional sources of variance in the expected decision variable. When these effects were combined with a listener's unique weighting strategy, a specific prediction for observed sensitivity, d'_{pred}, was obtained. The intention was that any residual variance needed to bring this prediction closer to the observed sensitivity would be modeled as a nonintegrated constant associated with central noise. As it turned out, the inclusion of sensorineural noise was adequate at bringing the estimates of d'_{pred} very near the observed sensitivities. This indicates that SNHL did not influence central processes related to the implementation of a decision strategy across trials and conditions and that sensorineural noise was the primary component of the overall internal noise in this discrimination task for both groups of listeners.

B. Decision weights

Within our theoretical framework for the PLIO, ideal weights should be proportional to the reliability of information received, which takes into account each listener's unique profile of sensorineural noise. In this study, most listeners did not employ optimal weighting strategies. Instead, higher-frequency tones ($\geqslant$2212 Hz) were given very little weight by individuals from either group of listeners compared to lower-frequency tones ($\leqslant$1197 Hz). Furthermore, as hearing loss and sensorineural noise increased there was a tendency for the 257 Hz weight to increase at the expense of a decrease in the other weights. This resulted in significantly lower weighting efficiencies for the HI listeners.

There are at least a few explanations for why the lower-frequency tones were on average given significantly greater weight than the higher-frequency tones.[5] First, it is worth noting that there were a number of instances where the weights in the high-frequency region were significantly greater than zero for the NH listeners. This was especially true for the first author, a highly practiced listener on this task, who put the greatest weight on the highest-frequency tone. This indicates that attentional mechanisms influence the magnitude of the high-frequency weights to some extent, rather than a simple failure to hear the individual tones due to upward spread of masking, for example. A preference for low-frequency information, especially on a relative scale, has been demonstrated elsewhere.

In a series of experiments, Neff and Odgaard (2004) had listeners engage in sample discrimination of frequency experiments similar in nature to the present study except listeners were asked to focus selectively on a single frequency difference near 2000 Hz instead of integrating information across six simultaneous differences. Additional frequency information above and below the target frequency served as distracters and consisted of either random-frequency tones, fixed-frequency tones with and without varying level, or noise bands. For random-frequency tones, weighting analyses revealed a dominance of the lower-frequency distracters over the higher-frequency distracters. This relative patten persisted even when the distracters and targets were shifted two octaves higher in frequency. Furthermore, the critical feature of the low-frequency distracters was informational (variation) rather than energetic (merely being present) since fixed-frequency tones and noise bands produced little to no negative effects on target discrimination.

Alexander (2004, pp. 110–154; Alexander and Lutfi, 2003) reported that in tone detection tasks involving random-frequency distracters and fixed-frequency targets at 800, 2000, or 5000 Hz that (1) detection thresholds increased with increases in the informational content of distracters below the target frequency compared to those above and (2) distracters received significantly greater weight when they were immediately below the target frequency compared to when they were immediately above. Alexander also found that for a 2000 Hz target, randomly turning on and off the below-target distracters from trial to trial (informational masking) while keeping the above-target distracters always on significantly increased thresholds for target detection compared to the opposite when the below-target distracters were static (energetic masking) and the above-target distracters were random.

We carried out two related experiments using the same listeners and almost the same conditions. In one experiment, the high-frequency tones were randomly selected as before except that they were selected from the exact same distribution for the signal and noise intervals so that the expected frequency difference for these tone pairs was zero. Since the high-frequency tones had no informational value across the experiment (i.e., w_{PLIO} equaled zero), they were distracters and listeners were reminded to ignore them. Likewise, in a second experiment, the low-frequency tones had expected frequency differences of zero and served as distracters. Performance in the high-frequency distracter experiment yielded weighting functions and sensitivity estimates very close to the current experiment. However, in the low-frequency distracter experiment d'_{obs} and d'_{pred} were very close to zero for most of the listeners from both groups since the weighting patterns were still low-frequency dominant. Like the Neff and Odgaard (2004) and Alexander (2004; Alexander and Lutfi, 2003) studies, even when it was highly disadvantageous to attend to the low-frequency information, listeners could not ignore it.

It seems clear that listeners enter the laboratory setting with a preset weighting pattern for pitch-like discriminations. Relatively low frequencies tend to dominate the perception of tonal complexes. Perhaps only with a lot of practice and motivation can this bias be overcome. This preset weighing strategy is determined by a lifetime of experience attending to real-world sounds, especially the most functionally relevant sounds like speech. It is known that the energy of real environments and especially of speech and music is distributed much like pink noise and decreases as a function of frequency (e.g., Voss and Clarke, 1975, 1978). A person with a longstanding high-frequency sensorineural hearing loss experiences an exaggerated form of this low-frequency dominance. While the above explanation is speculative, the fact remains that people will bring to the laboratory setting an ingrained listening bias that may not conform to our task demands. An ecologically informed ideal observer model would appreciate the statistics of the acoustics of the everyday world. Within this framework, predictions for sensitivity will better accord with listeners' performance. Future research can benefit from exploring how experience with hearing impairment shapes the way people listen to complex sounds that vary along several acoustic dimensions, like speech. Casting performance in terms of an ecologically informed ideal observer that includes peripheral limitations as well as environmental experience will improve on models of speech intelligibility, including the effects of noise, the spatial distribution of multiple talkers, and various hearing aid and cochlear implant processing algorithms.

ACKNOWLEDGMENTS

This work was part of a dissertation submitted by the first author in partial fulfillment of the requirements for the Ph.D. in Communicative Disorders at the University of Wisconsin-Madison. The authors would like to thank Karen Malott and Laurie Canadeo for their assistance during data collection. The authors would especially like to thank Dr. Bruce Berg for his generous comments that helped to clarify the technical details of the decision theory equations and notation and an anonymous reviewer whose comments on earlier drafts proved extremely helpful. This research was supported by a grant to the first author from The National Organization for Hearing Research Foundation and a grant to the second author from the NIDCD (R01 DC1262-10). This manuscript was written while supported by NIDCD grants R01 DC04072 (Keith R. Kluender) and T32 DC000013 (BT-NRH)

APPENDIX: THE PARTITION OF NOISE EFFICIENCY

For the sake of simplicity, assume that a hypothetical listener has equal σ_P for each tone so that the weights for d'_{PLIO} are equal to each other and are identical to the weights for d'_{ideal}. Furthermore, assume that the listener employs an optimal listening strategy so that the weights can be treated as constants and disregarded altogether. Therefore, from Eq. (3) we have

$$\eta_{\mathrm{nos}(P)} = \left(\frac{d'_{\mathrm{PLIO}}}{d'_{\mathrm{ideal}}}\right)^2 = \frac{2(m\delta)^2}{m2\sigma_S^2 + m2\sigma_P^2} \times \frac{m2\sigma_S^2}{2(m\delta)^2} = \frac{m2\sigma_S^2}{m2\sigma_S^2 + m2\sigma_P^2}, \tag{A1}$$

$$\eta_{\mathrm{nos}(C)} = \left(\frac{d'_{\mathrm{obs}}}{d'_{\mathrm{pred}}}\right)^2 = \frac{2(2m\delta)^2}{m2\sigma_S^2 + m2\sigma_P^2 + \sigma_C^2} \times \frac{m2\sigma_S^2 + m2\sigma_P^2}{2(m\delta)^2} = \frac{m2\sigma_S^2 + m2\sigma_P^2}{m2\sigma_S^2 + m2\sigma_P^2 + \sigma_C^2}, \tag{A2}$$

$$\eta_{\mathrm{nos}} = \eta_{\mathrm{Pnos}} \times \eta_{\mathrm{Cnos}} = \frac{m2\sigma_S^2}{m2\sigma_S^2 + m2\sigma_P^2 + \sigma_C^2}. \tag{A3}$$

[1]Because there was no overlap in age between the NH and HI listeners, we cannot be sure to what extent the obtained results, such as estimates of frequency resolution, reflect processes related to aging *per se*. Fortunately, this aspect is accounted for because sensorineural resolution is input for the different decision models. It is also possible that age-related factors influenced how listeners weighted information in the multiple-observation, sample discrimination experiment, although, at this time, it is not entirely clear how or why.

[2]Sometimes over the course of the four blocks there was improvement for one or more frequencies. The blocks and frequencies for which this occurred did not seem systematic within or between listeners. Additional blocks were run until there was no substantial improvement for four successive blocks. Paired t tests comparing DLFs obtained from the two trial blocks immediately before the sample discrimination experiment and from the last two trial blocks obtained after the sample discrimination experiment were not statistically significant for any of the six tones in either group ($p \geqslant 0.10$ for all comparisons). This indicates that there was a lack of a practice effect over the course of the experimental protocol.

[3]An index of sensitivity can at least be calculated by (l) the difference in the Z-transformed hit and false-alarm rates, or (2) the difference in the means of the signal and noise distributions of the decision variable (Δ) divided by their common standard deviation (σ). For a yes-no (y/n) task these two computations are the same. However, for a 2IFC task the mean difference (the numerator) is integrated by a factor of 2, while the standard deviation (the denominator) is integrated by $\sqrt{2}$. This leads to an overall increase of $\sqrt{2}$ between the difference in the Z-transformed hit and false-alarm rates for the 2IFC task compared to the y/n task. Because it is desirable to have an index of sensitivity that is task independent, the Δ/σ ratio is the preferred index for d', meaning that the difference in the Z-transformed hit and false-alarm rates for the 2IFC task must divided by $\sqrt{2}$. The difference in the Z-score computation of sensitivity between the two tasks reflects the fact that greater sensitivity is needed to achieve the same performance or percent correct on a y/n task compared to a 2IFC task.

[4]While at first it seems that central noise played little role in this multiple-observation, sample discrimination task, it is likely that each estimate of sensorineural noise collected during the DLF procedure already included some central noise. The actual central noise during the sample discrimination experiment might be only some small factor of the central noise collected during the estimate of one DLF. Therefore, analytically integrating sensorineural noise (and central noise) across the six frequency regions might lead to an over estimate of the total amount of sensorineural noise. In this regard, d'_{pred} and d'_{wgt} might be considered as lower and upper bounds, respectively, for modeling the influence of sensorineural noise on observed sensitivity. This combined with the fact that predictions for sensorineural noise and sensitivity were analytically derived from expected values might explain the few instances where d'_{pred} under-predicts d'_{obs} and where $\eta_{\mathrm{nos}(C)}$ is greater than one.

[5]One review noted that the level of the broadband stimuli in this experiment (about 50 and 85 dB SL for the HI and NH listeners, respectively) possibly elicited the acoustic reflex in some or most listeners. Because maximal attenuation from the graded acoustic reflex occurs between 500 and 1000 Hz (about 20–25 dB with 10 dB/oct. slopes), it could be responsible in part for the observed low-frequency dominance in the weighting patterns. Presumably, their reduced loudness would cause them to "pop-out" perceptually (unlikely) or their sensorineural resolution would be enhanced because the broadening of auditory filters associated with high-presentation levels would be somewhat mitigated by the attenuation. The following results suggest that the acoustic reflex was not responsible for the low-frequency dominance of the observed weights. Weights were also obtained from a subset of listeners ($n=10$ for each group) who replicated this study using a shorter stimulus duration (40 ms versus 120 ms) in Alexander (2004). Because the acoustic reflex is dependent on a temporal summation of energy, its effects should have been diminished at the shorter duration. This being said, paired-sample t tests revealed that no comparisons were statistically significant except for the HI listeners for the 257 and 603 Hz weights. The shorter duration was associated with significantly lower weights at 257 Hz and with significantly higher weights at 603 Hz compared to the longer duration ($p<0.05$). However, this might be because sensorineural resolution was significantly worse at 257 Hz for the shorter duration compared to the longer duration.

Alexander, J. M., and Lutfi, R. A. (**2003**). "Upward spread of informational masking in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **113**, 2287.

Alexander, J. M., and Lutfi, R. A. (**2004**). "Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights," J. Acoust. Soc. Am. **116**, 2234–2247.

Alexander, J. M. (**2004**). Decision Factors in Multitone Detection and Discrimination by Listeners with Normal and Impaired Hearing (Unpublished dissertation, University of Wisconsin-Madison).

ANSI. (**1996**). ANSI S3.6-1996, "American National Standards specification for audiometers," (American National Standards Institute, New York).

Berg, B. G. (**1990**). "Observer-efficiency and weights in a multiple observation task," J. Acoust. Soc. Am. **88**, 149–l58.

Buus, S., Florentine, M., and Zwicker, T. (**1995**). "Psychometric functions for level discrimination in cochlearly impaired and normal listeners with equivalent-threshold masking," J. Acoust. Soc. Am. **98**, 853–861.

Doherty, K. A., and Lutfi, R. A. (**1996**). "Spectral weights for overall level discrimination in listeners with sensorineural hearing loss," J. Acoust. Soc. Am. **99**, 1053–1057.

Doherty, K. A., and Lutfi, R. A. (**1999**). "Level discrimination of single tones in a multitone complex by normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **105**, 1831–1840.

Florentine, M., Reed, C. M., Rabinowitz, W. M., Braida, L. D., Durlach, N. I., and Buus, S. (**1993**). "Intensity perception. XIV. Intensity discrimination in listeners with sensorineural hearing loss," J. Acoust. Soc. Am. **94**, 2575–2586.

Freyman, R. L., and Nelson, D. A. (**1991**). "Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **34**, 1371–1386.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Jesteadt, W., Nizami, L., and Schairer, K. S. (**2003**). "A measure of internal noise based on sample discrimination," J. Acoust. Soc. Am. **114**, 2147–2157.

Lentz, J. J., and Leek, M. R. (**2002**). "Decision strategies of hearing-impaired listeners in spectral shape discrimination," J. Acoust. Soc. Am. **111**, 1389–1398.

Lentz, J. J., and Leek, M. R. (**2003**). "Spectral-shape discrimination by hearing-impaired and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 1604–1616.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A. (**1989**). "Informational processing of complex sound. I: Intensity discrimination," J. Acoust. Soc. Am. **86**, 934–944.

Lutfi, R. A. (**1990**). "Informational processing of complex sound. II: Cross-dimensional analysis," J. Acoust. Soc. Am. **87**, 2141–2148.

Lutfi, R. A. (**1992**). "Informational processing of complex sound. III: Interference," J. Acoust. Soc. Am. **91**, 3391–3401.

Lutfi, R. A., Doherty, K. A., and Oh, E. (**1996**). "Psychometric functions for the discrimination of spectral variance," J. Acoust. Soc. Am. **100**, 2258–2265.

Macmillan, N. A., and Creelman, C. D. (**2005**). *Detection Theory: A User's Guide*, 2nd ed. (Erlbaum, Mahwak, NJ).

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing*, 4th ed. (Academic, San Diego).

Neff, D. L., and Odgaard, E. C. (**2004**). "Sample discrimination of frequency differences with distracters," J. Acoust. Soc. Am. **116**, 3051–3061.

Sek, A., and Moore, B. C. J. (**1995**). "Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am. **97**, 2479–2486.

Sorkin, R. D., Robinson, D. E., and Berg, B. G. (**1987**). "A detection theory method for the analysis of visual and auditory displays," in *Proc. Human Factors Soc., 31st Annual Meeting*, pp. 1184–1188.

Tanner, W. P., and Birdsall, T. G. (**1958**). "Definitions of d' and η as psychological measures," J. Acoust. Soc. Am. **30**, 922–928.

Voss, R. F., and Clarke, J. (**1975**). "'$1/f$ noise' in music and speech," Nature (London) **258**, 317–318.

Voss, R. F., and Clarke, J. (**1978**). "'$1/f$ noise' in music: Music from $1/f$ noise," J. Acoust. Soc. Am. **63**, 258–263.