Binaural detection with narrowband and wideband reproducible noise maskers: I. Results for human

Mary E. Evilsizer
*Hearing Research Center, Department of Biomedical Engineering, Boston University,
44 Cummington Street, Boston, Massachusetts 02215*

Robert H. Gilkey
Hearing Research Center, Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, Department of Psychology, Wright State University, Dayton, Ohio, and Human Effectiveness Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio

Christine R. Mason, H. Steven Colburn, and Laurel H. Carney[a]
*Hearing Research Center, Department of Biomedical Engineering, Boston University,
44 Cummington Street, Boston, Massachusetts 02215*

(Received 12 February 2001; revised 29 September 2001; accepted 1 October 2001)

This study investigated binaural detection of tonal targets (500 Hz) using sets of individual masker waveforms with two different bandwidths. Previous studies of binaural detection with wideband noise maskers show that responses to individual noise waveforms are correlated between diotic (N_0S_0) and dichotic (N_0S_π) conditions [Gilkey *et al.*, J. Acoust. Soc. Am. **78**, 1207–1219 (1985)]; however, results for narrowband maskers are not correlated across interaural configurations [Isabelle and Colburn, J. Acoust. Soc. Am. **89**, 352–359 (1991)]. This study was designed to allow direct comparison, in detail, of responses across bandwidths and interaural configurations. Subjects were tested on a binaural detection task using both narrowband (100-Hz bandwidth) and wideband (100 Hz to 3 kHz) noise maskers that had identical spectral components in the 100-Hz frequency band surrounding the tone frequency. The results of this study were consistent with the previous studies: N_0S_0 and N_0S_π responses were more strongly correlated for wideband maskers than for narrowband maskers. Differences in the results for these two bandwidths suggest that binaural detection is not determined solely by the masker spectrum within the critical band centered on the target frequency, but rather that remote frequencies must be included in the analysis and modeling of binaural detection with wideband maskers. Results across the set of individual noises obtained with the fixed-level testing were comparable to those obtained with a tracking procedure which was similar to the procedure used in a companion study of rabbit subjects [Zheng *et al.*, J. Acoust. Soc. Am. **111**, 346–356 (2002)]. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1423929]

PACS numbers: 43.66.Pn, 43.66.Dc [LRB]

I. INTRODUCTION

The task of detecting a pure-tone signal in a noise masker has been a critical tool used by psychophysicists to probe the mechanisms of hearing (e.g., Fletcher, 1940; Helmholtz, 1863). This simple task has been a building block of auditory theory, playing a role in the development of concepts, such as the critical band filter, and models for the integration of information across time as well as for the fundamental mechanisms of binaural hearing. Nevertheless, the mechanisms with which normal-hearing listeners perform this basic task are still not completely understood. Although it is often assumed that listeners base their decision on stimulus energy or a closely related statistic, this assumption has repeatedly been shown to conflict with observed data (see, for example, Gilkey, 1987; Kidd, 1987; Kidd *et al.*, 1989; Gilkey and Robinson, 1986; Isabelle and Colburn, 1991; Isabelle, 1995; Richards *et al.*, 1991; Richards, 1992; Richards and Nekrich, 1993). Moreover, evidence supporting the processing mechanisms incorporated in the classical psychophysical models are not obvious in recent physiological measurements (Young and Barta, 1986; Miller *et al.*, 1987; Rees and Palmer, 1988; Jiang *et al.*, 1997a, b; Palmer *et al.*, 1999, 2000). The limits of our understanding of the tone-in-nose detection task are perhaps most striking in the context of binaural masking experiments, where the relation between binaural and monaural processing is still not understood. Classical critical-band based models of binaural interaction fail to predict several significant features of the observed data, which suggests that integration of information across auditory channels must be included to explain the results (Zwicker and Henning, 1984; van de Par and Kohlrausch, 1999; Breebaart *et al.*, 2001; Trieurniet and Boucher, 2001). The failures of the classical models are even greater when predicting the results of binaural detection with reproducible noises, in which the details of responses to a set of individual

---

[a] Author to whom correspondence should be addressed. Present address: Department of Bioengineering and Neuroscience, Institute for Sensory Research, Syracuse University, 621 Skytop Rd., Syracuse, NY 13244; electronic mail: laurel_carney@isr.syr.edu

repeated noise waveforms is investigated (Gilkey *et al.*, 1985; Isabelle and Colburn, 1991; Isabelle, 1995; Gilkey, 1990; Colburn *et al.*, 1997).

This article reports the first experiments in a series of studies that will utilize psychophysical measurements from humans, psychophysical measurements from rabbits, physiological recordings from the inferior colliculus of rabbits, and computational modeling to explore the processing governing tone-in-noise detection. This set of studies is linked by a common set of stimulus manipulations (masker bandwidth and interaural signal phase) and by a common set of reproducible noise maskers, which will allow direct comparison across studies of psychophysical, physiological, and model responses to individual noise-alone and signal-plus-noise waveforms.

This article examines human monaural and binaural detection for both narrowband and wideband maskers, which were generated from the same 25 noise waveforms such that the spectral components in the 100-Hz frequency region surrounding the 500-Hz tone were identical under wideband and narrowband conditions. Most efforts to describe tone-in-noise detection have considered only the parameters of the noise process and have ignored the statistics of the particular noise waveforms presented. In a typical experiment, each noise waveform is presented only once, and the average performance across a large number of masker samples is studied. Green (1964) used the term ‘‘molar’’ to refer to performance averaged across the ensemble of masker waveforms in this way. Another method is to consider each stimulus and predict the subjects' responses on a trial-by-trial basis. Green argued that a complete understanding of tone-in-noise detection would allow the experimenter to predict this ‘‘molecular-level’’ performance. In practice, incomplete knowledge of the internal noise of the listener and the sequential dependencies across trials makes trial-by-trial predictions impractical. Instead, a ‘‘quasi-molecular’’ approach can be employed, in which a set of reproducible noise waveforms is presented on multiple trials and the average response to each individual masker is analyzed. Several investigators have studied tone-in-noise detection using this approach (e.g., Pfafflin and Mathews, 1966; Ahumada and Lovell, 1971; Ahumada *et al.*, 1975; Siegel and Colburn, 1983, 1989; Gilkey *et al.*, 1985; Isabelle and Colburn, 1991; Isabelle, 1995). In most of these quasi-molecular-level studies, performance is described in terms of the probability of a ‘‘target present’’ response in one-interval experiments. Thus one measures the probability of correct detection, or a ‘‘hit,’’ (P_h) and probability of a false alarm (P_f) for each sample in the set of reproducible noises.

Modeling studies have been only moderately successful at explaining subject performance in reproducible-noise experiments, but they have raised several interesting theoretical questions, especially when results are compared across experimental studies (Gilkey *et al.*, 1985; Isabelle and Colburn, 1991; Isabelle, 1995). To a first approximation, the differences in hit and false-alarm rates across the ensemble of reproducible noises in a monaural or diotic detection task with wideband maskers (Watson, 1962; Ahumada *et al.*, 1975; Gilkey and Robinson, 1986) can be explained by sample-to-sample differences in the output of a simple critical-band-based energy-detector model. Nevertheless, energy-based models do not account for a substantial portion of the total variance in subject responses (e.g., Gilkey and Robinson, 1986), suggesting that energy at the output of a narrowband filter tuned to the target frequency may contribute to, but does not completely determine, the differences in responses observed across samples. Moreover, simple manipulations of stimulus parameters, such as masker bandwidth and the interaural phase of the signal tone, yield results that are not predictable by energy-related models.

The current study was designed in part to explore an apparent incongruity between studies with different masker bandwidths, notably the studies of Gilkey *et al.* (1985) and Isabelle and Colburn (1991). With a 500-Hz target and wideband (100–3000 Hz) maskers, Gilkey *et al.* (1985) found that the across noise-sample pattern of responses (hit rates and false-alarm rates) for the diotic (N_0S_0) condition was correlated with the pattern for the dichotic (N_0S_π) condition, even though the signal level was 10–15 dB lower for the dichotic condition. In contrast, with a 500-Hz target and narrowband ($\frac{1}{3}$-oct band) maskers, Isabelle and Colburn (1991) found that hit and false-alarm rates were uncorrelated between the N_0S_0 and N_0S_π conditions for two of their three subjects. In subsequent work, they showed that the across-sample differences in hit rate with narrowband maskers were not well predicted by energy-related models, including the equalization-cancellation (EC) model (Durlach, 1963) and cross-correlation models (Isabelle, 1995; Colburn *et al.*, 1997). The difference in masker bandwidths used in the two studies with reproducible noise was hypothesized to be the most likely reason for the discrepancy between these results (Isabelle and Colburn, 1991). Yet, if detection is based on energy in the response of narrow (i.e., critical-band) filters, then performance for tone-detection tasks with narrowband and with wideband maskers should be similar. Studies of binaural detection using random (nonreproducible) noise maskers have also concluded that there are differences in processing strategies between wideband and narrowband masker conditions, and between N_0S_0 and N_0S_π, especially for narrowband maskers (e.g., van de Par and Kohlrausch, 1999; Breebaart *et al.*, 2001).

The hypothesis that differences in bandwidth explain the differences in results between studies with reproducible noises can be tested by using pairs of narrowband and wideband maskers generated such that they are identical in the narrow band around the target frequency (i.e., approximately a critical band) and differ only outside the frequency range of the narrowband masker. Gilkey (1990) reported preliminary results comparing wideband and narrowband maskers in the same subjects; unlike Isabelle and Colburn (1991), Gilkey found that false-alarm rates were correlated across interaural configurations even with the narrowband maskers (although less so than with the wideband maskers). However, some of the subjects in Gilkey's study had unusual thresholds under the narrowband N_0S_π condition, which were substantially higher than those of Isabelle and Colburn's subjects. Said differently, Gilkey's subjects had similar thresholds under the N_0S_0 and N_0S_π conditions, suggesting that they may have been using similar strategies under both conditions and may not have taken full advantage of the additional binaural cues available in the N_0S_π condition.

The current study consisted of two experiments. The first experiment tested both diotic (N_0S_0) and dichotic (N_0S_π) detection of fixed-level tones in narrowband and wideband noise maskers. In the second experiment, a tracking procedure was used to control the signal level. In both experiments, performance was examined across the ensemble of noise samples (molar level) and on a sample-by-sample basis (quasi-molecular level). Because a companion study in rabbit (Zheng *et al.*, 2002) used a tracking procedure to study behavioral performance across noises, it was important to determine whether sample-level data obtained with two different procedures were comparable.

II. GENERAL METHODS

The subjects were four undergraduate students (three male and one female) aged 18–20 years with normal hearing. None of the subjects had prior experience with auditory experiments. The subjects were tested individually in an IAC (Industrial Acoustics Co., Bronx, NY) double-walled sound-attenuating booth. Both the masker and target stimuli were generated and combined using TDT (Tucker-Davis Technologies, Gainesville, FL) programmable equipment and presented to the subject via TDH-39 (Telephonics Corp., Farmington, NY) headphones.

A. Stimuli

To compare subject performance across masker bandwidths, narrowband and wideband noise maskers were created with related spectra. Twenty-five independent wideband, reproducible noise maskers were created that had a rectangular spectral envelope with a bandwidth of 100 Hz to 3 kHz, chosen to be consistent with the wideband masker bandwidth used in the study by Gilkey *et al.* (1985). The narrowband noise samples were obtained by digitally filtering the wideband noise samples so that the narrowband and wideband noise samples had identical phase and power spectra, component-by-component, in the 100-Hz frequency band, geometrically centered around 500 Hz (452 to 552 Hz). The narrowband masker was chosen to be similar to that of the narrowband masker employed by Isabelle and Colburn (1991). The long-term expected spectrum level of both the wideband and narrowband maskers was 40-dB SPL. Both the 500-Hz target and the masker were 300 ms in duration including a 10-ms rise/fall time with a cosine-squared ramp.

For each bandwidth, the N_0S_0 and N_0S_π stimulus conditions were each presented for approximately the same number of trials. Two different starting phases were studied for the N_0S_0 condition; half of the N_0S_0 trials were created by adding the tone with a 0° starting phase [$N_0S_0(0°)$], and the other half were created using a 180° starting phase [$N_0S_0(180°)$]. These two starting phases represent the stimulus combinations that make up the N_0S_π stimulus; the results for both starting phases of the N_0S_0 stimuli are useful for understanding and modeling the relation between diotic and dichotic conditions. Performance in the diotic case varies with the starting phase of the target tone (Gilkey *et al.*, 1985; Isabelle and Colburn, 1991). When analyzing the responses to the N_0S_0 stimuli, 50 target-masker combinations (the two starting phases for 25 different reproducible noises) were considered.

B. Experiment 1: Binaural detection with narrowband and wideband maskers

In experiment 1, responses were collected for detection of a 500-Hz tone under diotic and dichotic conditions with the wideband and narrowband sets of reproducible noise maskers.

1. Methods

a. Training. Training consisted of three tasks: a two-interval two-alternative forced-choice (2I,2AFC) tracking task with feedback, a one-interval fixed-level task with feedback, and a one-interval fixed-level task without feedback. Within each task, the interaural configuration and masker bandwidth varied across sessions according to a balanced Latin square, but were held constant within sessions. Each of these training tasks used random noise maskers (i.e., not reproducible noise).

First, the 2I,2AFC task with feedback was used to familiarize the subjects with the listening conditions and to provide an initial estimate of each subject's threshold, which was used to determine the initial tone level for subsequent fixed-level testing. The subject's task was to decide which of two stimulus intervals containing noise also contained a tone. The two-down–one-up tracking procedure estimated the 70.7% correct point on the psychometric function (Levitt, 1971). This procedure used 4-dB steps through the first two reversals and 2-dB steps for the remainder of the run. Ten to 15 runs of the 2I,21AFC task were completed; the exact number of runs depended on the variance of the threshold estimates. Each run consisted of 100 pairs of stimuli in which each interval of the pair had the same masker waveform.

Second, a one-interval, fixed-level task with feedback was employed to familiarize the subject with the task and to determine a signal level for each subject under each condition that would lead to a value of d' near unity, where $d' = z_h - z_f$, and z_h and z_f were the z-scores derived from the overall probability of a hit (P_h) across samples and the overall probability of false alarms (P_f) across samples, respectively (MacMillan and Creelman, 1991). Each run consisted of 100 trials using random noise at the bandwidth being tested. The tone levels in this task were $+3$, $+1$, and -1 dB with respect to the threshold determined by the 2I,21AFC tracking task. Two runs at each tone level were completed for each of the bandwidths and interaural configurations being tested. This sequence was repeated multiple times; levels were adjusted (with 1.0-dB resolution) until performance was stable and d' was approximately unity for the intermediate level tested. Random noise at the bandwidth being tested was used in all training trials that had feedback to prevent subjects from learning the unique characteristics of the reproducible noises.

Finally, the same one-interval task was repeated without feedback to determine if the levels estimated from the psy-

TABLE I. Tone level $(E_S/N_0)^1$ in dB, d', and β are shown for each subject, interaural configuration, and bandwidth (NB: narrowband, 100 Hz geometrically centered around 500 Hz; WB: wideband, 100–3000 Hz). The χ^2 and N^a values are given for performance across reproducible noise samples for both P_h and P_f. The r values are Pearson product-moment correlations for the first half of the trials versus the last half of the trials for each subject and condition. All of the χ^2 values and r values are significant $(p<0.01)$.[b]

Interaural configuration	BW	S	E_s/N_0	d'	β	P_h			P_f		
						χ^2	N	r	χ^2	N	r
N_0S_0	NB	S1	11.8	1.14	1.34	1115.0	96	0.88	907.1	96	0.82
		S2	11.8	0.98	1.14	1355.7	96	0.83	1397.2	96	0.81
		S3	12.8	0.77	1.00	309.0	96	0.60	216.2	96	0.59
		S4	12.8	1.32	0.75	498.6	62	0.71	605.4	62	0.56
	WB	S1	10.8	0.94	1.13	1778.8	96	0.93	932.0	96	0.84
		S2	9.8	1.01	1.09	1811.3	80	0.94	1406.4	80	0.88
		S3	13.8	1.05	1.01	499.3	64	0.66	200.8	64	0.52
		S4	10.8	0.72	0.89	882.8	64	0.84	590.4	64	0.77
N_0S_π	NB	S1	−6.2	0.82	0.86	186.6	64	0.72	212.3	64	0.69
		S2	3.8	1.07	1.12	224.4	64	0.73	357.4	64	0.91
		S3	1.8	0.71	0.97	107.7	88	0.76	93.2	88	0.74
		S4	−6.2	0.96	0.86	278.4	64	0.71	301.5	64	0.68
	WB	S1	−0.2	1.02	1.08	674.2	96	0.93	659.3	96	0.94
		S2	−2.2	1.11	1.02	370.1	56	0.87	552.9	56	0.96
		S3	4.8	0.80	0.94	133.4	96	0.64	164.3	96	0.68
		S4	0.8	0.89	0.90	316.3	64	0.79	342.1	64	0.87

[a] N is the number of trials per reproducible noise sample. The N for the N_0S_0 conditions is a combination of trials from the two tone phases.

[b] Because both of the tone phases are included in the N_0S_0 condition, there are 50 items; there are 25 items in the N_0S_π condition. Therefore, the degrees of freedom in the two conditions are 49 and 24 for N_0S_0 and N_0S_π, respectively, for the χ^2 test; 48 and 23 for N_0S_0 and N_0S_π, respectively, for Pearson's r.

chometric function would still yield a value of d' near unity with feedback turned off. If it did not, the tone level was again adjusted (with 1-dB resolution) to obtain a value of d' that was again near unity.

b. Testing. Four subjects completed a one-interval tone-in-noise detection task for which the subject had to respond either ''yes, the tone was present'' or ''no, the tone was not present.'' The 500-Hz tone was fixed at a level determined by the training tasks described above. Final analyses were conducted on results for a single tone level at which stable performance with a d' near unity was maintained over a complete set of runs (Table I).

The bias parameter β was calculated as a measure of the subject's tendency toward one response using the expression $\beta = e^{-0.5(z_h^2 - z_f^2)}$ (MacMillan and Creelman, 1991). A β value of 1 corresponds to no bias, so that the subject will respond ''tone'' and ''no tone'' equally often. β values greater than 1 indicate that the subject responds ''no tone'' more often; β values less than 1 indicate that the subject responds ''tone'' more often. The experimenter gave the subjects verbal feedback on the bias of their responses if the value of β for a session strayed more than 15% from unity.

Each testing session consisted of four identical sets of trials. Each set began with 20 practice trials with tone stimuli at a level 2 dB above the level that resulted in a d' of unity. Listeners were given feedback after each of the practice trials. Random noise at the bandwidth being tested was used in these practice trials to prevent subjects from learning the unique characteristics of the reproducible noises. Each set then continued with four runs consisting of 100 trials without feedback at the tone level chosen during the preliminary testing. Twenty-five reproducible noise masker waveforms were used in testing tone detection in each condition. Within each run, each noise sample was randomly presented exactly four times, two times with the tone and two times alone, so that each run consisted of 100 trials with no feedback. Each bandwidth and condition was tested for two to three sessions, which resulted in 56 to 96 trials for each signal-plus-noise and each noise-alone sample in each condition. The interaural configurations and noise masker bandwidths were randomized across sessions using a balanced Latin square.

2. Results and discussion

The molar-level results (i.e., averaged across noise samples) are shown in Table I, including E_S/N_0 in dB, d', and β, for the four combinations of interaural configuration (N_0S_0 and N_0S_π) and masker bandwidth (NB: narrowband, 100 Hz bandwidth; WB: wideband, 100–3000 Hz band). The other entries in Table I, χ^2, N, and r-values for hits and false alarms trials, will be discussed later in this work. Subject performance was near the targeted levels ($d'=1$, $\beta=1$) in all cases. Moreover, the performance levels (E_S/N_0 in dB) are comparable to those typically observed in molar-level experiments employing similar stimuli (reviewed in Durlach and Colburn, 1978). Because the d' values were not exactly one, and psychometric functions were not obtained in this study, exact values of the MLD cannot be determined. Nevertheless, approximate MLDs from these results range from about 9 dB to about 12 dB for the wideband maskers, and from about 8 dB to about 19 dB for the narrowband maskers. Most of these values are compatible with those that have typically been observed for similar conditions in the literature (reviewed by Durlach and Colburn, 1978). However, it should be noted that the results for subject 2 under the N_0S_π condition are unusual and indicate substantially worse molar-level performance under the narrowband condition than under the wideband condition (this subject has a relatively small MLD, about 8 dB, for the narrowband condition). Although there is no obvious explanation for this anomaly, considerable variability across subjects has been reported for performance in binaural detection tasks, particularly for narrow bandwidths (Bernstein *et al.*, 1998). In summary, these molar-level results are representative of those that would be expected in an equivalent experiment that did not employ reproducible noise maskers.

A summary of the molecular-level data can be seen in Fig. 1, in which the results for each subject under each binaural presentation mode and each masker bandwidth are shown separately in receiver operating characteristic (ROC) space. The top three rows of plots are for responses to narrowband stimuli, and the bottom three rows are for wideband stimuli; each row represents a particular interaural configuration. Each plotted character shows the proportion of hits (P_h) and the proportion of false alarms (P_f) for a particular noise sample; each character refers to the same noise sample in all panels. As can be seen, the characters are distributed broadly throughout the upper half of ROC space. The χ^2 values shown in Table I indicate that in each panel these across-sample differences in P_h and P_f are significantly greater than would be expected by chance alone and thereby indicate that the subjects' responses were driven by the properties of the individual noise-alone and signal-plus-noise samples. Said differently, some noise-alone and signal-plus-noise samples ''sounded'' more like they contained the target tone than others did. For example, Sample O can be seen in the upper right-hand corner of most of the panels, implying that this sample sounded like it contained the target tone on both noise-alone and signal-plus-noise trials under most conditions. In contrast, some samples appear below the positive slope diagonal in some conditions (e.g., sample A for all subjects in the N_0S_0 wideband condition with 180° tone phase), indicating that the effect of adding the target in these cases was to reduce the probability of a ''yes'' response. Said differently, adding the target made the sample sound less like it contained a target. Gilkey (1981) found that these cases with lower values of P_h than P_f tend to occur when the phase angle of the signal is such that adding the signal to the noise tends to reduce the energy in the noise near 500 Hz. These cases occurred in the present study predominantly in the two wideband N_0S_0 conditions.

The differences in performance on the tone-detection task across the set of reproducible noises were comparable to those reported in previous studies, based on a comparison of ROC plots (Gilkey *et al.*, 1985; Isabelle and Colburn, 1991) and χ^2 values (Isabelle and Colburn, 1991). Greater differences in responses across the set of reproducible noises for N_0S_0 than for N_0S_π can be observed in Fig. 1, and are also reflected in the χ^2 values in Table I (note that the numbers of trials for each condition must be taken into account when comparing χ^2 values). These greater differences in responses across noise samples for the N_0S_0 condition were also reported by Isabelle and Colburn (1991) and are consistent with the greater dependence of detection threshold on target phase for the N_0S_0 condition, which has been reported in previous studies of detection in reproducible noises (Gilkey *et al.*, 1985; Langhans and Kohlrausch, 1992). Examining these data at the quasi-molecular level indicates statistically significant sample-by-sample differences in subject responses that are not, by definition, considered in a molar level analysis. The goal of this series of articles is to utilize these sample-by-sample differences to determine the processing that the observer uses to judge the presence or absence of the target.

a. Comparison of responses across bandwidths. If the subjects base their judgments only on information within the auditory filter centered at the 500-Hz target frequency, then the effective stimuli under the wideband and narrowband conditions are identical. If so, the patterns of responses seen in the panels in the upper half of Fig. 1 should be identical to the corresponding patterns in the lower half of Fig. 1.

To examine this prediction more closely, the correlation between responses under the narrowband and wideband conditions is shown in Table II, separately for P_h and P_f, and for each binaural presentation mode and subject. For the N_0S_0 condition, the correlations for all of the subjects are significant for P_h; three of the four subjects are significantly correlated for P_f and the fourth subject shows positive, but insignificant, correlation. These results indicate that subjects are, to some degree, using the same information in the wideband and narrowband maskers (e.g., the information contained in the critical band centered at the target frequency) to make their judgments about the presence of the target in the N_0S_0 condition. However, a measure of the strength of these correlations in the context of the stability of subjects' performance can be obtained by comparison to the intra-subject correlations in Table I. The observed correlations of responses for the two bandwidths, while significant, are lower than might be expected based on the correlation between each subject's responses during the first half and second half of the runs (Table I), which was significantly greater than that across the two bandwidths. Tests of significant differences for non-independent correlations were used to compare the across-bandwidth correlation to the first-half–last-half correlations for N_0S_0 hits. Fourteen of the 16 resultant comparisons (2 bandwidths×2 halves×4 subjects) were significant at the 0.05 level. The significantly decreased correlation across bandwidths relative to the first-half–last-half correlation implies that information outside the 100-Hz band centered at the signal frequency affects the subject responses in the wideband condition. Consistent with this result is Gilkey and Robinson's (1986) ability to explain more of the sample dependence with a model that combined seven 50-Hz bands over a range of frequencies than they could with a single frequency band.

For the N_0S_π condition, none of the subjects show values of P_h that were significantly correlated across the bandwidths (Table II). For two subjects, P_f was significantly cor-

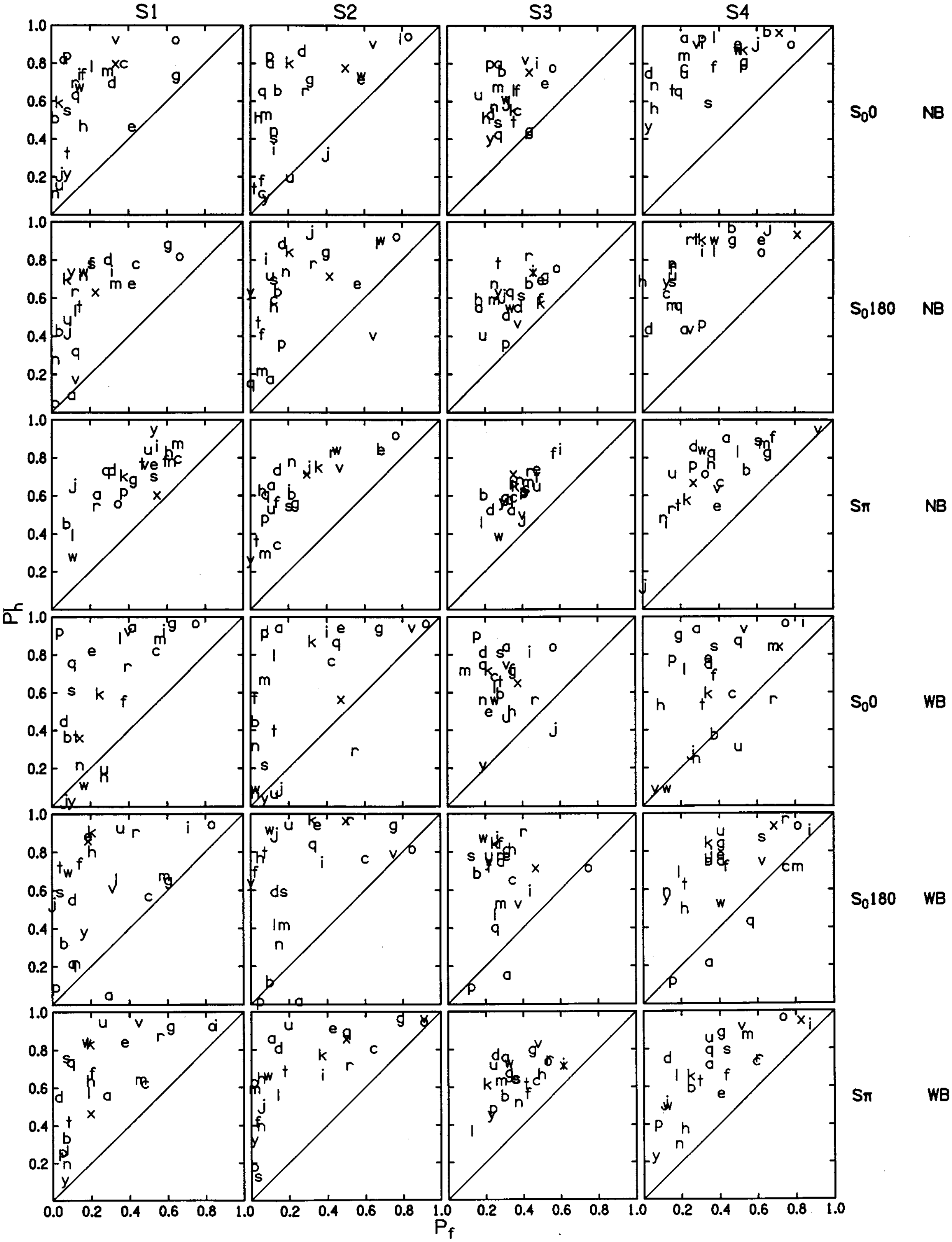

FIG. 1. P_h vs P_f for each reproducible noise sample ("a" through "y") for each subject (columns), bandwidth (top three rows versus bottom three rows), and interaural configuration (rows).

related across bandwidths, but these correlations were significantly lower than the comparable correlations between each subject's responses for the first and second halves of the runs (Table I). The weak or insignificant correlations across the two bandwidths implies that, for all subjects, N_0S_π responses in the wideband condition are influenced by information outside the critical band centered at the signal frequency, thereby lowering the correlation with responses in

TABLE II. Correlations between results for the wideband and the narrowband conditions. Pearson product-moment correlations are given for each subject and interaural configuration for both P_h and P_f. df=48 (N_0S_0), 23 (N_0S_π) (see note in caption for Table I).

Interaural configuration	Subject	r (P_h)	r (P_f)
N_0S_0	S1	0.72[a]	0.66[a]
	S2	0.51[a]	0.51[a]
	S3	0.43[a]	0.47[a]
	S4	0.38[a]	0.22
N_0S_π	S1	−0.06	0.20
	S2	0.26	0.48[b]
	S3	−0.03	0.47[b]
	S4	0.05	−0.08

[a] $p<0.01$.
[b] $p<0.05$.

TABLE III. Correlations between results for the two interaural configurations, N_0S_0 and N_0S_π. Pearson product-moment correlations are given for each subject and bandwidth. Note: Because the two tone phases were averaged for the N_0S_0 condition in the comparison to N_0S_π, all correlations have 23 degrees of freedom.

Bandwidth	Subject	r (P_h)	r (P_f)
NB	S1	−0.15	0.23
	S2	0.80[a]	0.92[a]
	S3	0.38	0.25
	S4	−0.50[b]	−0.17
WB	S1	0.77[a]	0.94[a]
	S2	0.75[a]	0.98[a]
	S3	0.43[b]	0.71[a]
	S4	0.86[a]	0.91[a]

[a] $p<0.01$.
[b] $p<0.05$.

the narrowband condition. Moreover, the effect of frequencies outside the critical band centered at the signal frequency appears to be substantially greater in the N_0S_π condition than in the N_0S_0 condition.

These results are compatible with the unpublished results presented by Gilkey (1990), who found significant correlations between values of P_f in wideband and narrowband N_0S_0 conditions, but substantially smaller (although still statistically significant for two out of three subjects) correlations between values of P_f in the wideband and narrowband N_0S_π conditions. The weak relation between wideband and narrowband responses in the N_0S_π condition is also consistent with subjective reports about the cues used in the N_0S_π condition for narrowband and wideband stimuli. Specifically, in the narrowband case, the ''width'' or ''shape'' of the binaural image is generally reported to provide a cue for detection; in the wideband case, the strength of the tonelike percept is reported to provide a cue for detection. These results are also consistent with the conclusions of van de Par and Kohlrausch (1999) and Breebaart *et al.* (2001) using random (nonreproducible) noise maskers across a range of bandwidths. However, they concluded that subjects were performing the task based on a single auditory filter centered at the target frequency in the wideband case, whereas integration across a number of different auditory filters was used in the narrowband case. The reproducible noise results presented here provide a means to test specific predictions of this and other models in future modeling studies.

c. Comparison of Responses across Interaural Configurations. The potential difference between results for wideband and narrowband conditions was first indicated when values of P_f and P_h were compared across interaural configurations for sets of reproducible narrowband maskers by Isabelle and Colburn (1991), who found weak and typically insignificant correlations, and for wideband maskers by Gilkey *et al.* (1985), who found significant correlations. However, these results were for different subjects and different reproducible noise samples. The current study allows this comparison within a single experiment. Based on the previous reports, it was expected that performance for N_0S_0 and N_0S_π conditions would be correlated under the wideband conditions and uncorrelated under the narrowband conditions. Table III shows that the responses of all four subjects were significantly correlated across interaural configurations with the wideband maskers, which is consistent with Gilkey *et al.* (1985). With the narrowband maskers, the responses of two subjects (S1, S3) were not significantly correlated for either signal-plus-noise trials (i.e., P_h) or noise-alone trials (i.e., P_f), the responses of one subject (S2) were positively correlated for both noise-alone and signal-plus-noise trials, and the responses of one subject (S4) were negatively correlated (but only for signal-plus-noise trials). These results are consistent with the diversity in response patterns across subjects reported by Isabelle and Colburn (1991) for narrowband maskers. (They reported one subject with significant positive correlations for both hits and false alarms, and two subjects with negative, but not significant, correlations.) These results suggest either that different processing strategies are used for different bandwidths or that masker components outside the critical band have a significant impact on the processing of stimulus components within the critical band.

d. Comparison of performance across subjects. The comparisons across bandwidths and interaural configurations presented above illustrate trends that were generally true for all subjects. For example, the P_h results for the N_0S_0 condition were strongly correlated across bandwidths for all subjects, and the P_h results for the N_0S_π condition were not correlated for any of the subjects

Intersubject correlations are shown in Table IV. The responses across reproducible noise maskers were significantly correlated for all pairs of subjects for the N_0S_0 condition for all cases, including both hits and false alarms and both narrowband and wideband maskers. N_0S_π results were also significantly correlated for all pairs of subjects for the wideband maskers, for both hits and false alarms. However, for the narrowband N_0S_π condition, only two subjects (S1 and S3) had significant positively correlated performance; one pair of subjects had a significantly negative correlation for performance on hits, and all other correlations were insignificant for the narrowband N_0S_π condition. The degree of variability in performance across subjects in this study was consistent with that reported for similar tasks by Bernstein *et al.* (1998).

TABLE IV. Correlations of responses between subjects. Pearson product-moment correlations are given for each subject pair under each interaural configuration and bandwidth. df=48(N_0S_0), 23 (N_0S_π) (see note in caption for Table I).

Interaural configuration	Bandwidth	Subject pair	r (P_h)	r (P_f)
N_0S_0	NB	S1-S2	0.54^{a}	0.44^{a}
		S1-S3	0.45^{a}	0.56^{a}
		S1-S4	0.45^{a}	0.31^{b}
		S2-S3	0.35^{b}	0.50^{a}
		S2-S4	0.55^{a}	0.60^{a}
		S3-S4	0.69^{a}	0.61^{a}
	WB	S1-S2	0.81^{a}	0.63^{a}
		S1-S3	0.66^{a}	0.55^{a}
		S1-S4	0.82^{a}	0.56^{a}
		S2-S3	0.69^{a}	0.58^{a}
		S2-S4	0.73^{a}	0.58^{a}
		S3-S4	0.71^{a}	0.47^{a}
N_0S_π	NB	S1-S2	-0.59^{a}	-0.34
		S1-S3	0.50^{b}	0.55^{a}
		S1-S4	-0.05	0.19
		S2-S3	-0.19	-0.09
		S2-S4	-0.32	-0.16
		S3-S4	-0.18	-0.25
	WB	S1-S2	0.56^{a}	0.69^{a}
		S1-S3	0.63^{a}	0.60^{a}
		S1-S4	0.68^{a}	0.77^{a}
		S2-S3	0.62^{a}	0.54^{a}
		S2-S4	0.51^{a}	0.61^{a}
		S3-S4	0.62^{a}	0.78^{a}

[a] $p < 0.01$.
[b] $p < 0.05$.

C. Experiment 2: Comparison of tracking and fixed-level procedures

1. Methods

Experiment 1 was conducted at a fixed-tone level; however, a companion study (Zheng *et al.*, 2002) used rabbits that were studied with a tracking procedure. To explore potential differences in performance that might be related to these different testing procedures, a tracking procedure was used to retest two of the subjects from experiment 1 (S1 and S4), and the results of the two procedures were compared.

The same one-interval, yes–no task that was used in the fixed-level procedure of experiment 1 was used for the tracking procedure here, and the same fixed number of trials were included in each run. However, the tone-level in each tone-plus-noise trial was adjusted by following a two-down–one-up rule (Levitt, 1971). Tone levels were adjusted based on the subject's responses for tone-plus-noise trials only. For each track, 4-dB steps were used until there were two reversals and then 2-dB steps were used for the remainder of the run. The 70.7% correct detection threshold was calculated by averaging the reversals (after the step-size change) of each track. Signal trials that were presented at levels between one step above and two steps below the mean reversal level were used for the reproducible noise analysis, consistent with the analysis used in the companion study (Zheng *et al.*, 2002). Subject S1 was tested on narrowband conditions and subject S4 was tested on wideband conditions.

2. Results and discussion

Table V shows the summary of the data collected from these two subjects, which can be compared to their results in Table I for Experiment 1. With the tracking procedure, subject 1 shows a MLD of 16 dB for the narrowband condition and subject 4 has a MLD of 12 dB for the wideband condition. As can be seen, both molar-level (threshold values of E_S/N_0 in dB) and quasi-molecular-level (χ^2 and correlations between the first and last halves of the trials) results are comparable for the two experiments. Subject 1's threshold for the N_0S_π condition was higher by approximately 3 dB whereas subject 4's threshold was lower by 2 dB when performing the fixed-level experiment. Table VI shows that the sample-by-sample correlations between experiments 1 and 2 for both subjects were significant and positive for both hits and false alarms for all four combinations of bandwidth and interaural configuration. The results from the two testing procedures were strongly correlated for both subjects and for all conditions tested. Although these subjects were all tested extensively using fixed-level procedures before testing with the tracking procedures, it appears that tracking and fixed-level procedures yield similar results and that the results of our planned across-species comparisons would not be substantially obscured by this difference in procedure.

TABLE V. Experiment 2 results. Tone level (E_s/N_0) in dB is shown for each subject and condition. The values of χ^2 and N^{a} are given for performance across reproducible noise samples for both P_h and P_f. The r values are Pearson product-moment correlations for the first half of the trials versus the last half of the trials for each subject and condition. All of the χ^2 values and r values are significant ($p < 0.01$).[b]

Interaural configuration	BW	S	E_s/N_0	P_h			P_f		
				χ^2	N	r	χ^2	N	r
N_0S_0	NB	S1	12.8	302.1	88	0.62	370.6	160	0.71
N_0S_π			-3.2	291.6	94	0.91	204.3	160	0.74
N_0S_0	WB	S4	10.8	493.3	70	0.74	448.1	120	0.87
N_0S_π			-1.2	268.6	98	0.84	730.3	160	0.97

[a] N is the number of trials per reproducible noise sample. The N for the N_0S_0 conditions is a combination of trials from the two tone phases.
[b] Because both tone phases are included in the N_0S_0 condition, there are 50 items; there are 25 items in the N_0S_π condition. Therefore, the degrees of freedom in the two conditions are 49 and 24 for N_0S_0 and N_0S_π, respectively, for the χ^2 test; 48 and 23 for N_0S_0 and N_0S_π, respectively, for Pearson's r.

TABLE VI. Comparison of performance for the fixed-level versus tracking results. Pearson product-moment correlations between results for the two paradigms are shown.

Bandwidth	Interaural configuration	Subject	r (P_h)	r (P_f)
NB	N_0S_0	S1	0.79[a]	0.76[a]
	N_0S_π		0.85[a]	0.84[a]
WB	N_0S_0	S4	0.84[a]	0.77[a]
	N_0S_π		0.86[a]	0.91[a]

[a] $p < 0.01$.

III. GENERAL DISCUSSION

This study tested subjects using a binaural detection task with wideband and narrowband noise maskers that had the same spectral components in the 100-Hz frequency region surrounding the 500-Hz tone. Comparisons of the results for the two bandwidth conditions reported here indicate that frequencies outside the 100-Hz band centered at the 500-Hz tone influence detection in both the N_0S_0 and N_0S_π conditions. The results are consistent with previous studies that focused on either wideband or narrowband reproducible noise maskers (e.g., Gilkey *et al.*, 1985; Isabelle and Colburn, 1991).

Comparison of the responses across the N_0S_0 and N_0S_π conditions for each of the masker bandwidths suggests that diotic and dichotic responses differ significantly due to the influence of frequencies outside a bandwidth that approximates the critical band. Dichotic processing is apparently much more influenced by the presence of frequencies away from the target frequency. These results provide motivation to extend models beyond the narrowband mechanisms that have been the focus of binaural detection models to date (Isabelle, 1995; Colburn *et al.*, 1997; cf. Breebaart *et al.*, 2001; Trieurniet and Boucher, 2001).

Future studies in this series will attempt to model these experimental results. Several challenges for such modeling studies are raised by these results. For example, responses across subjects generally were not correlated for the narrowband N_0S_π condition; therefore, it would not be possible to explain these data with a single model except by changing parameters from subject to subject. In general, performance across subjects was more highly correlated for the N_0S_0 condition than for the N_0S_π condition and was more correlated for the wideband condition than for the narrowband condition. The results of these comparisons suggest that subjects' listening strategies may change in a complex manner that is influenced by energy outside the critical band. In addition, different strategies for combining information across frequencies have been suggested by this and other studies. Whereas we conclude that energy outside the critical band influences detection in the wideband condition, others have concluded that cross-filter integration predominantly affects the narrowband condition (e.g., Breebaart *et al.*, 2001). These differences can be explored both by detailed modeling of the results across reproducible noise maskers, and with additional experimental studies in which the spectral contents are systematically varied both within and outside the critical band.

It has been established in detection studies with rabbits that differences in detection performance are observed across noise samples (Early *et al.*, 2001). A study of tone detection in rabbit using narrowband and wideband maskers and using the interaural configurations of the current study is the topic of the companion article (Zheng *et al.*, 2002). Similar trends across bandwidths and interaural configurations were found for the rabbits as were found for the human subjects in the current study. Future studies will pursue the problem of diotic and dichotic masking with signal-processing- and physiologically based models, and with physiological experiments.

ACKNOWLEDGMENTS

We acknowledge the helpful discussions and comments provided on earlier versions of this article by Susan J. Early, J. Michael Harrison, Armin Kohlrausch, and Fabio Idrobo. This work was supported by NIH NIDCD DC01641 (MEE, CRM, LHC) and NIH NIDCD DC00100 (HSC, RHG). Additional support was provided by AFOSR F49620-97-1-0231 (RHG).

---

[1] The level of the signal with respect to the masker was calculated as E_S/N_0=Tone level (dB SPL)−Noise spectrum level (dB SPL) $+10\ \log_{10}$ duration (s)/1 s.

Ahumada, A., and Lovell, J. (**1971**). "Stimulus features in signal detection," J. Acoust. Soc. Am. **49**, 1751–1756.

Ahumada, A., Marken, R., and Sandusky, A. (**1975**). "Time and frequency analysis of auditory signal detection," J. Acoust. Soc. Am. **57**, 385–390.

Bernstein, L. R., Trahiotis, C., and Hyde, E. L. (**1998**). "Inter-individual differences in binaural detection of low-frequency tonal signals masked by narrow-band or broadband noise," J. Acoust. Soc. Am. **103**, 2069–2078.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001**). "An explanation for the apparently wider critical bandwidth in binaural experiments," in *Proceedings of the 12th International Symposium on Hearing: Physiological and Psychophysical Bases of Auditory Function*, edited by D. J. Breebaart, A. J. M. Houstma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht, The Netherlands), pp. 153–160.

Colburn, H. S., Isabelle, S. K., and Tollin, D. J. (**1997**). "Modeling binaural detection performance for individual masker waveforms," in *Binaural and Spatial Hearing*, edited by R. H. Gilkey and T. Anderson (Earlbum, Englewood Cliffs, NJ), Chap. 25, pp. 533–555.

Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. **35**, 1206–1218.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena," In *Handbook of Perception: Hearing*, edited by E. Carterette and M. Friedman (Academic, New York), Vol. 4, pp. 365–466.

Early, S. J., Mason, C. R., Zheng, L., Evilsizer, M., Idrobo, F., Harrison, J. M., and Carney, L. H. (**2001**). "Studies of binaural detection in the rabbit using Pavlovian conditioning," Behav. Neurosci. **115**, 650–660.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47–65.

Gilkey, R. H. (**1981**). "Molecular psychophysics and models of auditory signal detectability," Ph.D. Thesis, Department of Psychology, Indiana University.

Gilkey, R. H. (**1987**). "Spectral and temporal comparisons in auditory masking," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 26–36.

Gilkey, R. H. (**1990**). "The relation between monaural and binaural tone-in-noise masking," Assoc. for Res. in Otolaryngol. (abstract) **13**, 165.

Gilkey, R. H. and Robinson, D. E. (**1986**). "Models of auditory masking: A molecular psychophysical approach," J. Acoust. Soc. Am. **79**, 1499–1510.

Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (**1985**). "Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise," J. Acoust. Soc. Am. **78**, 1207–1219.

Green, D. M. (**1964**). "Consistency of auditory detection judgments," Psychol. Rev. **71**, 392–407.

Helmholtz, H. L. F. von (**1863**). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Vieweg und Sohn, Braunschweig, Germany). Translated as *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, by A. J. Ellis from the 4th German edition, 1877 (Leymans, London, 1885) (reprinted by Dover, New York, 1954).

Isabelle, S. K. (**1995**). ‘‘Binaural detection performance using reproducible stimuli,’’ Ph.D. thesis, Boston University.

Isabelle, S. K. and Colburn, H. S. (**1991**). ‘‘Detection of tones in reproducible narrow-band noise,’’ J. Acoust. Soc. Am. **89**, 352–359.

Jiang, D., McAlpine, D., and Palmer, A. R. (**1997a**). ‘‘Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions,’’ J. Neurophysiol. **77**, 3085–3106.

Jiang, D., McAlpine, D., and Palmer, A. R. (**1997b**). ‘‘Detectability index measures of binaural masking level difference across populations of inferior colliculus neurons,’’ J. Neurosci. **17**, 9331–9339.

Kidd, G., Jr. (**1987**). ‘‘Auditory discrimination of complex sounds: The effects of amplitude perturbation on spectral shape,’’ in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 16–25.

Kidd, Jr., G., Mason, C. R., Brantley, M. A., and Owen, G. A. (**1989**). ‘‘Roving-level tone-in-noise detection,’’ J. Acoust. Soc. Am. **86**, 1310–1317.

Langhans, A., and Kohlrausch, A. (**1992**). ‘‘Differences in auditory performance between monaural and diotic conditions. I: Masked thresholds in frozen noise,’’ J. Acoust. Soc. Am. **91**, 3456–3470.

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

MacMillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User’s Guide* (Cambridge U.P., New York).

Miller, M. I., Barta, P. E., and Sachs, M. B. (**1987**). ‘‘Strategies for the representation of a tone in background noise in the temporal aspects of the discharge patterns of auditory-nerve fibers,’’ J. Acoust. Soc. Am. **81**, 665–679.

Palmer, A. R., Jiang, D., and McAlpine, D. (**1999**). ‘‘Desynchronizing responses to correlated noise: A mechanism for binaural masking level differences at the inferior colliculus,’’ J. Neurophysiol. **81**, 722–734.

Palmer, A. R., Jiang, D., and McAlpine, D. (**2000**). ‘‘Neural responses in the inferior colliculus to binaural masking level differences created by inverting the noise in one ear,’’ J. Neurophysiol. **84**, 844–852.

Pfafflin, S. M., and Mathews, M. V. (**1966**). ‘‘Detection of auditory signals in reproducible noise,’’ J. Acoust. Soc. Am. **39**, 340–345.

Rees, A., and Palmer, A. R. (**1988**). ‘‘Rate-intensity functions and their modifications by broadband noise for neurons in the guinea pig inferior colliculus,’’ J. Acoust. Soc. Am. **83**, 1488–1498.

Richards, V. M. (**1992**). ‘‘The detectability of a tone added to narrow bands of equal-energy noise,’’ J. Acoust. Soc. Am. **91**, 3424–3435.

Richards, V. M., and Nekrich, R. D. (**1993**). ‘‘The incorporation of level and level-invariant cues for the detection of a tone added to noise,’’ J. Acoust. Soc. Am. **94**, 2560–2574.

Richards, V. M., Heller, L. M., and Green, D. M. (**1991**). ‘‘Detection of a tone in a narrow band of noise: The energy model revisited,’’ Q. J. Psychol. **43A**, 481–503.

Siegel, R. A., and Colburn, H. S. (**1983**). ‘‘Internal and external noise in binaural detection,’’ Hear. Res. **11**, 117–123.

Siegel, R. A., and Colburn, H. S. (**1989**). ‘‘Binaural processing of noisy stimuli: Internal/external noise ratios for diotic and dichotic stimuli,’’ J. Acoust. Soc. Am. **86**, 2122–2128.

Trieurniet, W. C., and Boucher, D. R. (**2001**). ‘‘A masking level difference due to harmonicity,’’ J. Acoust. Soc. Am. **109**, 306–320.

van de Par, S., and Kohlrausch, A. (**1999**). ‘‘Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters,’’ J. Acoust. Soc. Am. **106**, 1940–1947.

Watson, C. S. (**1962**). ‘‘Signal detection and certain physical characteristics of the stimulus during the observation interval,’’ Ph.D. thesis, Indiana University, Bloomington, IN.

Young, E. D., and Barta, P. E. (**1986**). ‘‘Rate responses of auditory-nerve fibers to tone in noise near masked threshold,’’ J. Acoust. Soc. Am. **79**, 426–442.

Zheng, L., Early, S. J., Mason, C. R., Evilsizer, M. E., Idrobo, F., Harrison, J. M., and Carney, L. H. (**2002**). ‘‘Binaural detection with narrowband and wideband reproducible noise maskers: II. Results for rabbit,’’ J. Acoust. Soc. Am. **111**, 346–356.

Zwicker, E., and Henning, G. B. (**1984**). ‘‘Binaural masking-level differences with tones masked by noises of various bandwidths and levels,’’ Hear. Res. **14**, 179–183.