

Monaural level discrimination under dichotic conditions^{a)}

Daniel E. Shub^{b)}

Speech and Hearing Bioscience and Technology Program, Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Hearing Research Center, Biomedical Engineering Department, Boston University, Boston, Massachusetts 02215

Nathaniel I. Durlach

Hearing Research Center, Biomedical Engineering Department, Boston University, Boston, Massachusetts 02215 and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

H. Steven Colburn

Hearing Research Center, Biomedical Engineering Department, Boston University, Boston, Massachusetts 02215 and Speech and Hearing Bioscience and Technology Program, Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 25 May 2006; revised 19 March 2008; accepted 2 April 2008)

The ability to make judgments about the stimulus at one ear when a stimulus is simultaneously presented to the other ear was tested. Specifically, subjects discriminated the level of a 600 Hz target tone presented at the left ear while an identical-frequency distractor was simultaneously presented at the other ear. When there was no distractor, threshold was 0.7 dB. Threshold increased to 1.1 dB when a distractor with a fixed phase and level was introduced contra-aurally to the target. Further increases in threshold were observed when an across-presentation variability was introduced into the distractor phase (threshold of 1.6 dB) or level (threshold of 5.8 dB). When both the distractor level and phase varied, the largest threshold of 7.3 dB was obtained. These increases in threshold cannot be predicted by common binaural models, which assume that a target stimulus at one ear can be processed without interference from the stimulus at the nontarget ear. The measured thresholds are consistent with a model that utilizes two binaural dimensions that roughly correspond to the loudness and the position of a fused binaural image. The results show that, with binaurally fused tonal stimuli, subjects are unable to listen to one ear. © 2008 Acoustical Society of America.

[DOI: 10.1121/1.2912828]

PACS number(s): 43.66.Pn, 43.66.Ba, 43.66.Fe, 43.66.Rq [AK]

Pages: 4421–4433

I. INTRODUCTION

Most models of binaural processing assume that one can voluntarily listen to the signal at one ear (right or left) even when the acoustic stimulus is binaural. More specifically, in the most frequently used models of binaural hearing [see the review by Colburn and Durlach (1978)], it is assumed that the input from each ear bifurcates into two pathways. One of these pathways proceeds to a network in which binaural interaction occurs and the other proceeds up the auditory system independent of the stimulus to the opposite ear (usually referred to as a monaural channel). Furthermore, these models assume that the listener is able to combine the information from these channels (e.g., switch between monaural and binaural listening). One consequence of these assumptions is that it should be impossible to degrade the performance of a task, where correct responses are completely determined by the stimulus at a single ear, by introducing signals into the

other ear. In other words, there should be no “cross masking,” or “contralateral masking,” or “binaural disadvantage,” or, using the term we prefer, “contra-aural interference.”

In this study, the ability to access the monaural channel is measured by having subjects discriminate the level of a target tone presented at one ear while an identical-frequency distractor tone is simultaneously presented at the other ear. The objective task is based on the level of the stimulus at a single ear even though the natural perception of a binaural tone is a fused percept with a prominent loudness and position that are influenced by the stimuli at both ears. In addition to this dominant perception, however, there are also “secondary” percepts such as the image width or additional images (e.g., Hafter and Jeffress, 1968; Ruotolo *et al.*, 1979; Hartman and Constan, 2002). Even if there is no percept that corresponds to a monaural channel, the complexity of the binaural percepts makes modeling of the auditory system as having both monaural and binaural channels seem reasonable since the binaural percepts could provide information equivalent to that carried by the postulated monaural channels. Although data that are consistent with an accessible monaural channel (e.g., almost all of the data on the masking of tones by noise) certainly exist, it is by no means always the case.

^{a)}Parts of this work were presented at the 27th Midwinter Meeting of the Association of Research in Otolaryngology [Shub, D. E., and Colburn, H. S. (2004). “Monaural Intensity Discrimination Under Dichotic Conditions,” *Assoc. Res. Otolaryngol. Abstr.* 1521].

^{b)}Present address: Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104. Electronic mail: dshub@sas.upenn.edu

There have been experiments of several types that have illustrated contra-aural interference, most of which involved broadband stimuli. For example, in studies of informational masking (e.g., Brungart and Simpson, 2002; Kidd *et al.*, 2003), monaural speech intelligibility was reduced by adding a speech masker to the other ear. In a precedence-effect experiment, Zurek (1979) demonstrated that detection of a lagging sound could be better under diotic conditions than under dichotic conditions; this experiment, together with the usual observation that diotic performance is the same as monotic performance, suggests that the dichotic case is generating contra-aural interference. In discrimination experiments, Bernstein and Oxenham (2003) showed that the ability to discriminate changes in the fundamental frequency of a harmonic tone complex was reduced under dichotic conditions, and Heller and Trahiotis (1995) reported that subjects could discriminate different noise tokens under monotic conditions, but that the subjects were unable to discriminate the tokens under some dichotic conditions. These studies interpreted the measured (contra-aural) interference by using the concepts of the central spectrum (Bilsen, 1977; Bilsen and Raatgever, 2000) and nonoptimal across-frequency processing.

Other cases of contra-aural interference cannot be easily interpreted with the concepts of the central spectrum and nonoptimal across-frequency processing. In particular, contra-aural interference has been reported in both masked (Taylor and Clarke, 1971; Taylor *et al.*, 1971a; Taylor *et al.*, 1971b; Yost *et al.*, 1972; Koehnke and Besing, 1992) and absolute (Zwislocki, 1972; Mills *et al.*, 1996) detection experiments. Small amounts of contra-aural interference have also been demonstrated in experiments in which subjects discriminate the level of a target tone, which is presented at one ear, in the presence of a simultaneously presented contra-aural distractor tone, with the same frequency and duration as the target (Rowland and Tobias, 1967; Yost, 1972; Bernstein, 2004). These level discrimination experiments were primarily investigations of binaural, specifically interaural level difference (ILD), processing. For some of these binaural conditions, when one headphone was removed (as was done in the monotic control conditions included in these studies), performance improved; it is these conditions that lead us to describe these experiments as showing contra-aural interference.

Although inconsistent with reports of contra-aural interference, the conceptualization and modeling of the auditory system including both monaural and binaural processing channels persist. One rationale for the inconsistency, which was suggested by Durlach and Colburn (1978), is that the “monaural channels” are difficult to access in some circumstances, particularly when the percepts are complicated. In these cases, one might expect that the paradigm, training, feedback, and instructions are important. In fact, the results of experiments in which secondary binaural cues are important can be heavily dependent on the experimental procedure (Hafer and Jeffress, 1968; Hafer and Carrier, 1970; Trahiotis, 1992). Therefore, tasks that use these secondary binaural cues to access the output of monaural processing channels might also be dependent on the experimental procedure.

Here, a number of steps were taken in an attempt to optimize the subjects’ abilities to access possible monaural channel listening when discriminating the level of a target tone that was presented at one ear. The task was objectively defined with correct-answer feedback based on the level of the target. Subjects were instructed to optimize performance by using the correct-answer feedback to prevent the introduction of potential biases that might arise from the experimenter’s description of useful percepts. A four-interval (two-cue), two-alternative forced-choice experimental paradigm was selected as it is particularly easy for subjects in complicated situations (Trahiotis, 1992). An adaptive paradigm, which is initiated with a difference between the two target levels that was large enough for consistently good performance, was used so that subjects would be able to experience a variety of possible subjective cues. Finally, performance was measured under a number of different dichotic conditions, with the monotic (monaural) condition as a reference, in which potential cues were systematically eliminated.

In a series of stimulus conditions specified at the nontarget ear, we systematically manipulated the amount of variability in the distractor level and phase to eliminate the reliability of the dominant binaural perceptions (loudness and position) for judgments about the level of the target. The most extreme condition was chosen, such that a model based on loudness and lateral position would predict substantial contra-aural interference, and thus, if little contra-aural interference was measured, this would provide evidence for the “monaural processing channels” described above and assumed in most binaural models. A number of intermediate conditions were also evaluated, such that if there was substantial contra-aural interference, one could determine whether performance is consistent with optimal use of only the outputs of the binaural processing channels. In the modeling portion of this work, we present predictions based on a model in which the decision device only has access to binaural processing channels that provide estimates of the lateral position and overall loudness.

The model is a two-dimensional decision-theoretic model with decision variables based on the overall loudness and lateral position. This model, like the experiment reported here, does not include variations over the frequency or the time dimensions of the stimulus; it is limited to a fused, stationary stimulus. When the distractor level and phase are fixed from interval to interval, both the loudness and lateral position are reliable cues for discriminating the level of the target. When the phase of the distractor is random from interval to interval, the reliability of the lateral position for discriminating the target is greatly reduced. When the level of the distractor is randomized, the reliabilities of both the loudness and the lateral position are reduced; however, the loudness and lateral position together still specify the target level. When the level and phase of the distractor are independently randomized, neither the loudness nor the lateral position nor their combination is adequate to specify the target level.

The rest of this document is organized as follows. Section II presents the experimental methods and explains the data analysis procedures. In Sec. III, models based on loud-

TABLE I. Distractor properties in the five conditions. In all of the conditions, the target has a frequency of 600 Hz, a duration of 300 ms, a phase of zero, and a reference level of 50 dB SPL. The distractor was simultaneously presented but contra-aurally to the target. Roving of the level and phase of the distractor was done on an interval-by-interval basis with values chosen from uniform distributions.

	Frequency (Hz)	Duration (ms)	Phase (rad)	Level (dB SPL)
No distractor
Fixed	600	300	0	50
Roving phase	600	300	Uniform $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$	50
Roving level	600	300	0	Uniform (50,80)
Double rove	600	300	Uniform $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$	Uniform (50,80)

ness and position decision variables are described and related to similar models in the literature. In Sec. IV, the psycho-physical results are presented, along with the predictions of the models. The discussion in Sec. V includes comparisons between data and model predictions and explores the possibility that monaural processing channels exist. Finally, Sec. VI gives some concluding remarks.

II. EXPERIMENTAL METHODS

A. Subjects

Four subjects (S1–S4) completed the tasks. Subject S1 is the first author. With the exception of S1, subjects received an hourly wage for their participation. All subjects had pure tone thresholds below 20 dB HL at frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz in both ears. The subjects were between 19 and 31 years old. Subjects S1 and S2 had prior listening experience in similar tasks, while subjects S3 and S4 had no prior experience in psychoacoustic experiments.

B. Stimulus and procedures

The experimental task was designed to test the ability of subjects to discriminate the level of a 600 Hz target tone at the left ear in the presence of a 600 Hz distractor tone simultaneously presented at the right ear. A four-interval, two-alternative, forced-choice (4I-2AFC), adaptive paradigm was used and correct-answer feedback was given on every trial. Five different conditions were explored: the *no-distractor* condition and four conditions that differed in the presence/absence of interval-to-interval variation in the distractor phase and level. The four distractor conditions were *fixed* distractor, *roving-phase* distractor, *roving-level* distractor, and *double-rove* distractor (with both level and phase variation). Table I lists the properties of the distractor used in each condition.

The target and distractor tones had 300 ms durations and 25 ms rise/fall times, were simultaneously gated on and off, and had 500 ms of quiet between intervals. The target was presented at either the reference level of 50 dB SPL or the

reference level plus an increment ΔL (in decibels). Tone phases were defined relative to the onset of the stimulus ramp; the target phase was always zero and the distractor phase was specified relative to this zero-phase target. When the distractor level was fixed, it was held at 50 dB SPL (the reference level of the target). When the level was roved (roving-level and double-rove conditions), the level of the distractor was randomly chosen on an interval-by-interval basis from a uniform distribution between 50 and 80 dB SPL. When the distractor phase was fixed, it was held at zero phase, and when it was roved (roving-phase and double-rove conditions), the phase of the distractor uniformly roved between $\pm \pi/2$ on an interval-by-interval basis.

In the 4I-2AFC paradigm used here (Bernstein and Trahiotis, 1982), listeners must distinguish the patterns *ABAA* from *AABA*, where *A* and *B* are defined by the level of the tone at the left ear (i.e., the target). Stimuli labeled *A* have the reference level at the left ear and stimuli labeled *B* have the level at the left ear incremented by ΔL . Subjects were informed about the two temporal patterns and instructed to maximize the percent correct by utilizing the trial-by-trial feedback. We did not explicitly tell the subjects to respond according to whether the level of the tone at the left ear was relatively higher on the second or third interval (even though such instructions would have been consistent with the objective task and with the feedback provided) because we thought it was better to avoid restricting the subjects' attention to specific subjective cues. In our pilot experiments, when subjects tried to focus their attention directly on the loudness at the left ear rather than on the loudness and lateralization of the fused image (and on how these subjective cues related to the feedback), they performed much worse. When considering the instructions used, it is also worth noting that subjects who knew that the key to the feedback was the level at the left ear (e.g., the first author who served as subject S1 as well as subject S2) and subjects who knew only that they should respond in a manner that, according to the feedback, led to a correct answer (subjects S3 and S4), produced roughly comparable data after the initial training period was completed.

A two-down one-up adaptive procedure, which was modeled after Levitt (1971), estimated the minimum change in the target level required to achieve a probability of a correct response of 0.7 in this paradigm. Each adaptive run consisted of 16 reversals and began with a random (large) initial value of ΔL (uniformly chosen between 15 and 25 dB). According to the two-down one-up adaptive rule, ΔL was initially adjusted by multiplying/dividing its current value in decibels by a scale factor of 1.8. After two reversals occurred, the increment was adjusted by multiplying/dividing by a scale factor of 1.4. The magnitude of the scale factor was further reduced when the fourth, sixth, and eighth reversals occurred to values of 1.2, 1.1, and 1.05, respectively. After the eighth reversal, the scale factor remained at 1.05 for eight additional reversals, at which point the adaptive run was concluded. The adaptive trials were self-paced and the subjects had an unlimited time to respond. The subjects received correct-answer feedback after every trial.

During each testing session, subjects completed four adaptive runs for each of the five different conditions. The ordering in which the conditions were presented was the same for each session: no distractor, fixed, roving phase, roving level, and double rove. Since the perceptual cues for each condition may have been different, subjects were alerted to the condition by using a unique letter for each condition. For each subject, data from the four adaptive runs for each distractor condition were collected in succession. The blocking of the adaptive runs in this manner was done to allow the subjects to refamiliarize themselves with the relevant perceptual cues. Subjects were given a minimum of 10 h of training before the reported data were collected. *Post hoc* analyses suggest that asymptotic performance was obtained early in the training period and that alerting the subjects to the condition before the block of runs began was sufficient to obtain the asymptotic performance. Further, in pilot listening, prolonged experience with the double-rove condition did not improve performance. For each condition, the reported results are based on 16 post-training adaptive runs.

C. Apparatus and materials

During the experiment, subjects sat in a sound treated room in front of a computer monitor. They responded “interval 2” or “interval 3” through a graphical interface via a computer mouse. On each trial, “lights” on a liquid crystal display monitor displayed the current interval number. The experiment was self-paced and listening sessions lasted for no more than 2 h with frequent rest breaks. A PC and Tucker–Davis Technology System II hardware (AP2, PD1, PA4, and HB6) generated the experimental stimuli at a sampling rate of 50 kHz. Stimuli were presented over Sennheiser HD 265 headphones.

D. Data analysis

Level discrimination thresholds have been reported in many forms and the analysis reported here follows the recommendations of Buus and Florentine (1991). In accordance with their recommendations, the performance metric used is ΔL , which is the decibel change of the target level. Also in accordance with the conclusions of Buus and Florentine (1991), the value ΔL is displayed on a logarithmic scale. The analysis and statistics are therefore based on the logarithm of ΔL . For example, the threshold was estimated from the geometric mean of the reversals of the adaptive runs and when averaging thresholds across conditions, the geometric mean was used. The results were also analyzed by considering ΔL on a linear display (i.e., using the arithmetic mean) and both analyses lead to similar conclusions.

The data from the experiment are reported in two ways. First, level discrimination thresholds were determined from the geometric mean of the values of ΔL that occurred on the last eight reversals of each adaptive run. Second, psychometric functions were fitted to the data from all the trials of the adaptive runs. The binary data for single trials (correct or incorrect) were used in fitting the psychometric functions instead of estimates of the probability of a correct response for each value of ΔL based on multiple trials since only a

few trials were conducted with each value of ΔL . Even though there were hundreds of trials per condition, the experimental paradigm resulted in a large number of different values of ΔL being used; the adaptive runs began with a random initial value of ΔL and on each trial, ΔL was adjusted using the above-mentioned adaptive rule.

In accordance with Buus and Florentine (1991), d' is assumed to be proportional to ΔL . Converting the correct/incorrect data into d' is problematic since there are many values of ΔL for which there were only a small number of trials that the subjects got all correct, leading to a d' of infinity. By assuming that (1) d' is proportional to ΔL , (2) the observer is unbiased (which is reasonable in a 4I-2AFC task), and (3) performance is limited by Gaussian noise, one can relate d' to the probability of a correct response. Specifically, with these assumptions, the probability of a correct response depends on ΔL and can be expressed as

$$P_{\text{correct}}(\Delta L) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\Delta L} e^{-x^2/2\sigma^2} dx, \quad (1)$$

where σ is a fitting parameter and is related to the proportionality between d' and ΔL . Note that Eq. (1) implies that when ΔL is equal to zero (corresponding to no change in the target level) the probability of a correct response is 0.5, as is expected, and that σ is the only fitting parameter. When fitting psychometric functions to the subject data, the parameter σ was adjusted to minimize the root-mean-squared (rms) error between the predicted percent correct and all of the data for a given subject and condition (collapsed over adaptive runs).

For each subject and condition, analysis of the fitted psychometric function is based on the parameter σ and the rms error statistics. Confidence intervals for both σ and the rms error were calculated by randomly drawing, with replacement, the results of N trials (where N is the total number of trials for a given subject and condition) and fitting a psychometric function to these sampled data. For each subject and condition, 1000 random drawings were made. An additional estimate of the threshold is obtained by substituting the appropriate value of σ into Eq. (1) and then solving for the ΔL that yields a probability of a correct response of 0.7. These estimates of the threshold based on the psychometric functions agree with the measurements of the threshold based on the reversals of the adaptive runs and are not explicitly presented.

III. MODELING

The dominant perception of the stimulus when the target and distractor are present is a fused binaural image with a salient loudness and lateral position. It is unclear whether the subjects have access to the level of the monaural target through secondary cues, which are cues that are not necessarily direct perceptions of the stimulus at a single ear but that provide sufficient information to reconstruct the stimulus at a single ear. Our modeling is intended to provide evidence in regard to the existence (or absence) of the monaural processing channels that appear in most binaural models. A model that includes monaural processing channels would not

predict an effect of a contra-aural distractor, and the predictions of our experiments are trivial for that case. Instead, we investigate a model in which there are no separate monaural channels. Specifically, we consider a detection theoretic model based on a two-dimensional decision space that roughly corresponds to the loudness and lateral position of the fused binaural image. Other perceptual variables, such as the image width, are not included nor are temporal characteristics that would be expected to be important in some experiments (especially in studies of binaural masking level differences), such as distributions of interaural differences or lateral positions over time.

The two dimensions of the model denoted Λ and Θ are functions of the level at the left ear, L_{left} (in decibels), the level of the right ear, L_{right} (in decibels), the interaural time difference, T (in microseconds), and two internal noises N_{Λ} and N_{Θ} (both in decibels). The dimensions Λ and Θ are defined as

$$\Lambda = 10 \log_{10}(10^{L_{\text{left}}/10} + 10^{L_{\text{right}}/10}) + N_{\Lambda} \quad (2)$$

and

$$\Theta = L_{\text{left}} - L_{\text{right}} + kT + N_{\Theta}, \quad (3)$$

where k is the intensity-time trading ratio in $\text{dB}/\mu\text{s}$, which is related to the time-intensity trading ratios used by [Hafters \(1971\)](#) and [Yost \(1972\)](#). To aid comparisons between Λ and Θ , we have defined Θ in decibels even though it can be logically expressed in a number of other ways (e.g., interaural time difference or a dimensionless quantity related to a location along the interaural axis). Note that, under monotic conditions (i.e., either L_{left} or L_{right} is equal to negative infinity), Θ is equal to either positive or negative infinity and indicates an extreme position toward the stimulated ear.

Models based on position variables such as Θ have been used to predict the results of many binaural experiments ([Hafters, 1971](#); [Yost, 1972](#)). Although models based on position variables have a strong predictive power for some types of experiments, they are insufficient to predict all binaural effects. Importantly, [Bernstein \(2004\)](#) reported that subjects could outperform an ideal observer of Θ when discriminating the ILD of tones with a random interaural time difference (ITD). The discrepancy between [Bernstein's \(2004\)](#) measurements of ILD discrimination and predictions of a simple position variable may be overcome by a variety of modifications, as discussed by [Stern and Trahiotis \(1997\)](#) in their review of position variable models, such as the image width or a comparison of separate monaural levels [as in the level meter model of [Hartmann and Constan \(2002\)](#)]. In this study, we consider predictions for monaural level discrimination based on the use of the overall (binaural) level Λ and the unmodified position variable Θ [as defined in Eqs. (2) and (3)].

Models based on the Λ dimension (the overall binaural level) have been evaluated as an aspect of central spectrum models ([Bilsen, 1977](#); [Bilsen and Raatgever, 2000](#)). These studies of the central spectrum have focused on issues concerning across-frequency integration. Binaural loudness has been studied in a variety of ways (e.g., [Zwicker and Zwicker, 1991](#); [Sivonen and Ellermeier, 2006](#); [Whilby et al., 2006](#)).

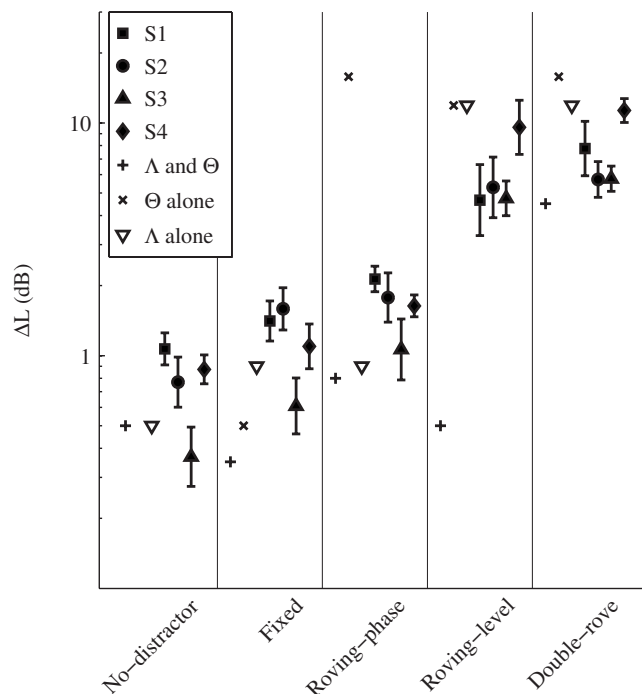


FIG. 1. Mean thresholds for the four subjects under the five different conditions. Error bars are two times the standard error of the mean. Note that the standard deviations are twice as large as two times the standard error of the mean based on 16 runs. The thresholds of the ideal observer of Λ and Θ , both alone and together, are also shown. In the no-distractor condition, the ideal observer of Θ never obtains threshold performance.

[Sivonen and Ellermeier \(2006\)](#) concluded that power summation (i.e., Λ) predicted binaural loudness best. Under monotic conditions, Λ mathematically reduces to the monaural level and models based on a monaural level detector have been explored for many monotic and diotic tasks.

Three models are explicitly considered here: the maximum likelihood observers of Λ alone, Θ alone, and Λ and Θ together. The dimensions Λ and Θ are specified with internal noises N_{Λ} and N_{Θ} , which are assumed to be zero-mean Gaussian random variables that are statistically independent across the dimensions, the observation intervals, and the trials. With these assumptions, there are three parameters, k , σ_{Λ} , and σ_{Θ} , which correspond to the intensity-time trading ratio and the standard deviations of the two internal noises. In order to predict generally accepted values for the just noticeable differences in the overall level, ILD, and ITD (cf. [Viemeister 1988](#); [Blauert, 1997](#)), the intensity-time trading ratio k is fixed at 1 dB per 20 μs , and the standard deviations of the internal noises (σ_{Λ} and σ_{Θ}) are fixed at 0.5 dB throughout all the modeling. The performance of these three ideal observers in a 4I-2AFC task is derived in Appendix A.

IV. RESULTS

A. Psychophysical results

Figure 1 contains the geometric mean of the thresholds (calculated from the reversals of the adaptive runs) of the four subjects for each of the five different conditions. A repeated measure analysis of variance test found statistically significant effects of distractor condition and subject and a

statistically significant interaction between distractor condition and subject ($p < 0.02$). The measured thresholds vary with the type of distractor and subject. Multiple planned paired t tests were used to test for statistically significant differences between conditions (all combinations). There are statistically significant differences ($p < 0.05$) in performance between all pairs of conditions except between the roving-level and double-rove conditions ($p = 0.09$). Performance in the no-distractor condition was the best with an average (across-subject geometric-mean) threshold value of ΔL of 0.7 dB. Performances with the roving-level and double-rove distractors were the worst with average thresholds of 5.8 and 7.3 dB, respectively. The fixed and roving-phase distractors showed small detrimental effects on performance, with average thresholds of 1.1 and 1.6 dB, respectively. Displaying the threshold value of ΔL (a decibel value) on a logarithmic scale, as opposed to a linear scale, expands the differences between the no-distractor condition and the fixed and roving-phase distractor conditions and compresses the differences between the no-distractor condition and the roving-level and double-rove distractor conditions. The subjects are clearly not basing their decisions on the level at the target ear; in other words, they are not “listening to one ear” since the addition of a distractor at the nontarget ear decreases performance.

Figure 2 shows example psychometric functions for a representative subject (S2) for all five conditions. The psychometric functions take into account all trials, whereas the threshold measurements take into account only trials at which reversals in performance occurred. Since the data were collected using an adaptive paradigm, most trials had values of ΔL near threshold. Traditionally, psychometric functions are not constructed from data collected with adaptive paradigms. We wished, however, to determine the extent to which performance was monotonic in ΔL . Visually, the data appear monotonic and generally consistent with the sigmoid function of Eq. (1). The values of σ (the single fit parameter for the psychometric function) that best fit the data are presented in the top panel of Fig. 3. Consistent with the threshold data presented in Fig. 1, the value of σ that best fits the data systematically varies across the conditions. Paired t tests show no statistically significant differences ($p > 0.05$) in the values of σ among the no-distractor, fixed, and roving-phase conditions nor are there statistically significant differences between the roving-level and double-rove conditions ($p > 0.05$). The differences in the values of σ for the no-distractor, fixed, and roving-phase conditions and for the roving-level or double-rove conditions, however, are statistically significant ($p < 0.025$).

In addition to the changes in σ , there are also changes in the rms error between the fitted psychometric function and the data. The bottom panel of Fig. 3 shows the rms error. The rms error and the visual agreement between the fitted psychometric functions and the data are similar for all of the subjects. Paired t tests show no statistically significant differences ($p > 0.05$) in the rms error for the no-distractor, fixed, and roving-phase conditions nor for the roving-level and double-rove conditions. The differences in the rms error for the no-distractor, fixed, and roving-phase conditions and for

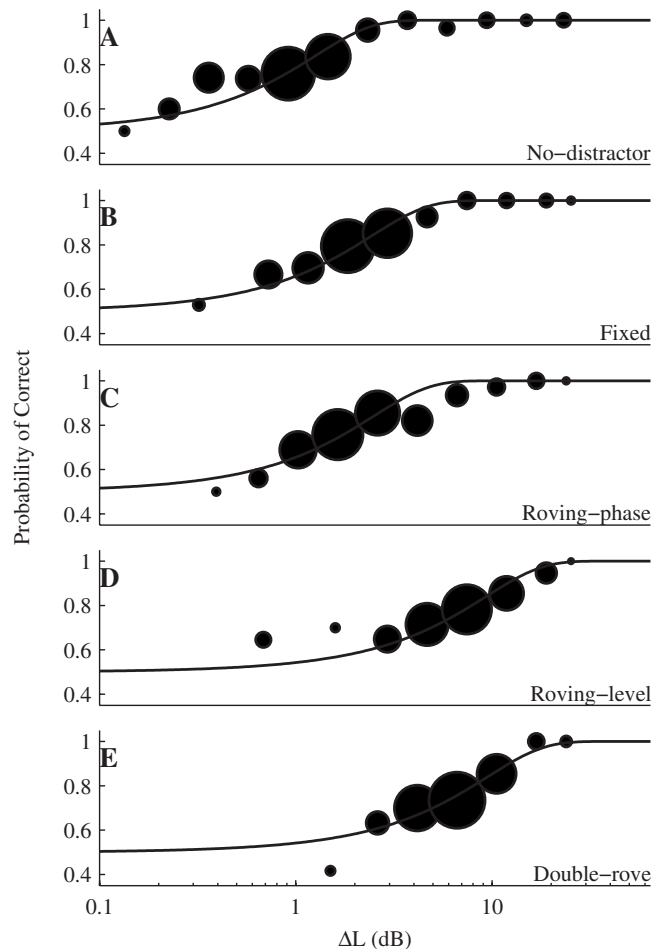


FIG. 2. Examples of the dependence of the probability of a correct response on the target increment for subject S2 for the five different conditions. Panels (A)–(E) correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively. The data have been binned according to ΔL , and the size of the symbol is proportional to the number of trials that occurred within the bin. The best fitting psychometric functions by using the single parameter function in Eq. (1) are also shown for each condition.

the roving-level and double-rove conditions, however, are statistically significant ($p < 0.025$). These changes in the rms error are indicative of a change in how well the fitted psychometric functions fit the data.

B. Decision variable distributions

In order to understand the information about the level of the target, which is carried by the two decision variables Λ and Θ , their probability densities for a given L_{target} are examined. We consider the joint density of Λ and Θ ($f_{\Lambda, \Theta|L_{\text{target}}}$) as well as the densities of Λ ($f_{\Lambda|L_{\text{target}}}$) and Θ ($f_{\Theta|L_{\text{target}}}$) in isolation. The manner in which these probability densities were computed is described in Appendix B. For simplicity, we define L_{target} as being equal to the reference level L_0 plus an increment ΔL and plot the density functions for five different values of ΔL (0, 2, 4, 8, and 16 dB), such that, for the unincremented target, ΔL is equal to zero.

Figure 4 shows $f_{\Lambda|L_{\text{target}}}$ for the five different conditions. Since the distractor phase has no effect on Λ , $f_{\Lambda|L_{\text{target}}}$ is identical in the fixed and roving-phase conditions as well as

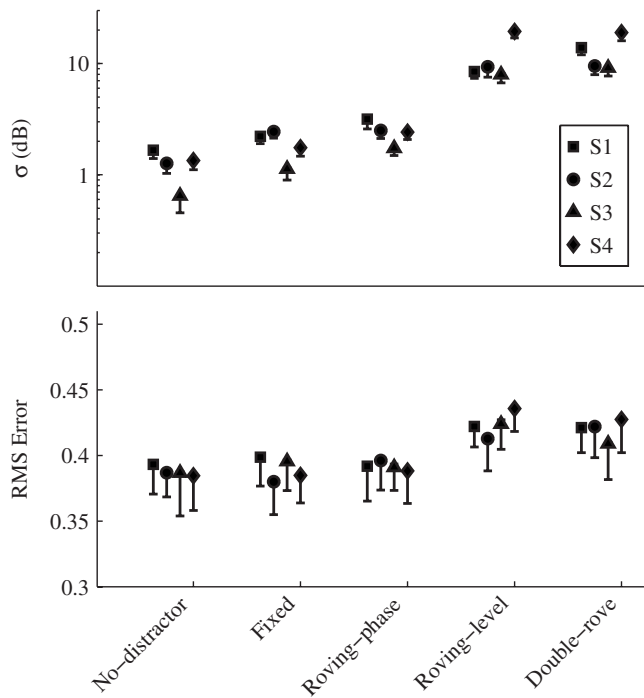


FIG. 3. Fit parameter σ (top panel) and the RMS error (bottom panel) for the psychometric functions that were fitted to the data of the four subjects under the five different conditions. Error bars are the 95% confidence intervals derived by resampling the data. When the error bars are absent, the confidence interval is on the order of the size of the symbol.

in the roving-level and double-rove conditions. In the no-distractor, fixed, and roving-phase conditions $f_{\Lambda|L_{\text{target}}}$ is Gaussian with standard deviation σ_{Λ} and a mean that depends on ΔL . The mean in the fixed and roving-phase conditions is higher than in the no-distractor condition due to the additional distractor energy; the effect of this additional energy decreases with increasing ΔL since Λ is calculated by adding the intensities (units of power per area) and not the decibel levels. In the roving-level and double-rove conditions, the random distractor level affects Λ , and therefore, $f_{\Lambda|L_{\text{target}}}$ is much broader. In these two conditions, changes to ΔL affect both the mean and the shape.

Figure 5 shows $f_{\Theta|L_{\text{target}}}$ in four of the conditions (the no-distractor condition is not included since Θ is undefined). Changes to ΔL again affect the mean in all conditions but do not affect the shape of $f_{\Theta|L_{\text{target}}}$ in any condition. In the fixed condition, $f_{\Theta|L_{\text{target}}}$ is Gaussian (with standard deviation σ_{Θ}). In the roving-phase and roving-level conditions, Θ depends on both the internal noise and a uniformly distributed random variable (either the distractor phase or level) and $f_{\Theta|L_{\text{target}}}$ is nearly uniform over a large range. In the double-rove condition, Θ is the sum of two uniformly distributed random variables (the distractor level and phase) and the internal noise and $f_{\Theta|L_{\text{target}}}$ is nearly trapezoidal in shape.

Figure 6 shows the region for which $f_{\Lambda, \Theta|L_{\text{target}}}$ is greater than 0.0001 for the roving-level and double-rove conditions, respectively. In the roving-level condition, the probability of Θ conditioned on Λ is narrow, and changes in ΔL substantially shift the distribution, making monaural level discrimination possible for the ideal observer (i.e., the distributions for the unincremented and incremented targets do not sub-

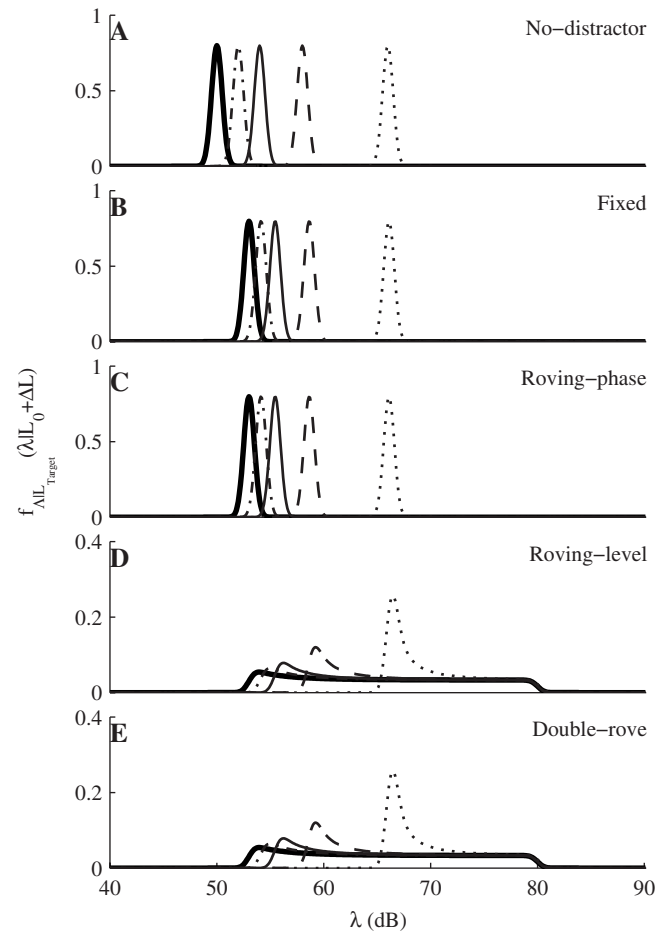


FIG. 4. Probability distributions of the overall level variable Λ . Panels (A)–(E) correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively. Note that each panel plots the density function of Λ as a function of its argument λ for the target level L_{target} equal to the reference level L_0 plus the monaural level increment ΔL . Specifically, $f_{\Lambda|L_{\text{target}}}$ is plotted for five values of ΔL : 0 dB (thick solid), 2 dB (dot dash), 4 dB (thin solid), 8 dB (dashed), and 16 dB (dotted).

stantially overlap). However, the complexity of the distributions may lead a nonoptimal observer to have a substantially degraded performance. In the double-rove condition, the variables Λ and Θ together do not carry accurate information; the probability of Θ conditioned on Λ is broad (i.e., each value of Λ now corresponds to a range of Θ values) and changes in the increment size have only small effects. Therefore, the ideal observer of Λ and Θ together cannot discriminate between the unincremented and incremented target levels.

In summary, the Λ and Θ dimensions carry information both individually and jointly about the target level. Introducing variability into the distractor phase decreases the information in Θ but does not compromise the information in Λ . Introducing variability into the distractor level reduces the information in Λ and Θ individually but does not reduce the joint information. In order to substantially decrease the performance of the observer of both Λ and Θ together, and thereby making it advantageous to utilize a secondary perception such as the output of monaural processing channels (if it is available), variability must be introduced into both the distractor phase and level (i.e., the double-rove condition).

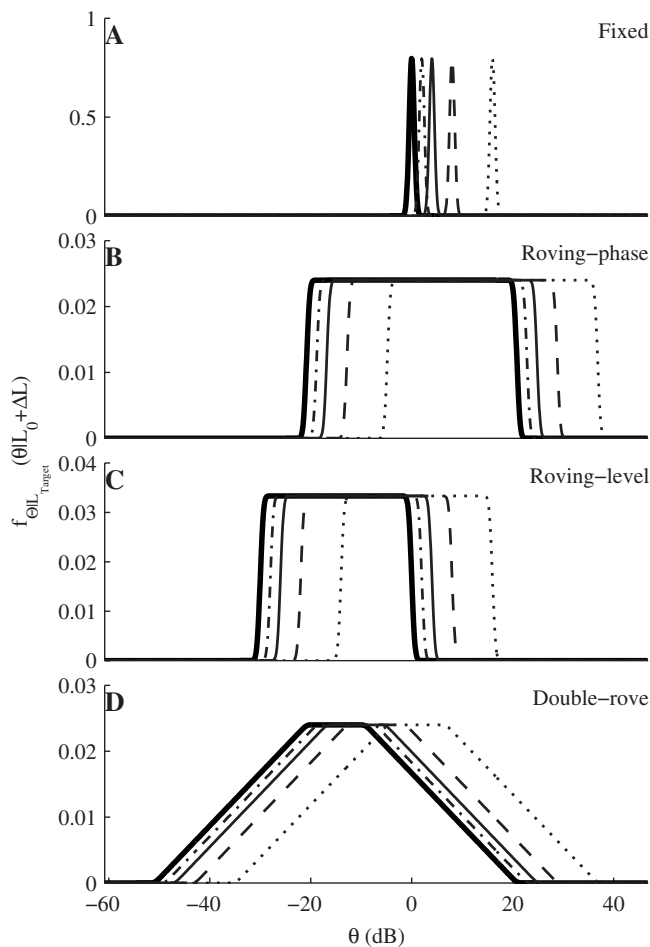


FIG. 5. Probability density functions $f_{\Theta|L_{\text{target}}}$ for the lateral position variable Θ for multiple monaural levels, which are similar to the plots in Fig. 4. In the no-distractor condition, Θ is infinite and is not shown. Panels (A)–(D) therefore correspond to the fixed, roving-phase, roving-level, and double-rove conditions, respectively.

C. Model predictions

The Λ alone, Θ alone, and Λ and Θ together ideal-observer models can be used to make predictions of both thresholds and psychometric functions. The predicted thresholds for the three models are included along with the measured thresholds in Fig. 1. The predicted psychometric functions for the five different distractor conditions are shown in Fig. 7 along with the average empirical psychometric functions for the subjects. We consider the predicted thresholds of the ideal observer of Θ alone first, which is followed by the ideal observer of Λ alone and then the ideal observer of Λ and Θ together. Finally, we consider the predicted psychometric functions for all three models.

The ideal observer of Θ alone performs best in the fixed condition with a threshold of σ_{Θ} (assumed here to be 0.5 dB). The ideal observer of Θ alone never obtains threshold performance in the no-distractor condition since Θ is equal to minus infinity in this case independent of the target level. This is consistent with the position providing no information about the target level in monotic conditions. In the roving-phase, roving-level, and double-rove conditions, the thresholds are 16.9, 12.8, and 16.9 dB, respectively. The Θ alone model is not a good predictor of performance in both

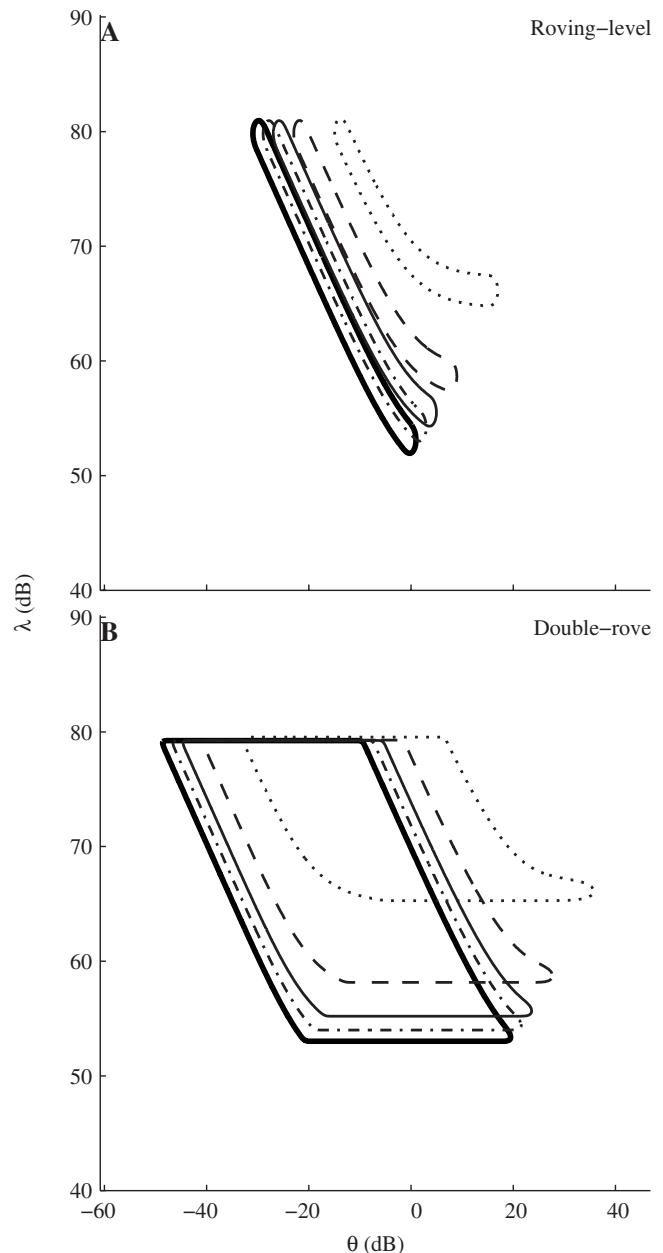


FIG. 6. Contours that enclose the region for which $f_{\Lambda, \Theta|L_{\text{target}}}$ is greater than 0.0001. Note that within these regions, $f_{\Lambda, \Theta|L_{\text{target}}}$ is not uniform. Panels (A) and (B) correspond to the roving-level and double-rove conditions, respectively. As in Figs. 4 and 5, five different values of ΔL are shown. The five different values of ΔL are 0 dB (thick solid), 2 dB (dot dash), 4 dB (thin solid), 8 dB (dashed), and 16 dB (dotted).

the no-distractor condition and the roving-phase condition. In all but the fixed condition, there is insufficient information in Θ to predict the measured thresholds. Although the perceived lateralization of the stimulus may be influencing the subjects' decisions, the Θ alone model is clearly an exceedingly poor predictor of the measured discrimination thresholds.

The ideal observer of Λ alone performs best in the no-distractor condition with a threshold equal to σ_{Λ} (assumed here to be 0.5 dB). In both the fixed and roving-phase conditions, the predicted thresholds are 0.9 dB; the predicted performance is slightly worse due to the added energy of the distractor. In both the roving-level and double-rove condi-

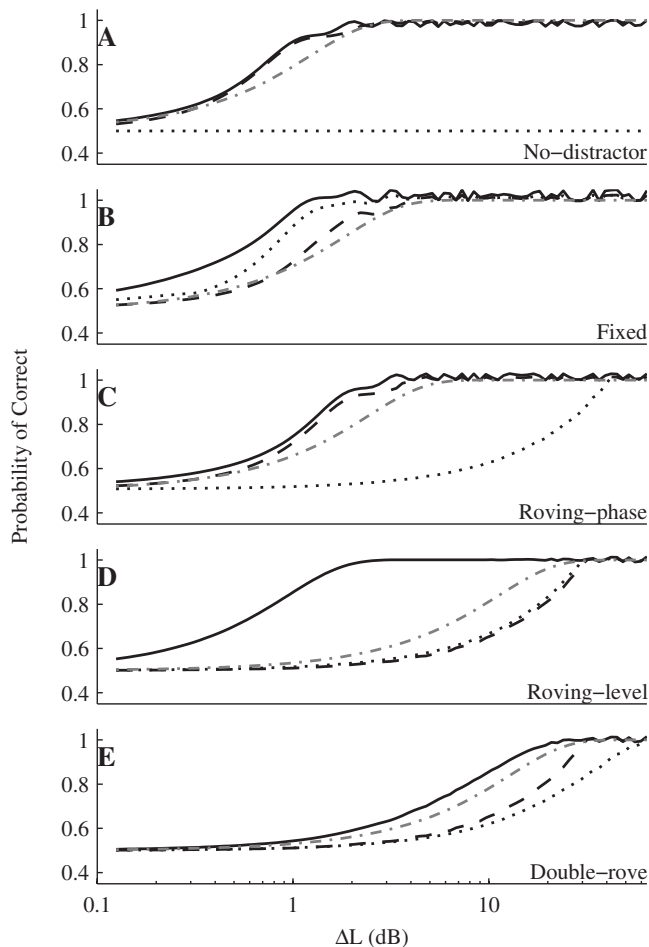


FIG. 7. Predicted psychometric functions for the five conditions for the ideal observer of Λ alone (dashed), Θ alone (dotted), and Λ and Θ together (solid). Also shown is the across-subject average fitted psychometric function (dot dash). Panels (A)–(E) correspond to the no-distractor, fixed, roving-phase, roving-level, and double-rove conditions, respectively.

tions, the predicted thresholds are 13.7 dB and are predominantly determined by the range of the level rove. For all the conditions, the predicted thresholds are in general agreement with the measured thresholds. The empirical fact that roving the level (with or without roving the phase) causes such a large elevation in the measured thresholds suggests that the subjects heavily relied on the overall loudness for some of their decisions. Although the Λ alone model is a better predictor of performance than the Θ alone model, it is important to note that the predictions in the roving-level and double-rove conditions are significantly worse than the measured thresholds. This means that the subjects could not have been basing their decisions solely on Λ .

The ideal observer of Λ and Θ together performs best in the fixed condition with a threshold of 0.3 dB since the two internal noises N_Λ and N_Θ are statistically independent. In the no-distractor condition, the threshold is 0.5 dB. In the roving-phase, roving-level, and double-rove conditions, the thresholds were, respectively, 0.9, 0.5, and 4.8 dB. For the Λ and Θ together model (unlike the Θ alone and Λ alone models), the predicted thresholds are seen to be less than or equal to the average measured thresholds in all of the conditions. In other words, there is enough information in Λ and Θ taken

together to perform as well or better than the subjects. However, as indicated by the predicted threshold being substantially lower than the measured thresholds in the fixed and roving-level conditions, the subjects are failing to make use of all the available information.

Figure 7 shows the predicted psychometric functions of the ideal observer of Λ alone (dashed), Θ alone (dotted), and Λ and Θ together (solid), along with the average empirical psychometric functions (dash dot) for the subjects. The differences among the psychometric functions predicted by the three models in the five conditions are indicative of changes in the information carried by Λ and Θ across the conditions. By comparing the model predictions and the empirical data over the whole range of ΔL (not only at the threshold), additional discrepancies between the predictions and empirical data become evident. Some of the theoretical psychometric functions differ from the empirical ones not only in the position on the abscissa (i.e., threshold) but also in shape.

V. DISCUSSION

It is clear from the results shown in Fig. 1 that the ability to discriminate the level of a monaural target can be severely degraded by the introduction of a contra-aural distractor. This contra-aural interference cannot be predicted by a model that includes monaural processing channels, which by definition are unaffected by stimulation of the other ear. In the models evaluated in this study, the measured performance is not predicted by the model based on lateral position; however, both models based on overall loudness and on loudness and position together show some promise in predicting the measured effects.

Although the Λ alone model reasonably predicts much of the data, there are fundamental problems. The main problem with this model occurs in the roving-level and double-rove conditions, where the subjects are seen to do modestly (but significantly) better than the model. Although the deviations between the model and data for these conditions are not large, they cannot be eliminated simply by decreasing the internal noise parameter. Even if all the internal noise were eliminated, the predicted performance in the double-rove condition would still be worse than that achieved by the subjects. Within the context of our models, this in turn implies that the subjects' decisions cannot be based solely on Λ but that Θ must also be considered.

The Λ and Θ together model predictions are reasonably consistent with the data in the no-distractor, roving-phase, and double-rove conditions (although the predictions are slightly lower than the average data in all three of these conditions). However, for the fixed condition, the prediction is substantially too low, and for the roving-level condition, it is monstrously too low. Thus, although it appears (as discussed above) that the subjects must be making some use of Θ as well as Λ , it is clear that their ability to use both together is clearly suboptimal. It seems unlikely that the less-than-optimum performance can be simply explained by assuming a degradation (i.e., increased internal noise) in performance caused by the need to simultaneously estimate two variables and the resulting problem of divided attention (e.g.,

Bonnel and Hafter, 1988). Rather, consistent with Fig. 6, it appears that to obtain a good performance in the roving-level condition requires the observer of both Λ and Θ to precisely combine the two observations. Therefore, in order to accurately predict the psychophysical results, the modeled observer must be modified such that it cannot precisely combine two observations.

The Λ and Θ together model assumes that the contra-aural interference arises because the subjects do not have access to the output of monaural processing channels. In the next paragraphs, several alternative hypotheses for the causes of poor performance are discussed and rejected. Specifically, we consider the effects of acoustic cross talk, subject confusion, and inadequate training.

The measured contra-aural interference could occur because the distractor corrupts the inputs to the monaural processing channels. Specifically, that acoustic cross talk results in the level at the target ear being influenced by the distractor. The effects of cross talk can be approximated by assuming that an attenuated and delayed version of the distractor is added (in units of pressure) to the target. Since the amplitude of the sum of two identical-frequency sinusoids depends on both the relative levels and phases, the level of the corrupted target is variable in the roving-phase, roving-level, and double-rove conditions. To limit the performance of the ideal observer, which is degraded only by the cross talk, to the empirically measured threshold in the double-rove condition, the amount of attenuation across the head would need to be less than 15 dB (substantially less than the typically assumed 40 dB).

It is also unlikely that the measured contra-aural interference is simply a result of the subjects being confused. The measured thresholds in the no-distractor condition agree with previous measures of monotic level discrimination thresholds (Viemeister, 1988). The thresholds in the fixed condition are similar to ILD discrimination thresholds, where the ILD is imposed by incrementing the level at one ear and decrementing the level at the other ear (Blauert, 1997). Additionally, if the subjects were confused, then introducing across-presentation variability to the distractor should degrade performance, but roving the distractor phase (i.e., perceived lateral position) did not substantially increase the thresholds of the subjects. Performance was only degraded by introducing across-presentation variability to the distractor level. However, introducing across-presentation variability to the overall level has little effect on spectral profile analysis (Green, 1988) or ILD discrimination (Bernstein, 2004). It seems that in the current task, performance is only degraded when the across-presentation distractor variability affects the information about the target level in both the overall loudness and the lateral position.

Finally, it also seems unlikely that subjects have been simply inadequately trained. The testing scheme that was employed in these experiments (the 4I-2AFC paradigm with trial-by-trial correct-answer feedback) is known to provide rapid learning and good performance in situations that may be confusing or difficult to learn (Trahiotis, 1992). Consistent with this established property of the testing scheme is that despite large differences in the relevant previous expe-

rience of the subjects (e.g., S1 was the first author of this article while S3 and S4 were naive), the measured discrimination performance did not appreciably vary across the subjects. Further, all our measured psychometric functions appear smooth and monotonically increasing with ΔL . Thus, the measured increases in the threshold due to the inclusion of the distractor are unlikely to be caused by artifacts associated with the adaptive procedure or by major perceptual changes as ΔL is varied. Rather, it seems likely that the contra-aural interference occurs because the subjects do not have access to the output of monaural processing channels.

VI. CONCLUDING REMARKS

The results of the experiments presented here strongly support the idea that when judging changes in the level of a target presented at one ear, listeners are unable to ignore a contra-aural distractor. Therefore, models in which it is assumed that the listener has access to monaural processing channels (either directly or through secondary perceptual attributes such as the time image or spatial width combined with the overall level) cannot be successfully applied to this experiment. Specifically, listeners' thresholds were increased by an order of magnitude (relative to the measured monaural performance) in some conditions for our 600 Hz stimuli. In general, subjects perceived a single compact image and reported that they used the loudness and the lateral position of this image for many of their judgments. A model based on the joint use of decision variables that correspond to the loudness and the lateral position of the primary image was consistent with many of the results, although performance in some conditions was notably poorer than that predicted when the internal noise was chosen to be consistent with the typical discrimination threshold for ITD and overall level. It is speculated that, although both the loudness and the lateral position can be used, there are difficulties in simultaneously using both for refined decisions. Furthermore, when a wider set of experiments is considered, it becomes clear that a model that only includes the loudness and the lateral position is not complete and that additional decision variables are needed to match the observed performance in other experiments. Overall, one can conclude that an adequate theory will likely involve degraded processing of the loudness and lateral position variables, the inclusion of further binaural decision variables, and the exclusion (at least in some conditions) of access to monaural channels.

ACKNOWLEDGMENTS

This research was supported by NIH/NIDCD Grant Nos. R01 DC00100, P30 DC004663, and F31 DC006769. The authors would also like to thank Dr. Frederick Gallun, Dr. Andrew Oxenham, and Dr. Bertrand Delgutte for reading a previous version of this manuscript. The comments by Dr. Armin Kohlrausch and two anonymous reviewers were also very helpful.

APPENDIX A

In this appendix, a general expression for the probability of a correct response in a 4I-2AFC task is calculated for

three observers in the $\{\Delta, \Theta\}$ space, where Λ and Θ are defined by Eqs. (2) and (3) in the text. Three models are considered: the ideal observer of Λ and Θ together, the ideal observer of Λ alone, and the ideal observer of Θ alone. The task requires the observer to discriminate the level of the target, L_{target} . On a given interval, the target is either unincremented, such that L_{target} is equal to L_0 (the reference level), or the target is incremented, such that L_{target} is equal to the sum of L_0 and ΔL . To calculate performance we note the following: (1) on each interval, there is a single observation of both Λ and Θ and (2) on a single trial, there are eight total observations (four of Λ and four of Θ). Due to the experimental paradigm, the observations in the first and last intervals carry no information for the ideal observers,¹ and therefore, only four observed values (two pairs) are relevant. The observation of Λ on the second interval is denoted λ_2 ; the observation of Λ on the third interval is denoted λ_3 . Similarly, the observation of Θ on the second interval is denoted θ_2 and that on the third interval is denoted θ_3 . The ideal observer of Λ and Θ together is considered first since the performance of the ideal observers of Λ alone and Θ alone follow from the ideal observer of Λ and Θ together.

The ideal observer of both Λ and Θ depends on two four-dimensional joint probability functions. The first is the probability densities of λ_2 , λ_3 , θ_2 , and θ_3 given that an increment with size ΔL occurred on the second interval; the second is the probability density of λ_2 , λ_3 , θ_2 , and θ_3 given that an increment with size ΔL occurred on the third interval. These four-dimensional joint probability functions can be written as the product of two two-dimensional joint probability functions by noting that when the interval in which the increment occurred is given, the observation of λ_2 is independent of λ_3 and the observation of θ_2 is independent of θ_3 . Since there are two intervals in which the target level can be incremented, there are four relevant two-dimensional probability density functions.

The relevant two-dimensional joint probabilities are the probability of the observed values of Λ and θ on a particular interval given a target level. The log-likelihood ratio $\eta_{\Lambda, \Theta}$ is defined in terms of these probabilities as

$$\eta_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L) = 10 \log_{10} \left(\frac{f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_2, \theta_2|L_0) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_3, \theta_3|L_0 + \Delta L)}{f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_2, \theta_2|L_0 + \Delta L) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_3, \theta_3|L_0)} \right).$$

The notation in this equation is designed to distinguish the identity of the functions from the values and variable in the arguments. Thus, the subscripts on $\eta_{\Lambda, \Theta}$ identify that this likelihood ratio applies to the case when both variables are available and it has five arguments.

The ideal observer is represented by a binary indicator function $\psi_{\Lambda, \Theta}$, which is calculated from the likelihood ratio. Specifically, when the *a priori* probabilities of each interval are equal and when the goal is to maximize the probability of a correct response, the optimum decision is determined by the sign of the log-likelihood ratio $\eta_{\Lambda, \Theta}$: when $\eta_{\Lambda, \Theta}$ is positive, the second interval is most likely to have the incremented target and that is the optimum decision. When $\eta_{\Lambda, \Theta}$ is negative, the third interval is more likely to have the incremented target and that is the optimum decision. This decision rule can be represented by the binary indicator $\psi_{\Lambda, \Theta}$, which is equal to unity when the optimum decision is the second interval and zero otherwise. Mathematically, the indicator function is

$$\psi_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L) = \begin{cases} 1 & \text{when } \eta_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L) \geq 0 \\ 0 & \text{when } \eta_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L) < 0 \end{cases}.$$

Then, the probability that the ideal observer of both Λ and Θ achieves the correct answer is a function of ΔL and can be written as

$$P_{\text{correct}}(\Delta L) = \frac{1}{2} \int \int \int \int [\psi_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_2, \theta_2|L_0) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_3, \theta_3|L_0 + \Delta L)] d\theta_2 d\lambda_2 d\theta_3 d\lambda_3 \\ + \frac{1}{2} \int \int \int \int [(1 - \psi_{\Lambda, \Theta}(\lambda_2, \theta_2, \lambda_3, \theta_3, \Delta L)) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_2, \theta_2|L_0 + \Delta L) f_{\Lambda, \Theta|L_{\text{target}}}(\lambda_3, \theta_3|L_0)] d\theta_2 d\lambda_2 d\theta_3 d\lambda_3.$$

Thus, for the optimum decision rule, the probability of a correct response depends only on the joint probability density function of Λ and Θ given L_{target} since the indicator function is also defined in terms of this density function. This joint probability density function $f_{\Lambda, \Theta|L_{\text{target}}}$ is approximated with numerical methods and the details of this approximation are contained in Appendix B. The probability of a correct response for the ideal observers of Λ alone or of Θ alone is calculated in an analogous manner and the derivation is not presented.

APPENDIX B

In this appendix, analytical and numerical techniques are used to approximate the joint density function of Λ and Θ , as defined in Eqs. (2) and (3), for a target level L_{target} . As outlined in Appendix A, knowledge of the joint density allows the calculation of the probability of a correct response in our experiment. This appendix derives a relatively simple expression that can be evaluated using standard numerical functions and techniques. Before the details of the derivation of $f_{\Lambda, \Theta|L_{\text{target}}}$ are outlined, the model variables are related to

the experimental variables. Specifically, the values that are appropriate for the psychophysical experiment are substituted into Eqs. (2) and (3). All levels are in decibels. In the experiment, L_{left} is always the level of the target L_{target} and L_{right} is the level of the distractor. The level of the target is the sum of a reference level L_0 and an increment ΔL , such that $L_{\text{left}}=L_{\text{target}}=L_0+\Delta L$, where ΔL is zero when the target is not incremented. The distractor has a level that is equal to the reference level plus a random variable A ; the level of the right ear can, therefore, be written as $L_{\text{right}}=L_0+A$. The interaural time difference T is the phase delay, which is defined as the negative of the distractor phase (also a random variable) divided by the radian frequency ω (i.e., $T=-\Phi/\omega$). The psychophysical experiment specifies that A has a uniform probability density function between a_{min} and a_{max} . The experiment also specifies that Φ has a uniform probability density function between ϕ_{min} and ϕ_{max} . Using this notation, when the distractor level is roved, $a_{\text{min}}=0$ and $a_{\text{max}}=30$, and when the distractor level is fixed, $a_{\text{min}}=a_{\text{max}}=0$. Similarly, when the distractor phase is roved, $\phi_{\text{min}}=-\pi/2$ and $\phi_{\text{max}}=\pi/2$, and when the distractor phase is fixed, $\phi_{\text{min}}=\phi_{\text{max}}=0$. Making these substitutions into Eqs. (2) and (3) results in

$$\Lambda = 10 \log_{10}(10^{L_{\text{target}}/10} + 10^{(L_0+A)/10}) + N_{\Lambda} \quad (\text{B1})$$

and

$$\Theta = L_{\text{target}} - (L_0 + A) - \frac{k}{\omega}\Phi + N_{\Theta}. \quad (\text{B2})$$

Our derivation of $f_{\Lambda, \Theta|L_{\text{target}}}$ begins by using the definition of conditional probability to expand the joint density function to

$$f_{\Lambda, \Theta|L_{\text{target}}} = f_{\Lambda|L_{\text{target}}}(\lambda|L_0 + \Delta L) f_{\Theta|\Lambda, L_{\text{target}}}(\theta|\lambda, L_0 + \Delta L).$$

Then, by using the fact that $f_{\Theta|\Lambda, L_{\text{target}}}$ can be obtained by integrating $f_{\Theta, A, \Phi|\Lambda, L_{\text{target}}}$ over all values of a and ϕ representing the values of the variables A and Φ , one obtains

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= f_{\Lambda|L_{\text{target}}}(\lambda|L_0 + \Delta L) \\ &\times \int \int [f_{\Theta|A, \Phi, \Lambda, L_{\text{target}}}(\theta|a, \phi, \lambda, L_0 + \Delta L) \\ &\times f_{A, \Phi|\Lambda, L_{\text{target}}}(a, \phi|\lambda, L_0 + \Delta L)] d\phi da. \end{aligned}$$

Making a substitution of Θ based on Eq. (B2) and using the definition of conditional probability yields

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= f_{\Lambda|L_{\text{target}}}(\lambda|L_0 + \Delta L) \int \int f_{N_{\Theta}}\left(\frac{k}{\omega}\phi - \mu_{\Theta}(a)\right) \\ &\times f_{A, \Phi|\Lambda, L_{\text{target}}}(a, \phi|\lambda, L_0 + \Delta L) d\phi da, \end{aligned}$$

where $\mu_{\Theta}(a)$ is equal to $\Delta L - a - \theta$. Using the definition of conditional probability for $f_{A, \Phi|\Lambda, L_{\text{target}}}$ and then noting the independence of Φ and A , Λ , and L_{target} gives

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= f_{\Lambda|L_{\text{target}}}(\lambda|L_0 + \Delta L) \int \int f_{N_{\Theta}}\left(\frac{k}{\omega}\phi - \mu_{\Theta}(a)\right) \\ &\times f_{A|\Lambda, L_{\text{target}}}(a|\lambda, L_0 + \Delta L) f_{\Phi}(a) d\phi da. \end{aligned}$$

By using the definition of conditional probability on $f_{A|\Lambda, L_{\text{target}}}$, noting the statistical independence of $f_{\Lambda|L_{\text{target}}}$ and L_{target} , and simplifying, one obtains

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= \int f_A(a) f_{\Lambda|\Lambda, L_{\text{target}}}(\lambda|a, L_0 + \Delta L) \\ &\times \int f_{\Phi}(a) f_{N_{\Theta}}\left(\frac{k}{\omega}\phi - \mu_{\Theta}(a)\right) d\phi da. \end{aligned}$$

By making use of the uniform probability density functions of the random variables A and Φ , $f_{\Lambda, \Theta|L_{\text{target}}}$ can be rewritten as

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= c \int_{a_{\text{min}}}^{a_{\text{max}}} f_{\Lambda|\Lambda, L_{\text{target}}}(\lambda|a, L_0 + \Delta L) \\ &\times \int_{\phi_{\text{min}}}^{\phi_{\text{max}}} f_{N_{\Theta}}\left(\frac{k}{\omega}\phi - \mu_{\Theta}(a)\right) d\phi da, \end{aligned}$$

where c is equal to $1/(a_{\text{max}} - a_{\text{min}})(\phi_{\text{max}} - \phi_{\text{min}})$.

From Eq. (B1) it follows that

$$f_{\Lambda|\Lambda, L_{\text{target}}} = f_{N_{\Lambda}}(\lambda - 10 \log_{10}(10^{(L_0+\Delta L)/10} + 10^{(L_0+a)/10})).$$

Making substitutions for the density functions of N_{Λ} and N_{Θ} yields

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &= \frac{c}{2\pi\sigma_{\Theta}\sigma_{\Lambda}} \int_{a_{\text{min}}}^{a_{\text{max}}} e^{-((\lambda - \mu_{\Lambda}(a))^2/2\sigma_{\Lambda}^2)} \\ &\times \int_{\phi_{\text{min}}}^{\phi_{\text{max}}} e^{-(\phi - \omega/k\mu_{\Theta}(a))^2/2(\omega/k\sigma_{\Theta})^2} d\phi da, \quad (\text{B3}) \end{aligned}$$

where $\mu_{\Lambda}(a)$ is the conditional expected value of Λ given that A is equal to a , which is equal to $10 \log_{10}(10^{(L_0+\Delta L)/10} + 10^{(L_0+a)/10})$.

Further analytical manipulations of $f_{\Lambda, \Theta|L_{\text{target}}}$ do not appear to reduce the complexity of the solution, but $f_{\Lambda, \Theta|L_{\text{target}}}$ as represented by Eq. (B3) above can be numerically approximated. The first step of the numerical implementation is to approximate the definite integral over A through a finite summation. Let us denote $a[n]$ as a sampled version of the continuous random variable A . Further, let $a[1]$ equal a_{min} and $a[N]$ equal a_{max} . The probability density function $f_{\Lambda, \Theta|L_{\text{target}}}$ can then be numerically approximated as

$$\begin{aligned} f_{\Lambda, \Theta|L_{\text{target}}} &\approx \frac{1}{N(\phi_{\text{max}} - \phi_{\text{min}})2\pi\sigma_{\Theta}\sigma_{\Lambda}} \\ &\times \sum_{n=1}^N e^{-((\lambda - \mu_{\Lambda}(a[n]))^2/2\sigma_{\Lambda}^2)} \\ &\times \int_{\phi_{\text{min}}}^{\phi_{\text{max}}} e^{-(\phi - (\omega/k)\mu_{\Theta}(a[n]))^2/2((\omega/k)\sigma_{\Theta})^2} d\phi. \end{aligned}$$

Note that the integral of the exponential function can be represented as the error function so that the whole expression

is easily numerically evaluated. One should note that in the limit when $a_{\max} - a_{\min} = 0$ or $\phi_{\max} - \phi_{\min} = 0$, one cannot simply evaluate this expression at zero but must rather evaluate the expression in the limit as the difference approaches zero.

¹Although the first and last intervals convey no information for the ideal observer, these intervals may aid the nonideal subjects.

- Bernstein, J. G., and Oxenham, A. J. (2003). "Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?," *J. Acoust. Soc. Am.* **113**, 3323–3334.
- Bernstein, L. R. (2004). "Sensitivity to interaural intensive disparities: Listeners' use of potential cues," *J. Acoust. Soc. Am.* **115**, 3156–3160.
- Bernstein, L. R., and Trahiotis, C. (1982). "Detection of interaural delay in high frequency noise," *J. Acoust. Soc. Am.* **71**, 147–152.
- Bilsen, F. A. (1977). "Pitch of noise signals: Evidence for a 'Central spectrum,'" *J. Acoust. Soc. Am.* **61**, 150–161.
- Bilsen, F. A., and Raatgever, J. (2000). "On the dichotic pitch of simultaneously presented interaurally delayed white noises: Implications for binaural theory," *J. Acoust. Soc. Am.* **108**, 272–284.
- Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge).
- Bonnel, A. M., and Hafter, E. R. (1998). "Divided attention between simultaneous auditory and visual signals," *Percept. Psychophys.* **60**, 179–190.
- Brungart, D. S., and Simpson, B. D. (2002). "Within-ear and across-ear interference in a cocktail-party listening task," *J. Acoust. Soc. Am.* **112**, 2985–2995.
- Buus, S., and Florentine, M. (1991). "Psychometric functions for level discrimination," *J. Acoust. Soc. Am.* **90**, 1371–1380.
- Colburn, H. S., and Durlach, N. I. (1978). "Models of binaural interaction," in *Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), vol. IV.
- Durlach, N. I., and Colburn, H. S. (1978). "Binaural phenomena," in *Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), vol. IV.
- Green, D. M. (1988). *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, New York).
- Hafter, E. R. (1971). "Quantitative evaluation of a lateralization model of masking-level differences," *J. Acoust. Soc. Am.* **50**, 1116–1122.
- Hafter, E. R., and Jeffress, L. A. (1968). "Two-image lateralization of tones and clicks," *J. Acoust. Soc. Am.* **44**, 563–569.
- Hafter, E. R., and Carrier, S. C. (1970). "Masking-level differences obtained with a pulsed tonal masker," *J. Acoust. Soc. Am.* **47**, 1041–1047.
- Hartmann, W. M., and Constan, Z. A. (2002). "Interaural level differences and the level-meter model," *J. Acoust. Soc. Am.* **112**, 1037–1045.
- Heller, L. M., and Trahiotis, C. (1995). "The discrimination of samples of noise in monotic, diotic, and dichotic conditions," *J. Acoust. Soc. Am.* **97**, 3775–3781.
- Kidd, G. J., Mason, C. R., Arbogast, T. L., Brungart, D. S., and Simpson, B. D. (2003). "Informational masking caused by contralateral stimulation," *J. Acoust. Soc. Am.* **113**, 1594–1603.
- Koehnke, J., and Besing, J. M. (1992). "Effects of roving level variation on monaural detection with a contralateral cue," *J. Acoust. Soc. Am.* **92**, 2625–2629.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Mills, J. H., Dubno, J. R., and He, N.-J. (1996). "Masking by ipsilateral and contralateral maskers," *J. Acoust. Soc. Am.* **100**, 3336–3344.
- Rowland, R. C. J., and Tobias, J. V. (1967). "Interaural intensity difference limen," *J. Speech Hear. Res.* **10**, 745–756.
- Ruoto, B. R., Stern, R. M. J., and Colburn, H. S. (1979). "Discrimination of symmetric time-intensity traded binaural stimuli," *J. Acoust. Soc. Am.* **66**, 1733–1737.
- Sivonen, V. P., and Ellermeier, W. (2006). "Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation," *J. Acoust. Soc. Am.* **119**, 2965–2980.
- Stern, R. M., and Trahiotis, C. (1997). "Models of binaural perception," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey, and T. R. Anderson (Erlbaum, Mahwah).
- Taylor, M. M., and Clarke, D. P. J. (1971). "Monaural detection with contralateral cue (MDCC). II. Interaural delay of cue and signal," *J. Acoust. Soc. Am.* **49**, 1243–1253.
- Taylor, M. M., Clarke, D. P. J., and Smith, S. M. (1971a). "Monaural detection with contralateral cue (MDCC). III. Sinusoidal signals at a constant performance level," *J. Acoust. Soc. Am.* **49**, 1795–1804.
- Taylor, M. M., Smith, S. M., and Clarke, D. P. (1971b). "Monaural detection with contralateral cue (MDCC). IV. Psychometric functions with sinusoidal signals," *J. Acoust. Soc. Am.* **50**, 1151–1161.
- Trahiotis, C. (1992). "Developmental considerations in binaural hearing experiments," in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (American Psychological Association, Washington, DC), vol. 1.
- Viemeister, N. F. (1988). "Psychophysical aspects of auditory intensity coding," in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (J. Wiley, New York).
- Whilby, S., Florentine, M., Wagner, E., and Marozeau, J. (2006). "Monaural and binaural loudness of 5- and 200-ms tones in normal and impaired hearing," *J. Acoust. Soc. Am.* **119**, 3931–3939.
- Yost, W. A. (1972). "Tone-on-tone masking for three binaural listening conditions," *J. Acoust. Soc. Am.* **52**, 1234–1237.
- Yost, W. A., Penner, M. J., and Feth, L. L. (1972). "Signal detection as a function of contralateral sinusoid-to-noise ratio," *J. Acoust. Soc. Am.* **51**, 1966–1970.
- Zurek, P. M. (1979). "Measurements of binaural echo suppression," *J. Acoust. Soc. Am.* **66**, 1750–1757.
- Zwicker, E., and Zwicker, U. T. (1991). "Dependence of binaural loudness summation on interaural level differences, spectral distribution, and temporal distribution," *J. Acoust. Soc. Am.* **89**, 756–764.
- Zwislocki, J. J. (1972). "A theory of central auditory masking and its partial validation," *J. Acoust. Soc. Am.* **52**, 644–659.