Chapter 24

Models of Binaural Perception

Richard M. Stern
Carnegie Mellon University

Constantine Trahiotis
University of Connecticut Health Center

(Received March 1995; revised July 1995)

We review some of the major trends in binaural modeling, particularly with regard to models based on the interaural cross-correlation of the auditory-nerve responses to the stimuli. Emphasis is placed on providing an intuitive understanding of cross-correlation-based binaural models, combined with an appreciation of their capabilities and limitations in describing a variety of binaural phenomena. We focus on the seminal theory of binaural processing by Jeffress and its later elaboration and quantification by Colburn. This theory describes and predicts binaural phenomena in terms of the putative activity of central units that record interaural coincidences of firing from matched pairs of auditory-nerve fibers (one from each ear). The input auditory-nerve fibers are matched in characteristic frequency with a fixed time delay inserted on one side. The response of a number of such central units at a given characteristic frequency, plotted as a function of internal delay, is an approximation to the interaural cross-correlation function of the sound as processed by the auditory periphery. We discuss predictions for many of the simple and complex binaural stimuli that are commonly used in psychoacoustical experiments. These experiments include measurements of subjective lateral position, interaural discrimination, binaural detection, and dichotic pitch.

INTRODUCTION

In this chapter we review some of the major trends of research in binaural modeling, particularly with regard to models based on the interaural cross-correlation of the auditory-nerve response to the stimuli.

The human binaural system has attracted the attention of auditory theorists since Lord Rayleigh's seminal investigations (Rayleigh, 1907). The "modern era" of binaural modeling can be said to have begun in 1948 with Jeffress's prescient paper suggesting a neural coincidence mechanism to detect interaural time differences. During that same year the original descriptions of the binaural masking level difference were provided independently by Hirsh (1948) and Licklider (1948).

A convenient starting point for this discussion is the classic review chapter of binaural models by Colburn and Durlach (1978). They described in detail models based on explicit detection of interaural differences (e.g., Jeffress, Blodgett, Sandel, and Wood, 1956; Hafter and Carrier, 1970), models based on direct comparison of the amount of the "left-sided" and "right-sided" internal response to stimuli (e.g., van Bergeijk, 1962), models based on cancellation of binaural maskers (e.g., Durlach, 1972), models based on the direct cross-correlation of the stimuli (e.g., Sayers and Cherry, 1957; Osman, 1971), and models that perform interaural comparisons of explicit descriptions of auditory-nerve activity (e.g., Colburn, 1973, 1977). These models were developed primarily to describe the results of experiments measuring binaural masking level differences (BMLDs). Domnitz and Colburn (1976) later demonstrated that, despite their apparent differences in structure, most of these models provide similar predictions for BMLDs measured with diotic maskers. In their summary, Colburn and Durlach (1978) noted that most of the contemporary models could be considered to be different implementations of the general structure shown in Fig. 1. This generic structure includes a series of peripheral processing steps including bandpass filtering and rectification, comparison of interaural timing information over a limited range of internal delays using a correlation or coincidence mechanism, consideration of interaural intensity differences of the outputs of monaural processors, and a subsequent decision-making mechanism.

Since 1978, the basic structure described by Colburn and Durlach, and especially the cross-correlation mechanism used for the extraction of interaural timing information, has formed the basis of all subsequent models of binaural hearing. Other recent trends fostering the development of a broader theory of binaural perception include an increased reliance on computational (as opposed to analytical) approaches to predicting the phenomena, and efforts to make use of head-related transfer functions in attempts to understand the relative salience of the different available cues and to mimic realistic sound fields using stimuli presented through headphones. At the same time, there has been increased attention paid to the development of models of more central physiological processing mechanisms, which may bear directly on our understanding of binaural hearing. This work has recently been summarized by Colburn (1995).

The goal of this chapter is to provide the reader with an intuitive understanding of how cross-correlation-based binaural models work, and to provide an appreciation of their capabilities and limitations in describing a variety of binaural phenomena. In contrast to other recent reviews of binaural modeling (e.g., Colburn, 1995; Stern and Trahiotis, 1995), this chapter is less comprehensive and more centered around our own efforts. Using this "narrow" approach, we show

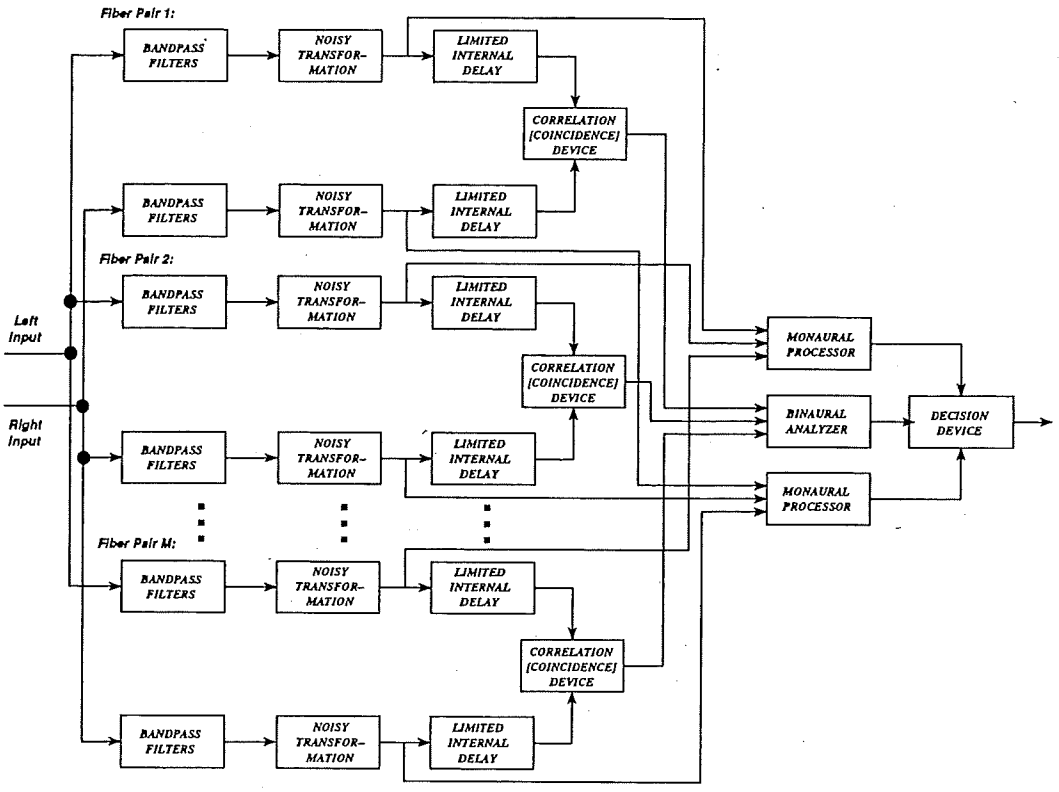

FIG. 1. Generic model of binaural processing proposed by Colburn and Durlach (1978). Three of many sets of fiber pairs are depicted.

how the same theoretical framework can be used to provide predictions spanning a wide variety of phenomena and empirical data. It should be understood that similar predictions would be provided by other cross-correlation-based models as well.

We begin by reviewing and discussing the Jeffress–Colburn model and selected extensions in Sections I and II. The response of the Jeffress–Colburn model to various kinds of simple stimuli is characterized in Section III, which also contains representative comparisons of the predictions of this kind of model to the corresponding experimental data.

I. CROSS-CORRELATION-BASED MODELS OF BINAURAL INTERACTION

A. The Jeffress–Colburn model

Modern binaural models are based on Jeffress's (1948) conception of a neural "place" mechanism that would enable the extraction of interaural timing information. Jeffress suggested that external interaural delays could be internally coded by central units that record coincidences of neural impulses from pairs of more peripheral nerve fibers. Each central unit was presumed to compare information from the two ears after a series of internal time delays. Licklider

(1959) later proposed a similar mechanism that could also be used to achieve an autocorrelation of neural signals for use in models of pitch perception. Jeffress's hypothesis was reformulated in a more quantitative form by Colburn (1973, 1977). Colburn's model consists of two parts: a characterization of auditory-nerve activity, and a central processor that analyzes and displays comparisons of neural activity from the two ears.

1. The model of auditory-nerve activity

The model of auditory-nerve activity used in the original Colburn model was adapted from an earlier formulation of Siebert (1970) and is depicted in the upper panel of Fig. 2. It consists of a bandpass filter (to depict the frequency selectivity of individual fibers), an automatic gain control (which limits the average rate of response to stimuli), a lowpass filter (which serves to limit phase-locking to stimulus fine structure at higher frequencies), and an exponential rectifier (which roughly characterizes peripheral nonlinearities). These elements were followed by a mechanism that generates neural impulses at an average rate that is proportional to the output of the rectifier and with temporal characteristics consistent with activity produced by a nonhomogeneous Poisson process. Predictions for this chapter follow the more recent formulation of Stern and Shear (Shear, 1987; Stern and Shear, 1996). They changed the shape of the nonlinear rectifier and interchanged the order of the rectifier and the lowpass filter in order to describe more accurately the response to noise stimuli and to high-frequency stimuli. Similar functional models have been used by others including Duifhuis (1973), Blauert and Cobben (1978), and Lindemann (1986a, 1986b).

Colburn used the nonhomogeneous Poisson process to characterize the response of auditory-nerve fibers to sound because it is the simplest stochastic process that can realistically be applied to model the neural firing times. Using an

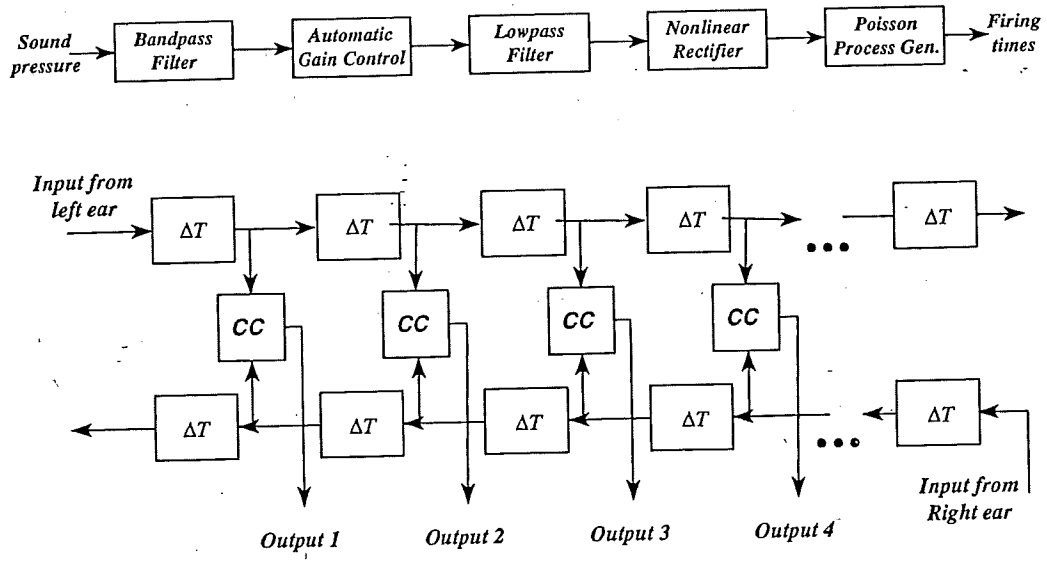

FIG. 2. Upper panel: Schematic representation of the auditory periphery. Lower panel: Schematic representation of the Jeffress place mechanism. The blocks labeled CC record coincidences of neural activity from the two ears (after the delays are incurred).

explicit analytical model like the Poisson process, one can predict discrimination and detection thresholds by application of the Cramer–Rao bound (cf. van Trees, 1968), or one can calculate means and variances of the predicted outputs of the coincidence counters and directly predict performance by assuming that the decision variable is normally distributed. The construction and evaluation of analytical models of neural activity is inevitably a compromise between faithfulness to the known physiological results and mathematical tractability. For example, it is well known that the peripheral auditory system is both time varying (e.g., due to the refractory nature of the auditory response) and nonlinear. The Poisson-process model ignores the refractoriness in the response, and the nonlinear rectifier in the model does not describe several known aspects of peripheral auditory nonlinearity. Furthermore, predictions are easily developed only for exponential and half-wave power-law rectification and only for a small set of stimuli (e.g., pure tones, tones in noise, and bandpass noise) with fixed interaural time delays (ITDs) and interaural intensity differences (IIDs).

In recent years, models of the peripheral auditory response to sound have become more computationally oriented (and more physiologically accurate) (e.g., Carney, 1993; Meddis, Hewitt, and Shackleton, 1990; Payton, 1988). These models have in turn served as front ends for binaural models. For example, the hair cell model of Meddis *et al.* has been incorporated into the binaural processing model of Shackleton, Meddis, and Hewitt (1992).

2. The model of central processing

One formulation of Colburn's quantification of Jeffress's hypothesis is depicted in the lower panel of Fig. 2. The heart of Colburn's (1973, 1977) model is an ensemble of units describing the interaction of neural activity from the left and right ears generated by auditory-nerve fibers with the same characteristic frequency (CF). The input from one side is delayed by an amount that is fixed for each fiber pair. The delay mechanism is commonly conceptualized in the form of a ladder-type delay line as in Fig. 2, but such a structure is not the only possible realization. The net interaural delay incurred by the two inputs to each fiber pair is the key parameter in the analysis of the outputs of the mechanism and is referred to using the variable τ. This ensemble of coincidence-counting units is similar in structure to the central processor of several other models including that of Blauert and Cobben (1978).

The relative number of coincidence counts of the Jeffress–Colburn model, considered as a function of the internal-delay parameter τ, is an estimate of the interaural cross-correlation of the auditory-nerve responses to the stimuli at each CF. In contrast, some previous models of binaural processing utilized the cross-correlation of the original stimuli (rather than the physiological response to the stimuli) to develop predictions (e.g., Sayers and Cherry, 1957; Osman, 1971).

Colburn and Durlach (1978) noted that the cross-correlation mechanism shown in Fig. 2 can also be regarded as a generalization of the equalization-cancellation (EC) model of Durlach (1963). Specifically, the EC model yields

predictions concerning binaural detection thresholds by applying a combination of ITD and IID that produces the best "equalization" of the masker components of the stimuli presented to each of the two ears. "Cancellation" of the masker is then achieved by subtracting one of the resulting signals from the other. Predictions provided by the EC model are generally dominated by the effects of the ITD-equalization component rather than the IID-equalization component. Because the interaural delays of the fiber pairs of the Jeffress–Colburn model perform the same function as the ITD-equalizing operation of the EC model, most predictions of detection thresholds for the two models are similar.

B. Physiological plausibility of the Jeffress–Colburn model

As summarized by Kuwada, Batra, and Fitzpatrick (Chapter 20, this volume) and Yin, Joris, Smith, and Chan (Chapter 21, this volume), a number of researchers have studied neural cells that have outputs that functionally resemble those of the coincidence-counting units schematized in Fig. 2. Particularly noteworthy are the cells having a "characteristic delay" first reported by Rose, Gross, Geisler, and Hind (1966) in the inferior colliculus. Such cells are maximally sensitive to inputs that have a specific interaural delay regardless of the frequency of the stimulation. Cells with similar responses have been reported by other researchers in other sites within the central auditory system.

The anatomical origin of the internal interaural delays has been the source of some speculation. The preponderance of evidence indicates that the delays are of neural origin, caused either by slowed conduction velocity or by synaptic delays (e.g., Smith, Joris, and Yin, 1993; Carr and Konishi, 1990; Young and Rubel, 1983). It has also been suggested (without evidence) that the internal delays could come about if higher processing centers were to compare timing information derived from auditory-nerve fibers with different CFs (Schroeder, 1977; Shamma, Shen, and Gopalaswamy, 1989). The anatomical validity of the models notwithstanding, the predictions of binaural models are unaffected by whether the internal delays are assumed to be caused by neural or mechanical phenomena.

C. Temporal integration of the coincidence display

Although the binaural system is known to resolve static ITDs as small as tens of microseconds, experiments measuring responses to time-varying ITDs (e.g., Licklider, Webster, and Hedlun, 1950; Grantham and Wightman, 1978) indicate a lower degree of temporal resolution, on the order of tens of milliseconds. For this reason, the binaural system is often characterized as being "sluggish." In order to understand threshold sensitivity to stimuli with either static or time-varying ITDs, one must note that discrimination between static ITDs reflects changes in the place of activity of the coincidence-counting units along the internal-delay axis. Such resolution is limited by the density of fiber pairs with respect to internal delay at each CF. On the other hand, resolution of time-varying ITDs reflects the averaging of instantaneous responses over running time (averaged across internal

delay). This type of averaging is often referred to as temporal integration and appears to be performed rather slowly.

It is helpful to think of the temporal averaging of the matrix of coincidence-counting units as resulting from a lowpass filtering of the instantaneous outputs of the coincidence counters with respect to running time. Figure 3 demonstrates how the expected number of instantaneous coincidences for fibers with a CF of 500 Hz varies as a simultaneous function of internal delay (τ) and running time (t), both with and without (running) lowpass filtering. The upper panel of Fig. 3 shows the instantaneous response of the coincidence counters depicted in Fig. 2 to a 500-Hz tone with zero ITD. Note that the peaks of activity are limited to particular intervals of the running time as well as to particular values of interaural delay. Said differently, there are areas of inactivity along both axes that reflect times for which the correlation function approaches zero. The lower panel of Fig. 3 shows the same function after temporal integration, realized by convolution with a simple lowpass filter. Note that the integration with respect to running time transforms the isolated peaks in the response of the coincidence counters to smoother ridges that are parallel to the running-time axis. We believe that such smoothing is necessary because it allows the binaural system to provide a stable spatial representation of the acoustic world.

It appears that many, if not all, of the data concerning the sluggishness phenomenon can be explained in terms of simple temporal integration of the coincidence-counter output (Grantham and Wightman, 1978; Bachorski, 1983;

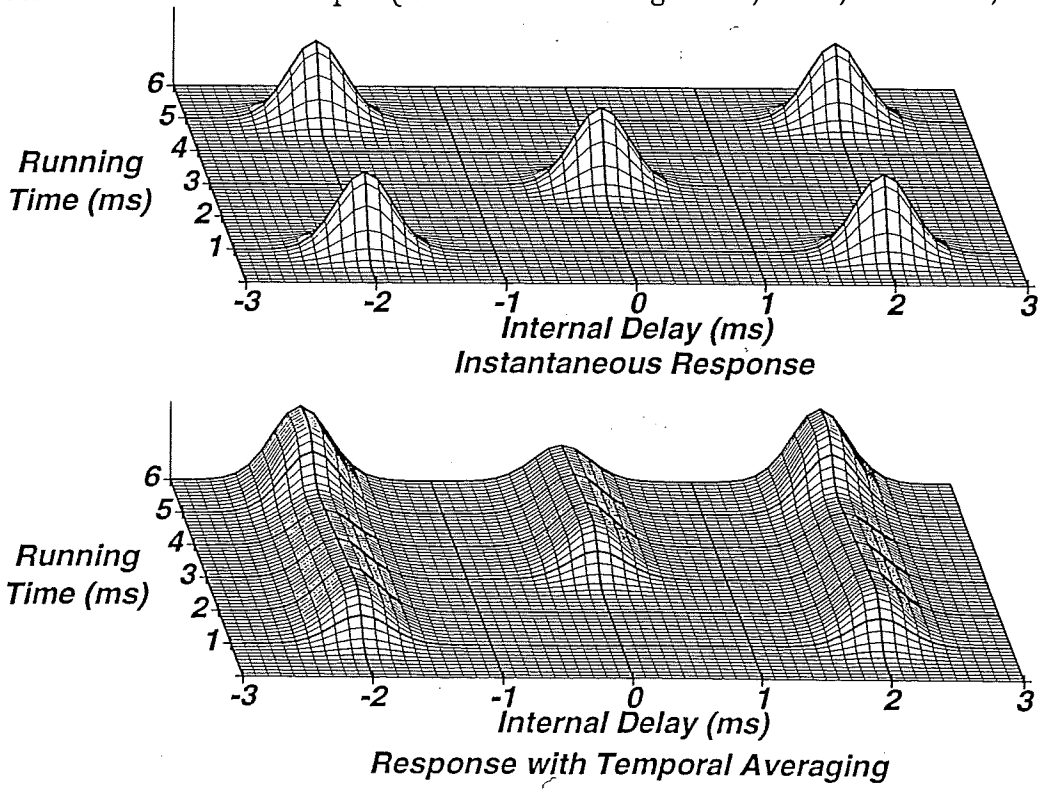

FIG. 3. Average value of the instantaneous number of coincidences as a simultaneous function of running time t and internal delay τ. The stimulus is a 500-Hz tone with zero ITD. The response is shown using no temporal integration (upper panel) and using temporal integration by an exponentially shaped temporal weighting function with an effective cutoff frequency of 5 Hz in the frequency domain.

Stern and Bachorski, 1983; Grantham, 1984). Various experimental results imply that the time constants for processing IIDs are much shorter than those that constrain the processing of ITDs (Grantham, 1984; Bernstein and Trahiotis, 1994). The necessity for including more than one time constant was also emphasized by Gabriel (1983). The type of temporal averaging that is likely to mediate binaural sluggishness also provides at least a qualitative explanation for the disappearance of binaural beats at high beat frequencies (Licklider *et al.*, 1950).

II. EXTENSIONS TO THE JEFFRESS–COLBURN MODEL

A. Extensions by Stern, Colburn, and Trahiotis

The original goal of work performed by Stern and his colleagues was to extend the Jeffress–Colburn formulation to provide predictions for the subjective lateral position of stimuli and to examine the extent to which the "position variable" could be used to describe detection and discrimination results. In order to do this, it was necessary to specify a means of combining effects produced by the ITDs and IIDs of the stimulus and to provide a way to predict subjective lateral position from the combined display. The resulting model is referred to as the position-variable model.

1. Combination of differences of interaural time and intensity

Because cross-correlation is a multiplicative operation, cross-correlation functions cannot be used to indicate which ear is receiving the more intense input. Hence, additional mechanisms are needed to describe how IIDs affect subjective lateral position. At one time it was felt that the effects of IIDs in binaural lateralization could be accounted for by the decrease in latency of the auditory-nerve response that occurs as the intensity of the signals is increased. This peripheral time–intensity trading mechanism, known as the latency hypothesis, was discussed by Jeffress in 1948 and later elaborated by David, Guttman, and van Bergeijk (1958) and Deatherage and Hirsh (1959). Although this hypothesis was at least qualitatively supported by early lateralization studies that utilized small ITDs and IIDs, it cannot describe lateralization data in which subjective lateral position is shown clearly to be a nonlinear function of ITD and IID when these two stimulus parameters are varied over a wider range of conditions (e.g., Sayers, 1964; Domnitz and Colburn, 1977; Bernstein and Trahiotis, 1985), as discussed in Sec. II.A.2. The latency hypothesis is also contradicted by the results of several interaural discrimination studies that indicate an inability to "trade" time and intensity differences completely. For example, Hafter and Carrier (1972) demonstrated that subjects could always discriminate between diotic 500-Hz tones and dichotic tones presented with a canceling combination of ITDs and IIDs that produced a centered primary image.

The position-variable model (Stern and Colburn, 1978) incorporates a more central mechanism to account for the effects of IID. The function describing the

number of coincidences as a function of internal delay is multiplied by a Gaussian-shaped function with a location along the internal-delay axis that depends on the IID of the stimulus. In Sec. II.B we describe an alternative intensity-weighting mechanism that was proposed by Lindemann (1986a), which incorporates lateral inhibition of the coincidence-counting response along adjacent delays. The Stern–Colburn and Lindemann models provide similar predictions for the lateralization of 500-Hz pure tones as a joint function of ITD and IID.

2. Lateral position predictions using the coincidence display

There are several ways of predicting lateral position from the outputs of the interaural coincidence-counting units. Stern and Colburn (1978) proposed that the predicted lateral position of a stimulus, $\hat{P}$, can be obtained by computing the centroid (or center of mass) along the internal-delay axis of the intensity-weighted function describing the number of coincidences, while integrating over frequency. This definition of predicted lateral position was originally adopted by Stern and Colburn for reasons of computational simplicity, and it has been employed by Blauert and his colleagues (e.g., Lindemann, 1986a, 1986b) as well. It should be noted, however, that the centroid computation by itself produces predictions for the intracranial location of only a single image. As a result, using the centroid of activity alone one cannot explain experimental results that suggest the existence of multiple images such as the studies by Moushegian and Jeffress (1959), Whitworth and Jeffress (1961), and Hafter and Jeffress (1968).

One plausible alternative is to predict lateral position by resorting to the locations of individual peaks of the cross-correlation function. Such locations allow one to account for multiple images that can occur for tonal stimuli presented interaurally out of phase (e.g., Sayers, 1964; Yost, 1981), as well as for the secondary "time image" observed when some stimuli are presented with conflicting ITDs and IIDs (e.g., Whitworth and Jeffress, 1961; Hafter and Jeffress, 1968). Shackleton *et al.* (1992) made predictions on the basis of either the centroid or the peaks of the responses, choosing the statistic that more accurately described the results for a given experiment. Although definitely not parsimonious, this type of approach may be necessary to account for the varieties of data in all their complexity.

The function specifying the density of internal delays along the internal-delay axis plays an important (but frequently unrecognized) role in developing predictions of subjective lateral position. The form of the function $p(\tau)$ derived by Colburn (1977) and later modified by Stern and Shear (1996) specifies that there are more coincidence-counting units with internal interaural delays of smaller magnitude. This has been verified by physiological measurements (e.g., Kuwada, Stanford, and Batra, 1987). Nevertheless, in order to describe many of the detection and lateralization data, a substantial fraction of the coincidence counters must be assumed to have internal delays much greater in magnitude than the largest delays that can be physically attained using free-field stimuli.

Colburn (1969, 1977) originally assumed that the density function for internal delays, $p(\tau)$, was independent of frequency, and he fitted the shape of $p(\tau)$ to

predict the relative masking level differences for two antiphasic conditions, N_0S_π versus $N_\pi S_0$. More recently, Stern and Shear (1996) made this function weakly dependent on frequency and changed its shape slightly. This allowed them to predict the lateralization of tonal stimuli with a fixed ITD as a function of stimulus frequency (Schiano, Trahiotis, and Bernstein, 1986).

The effects of the distribution of internal delay and the multiplicative intensity weighting function are illustrated in Figs. 4 and 5, which depict the representation of a 500-Hz pure tone with an ITD of +0.5 ms. Figure 4 shows the total number of coincidences recorded by the coincidence-counting units as a joint function of internal delay (along the horizontal axis) and CF (along the oblique axis). The upper panel shows the average number of coincidences per fiber pair. The center

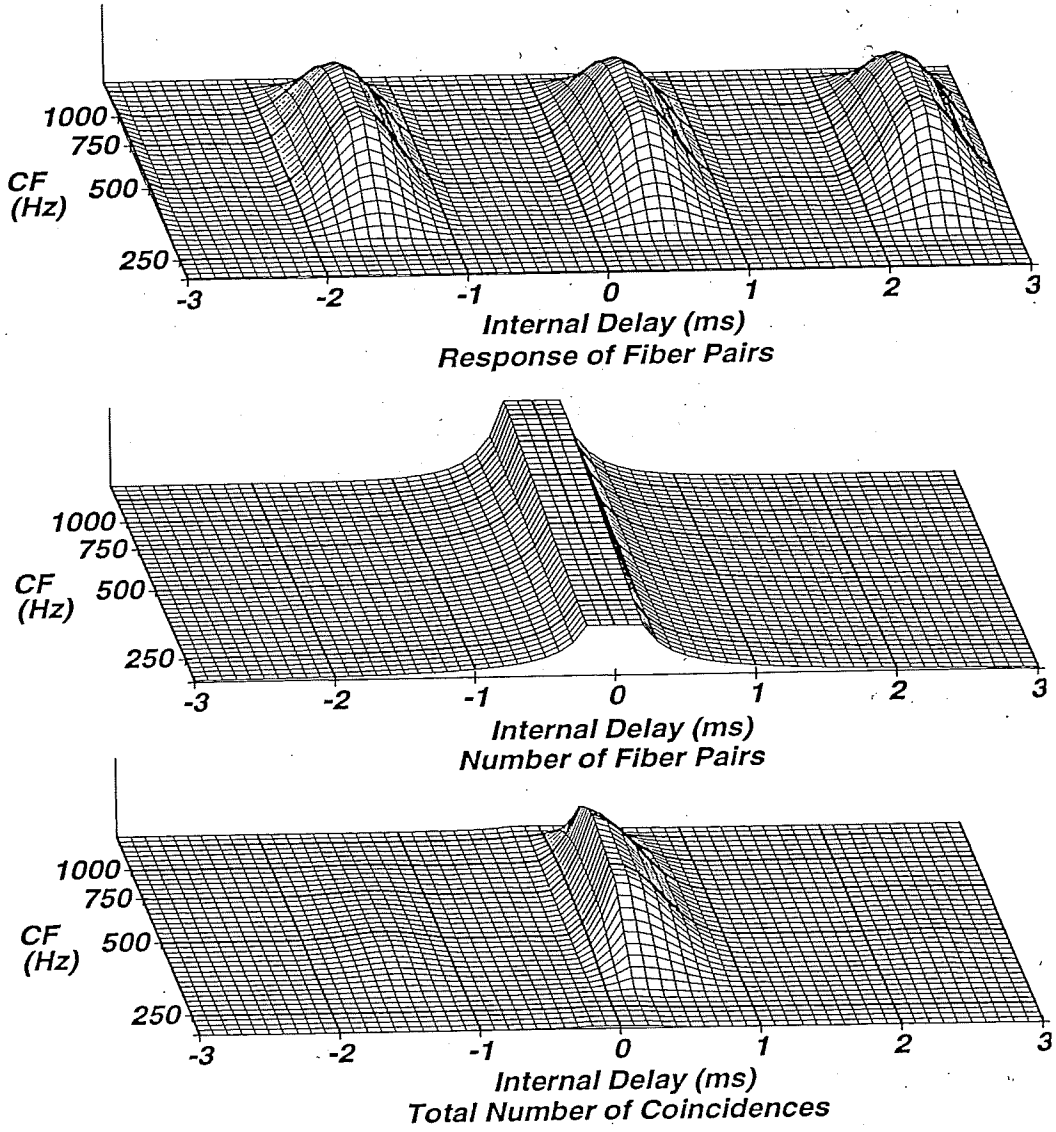

FIG. 4. Response of an ensemble of binaural coincidence-counting units to a 500-Hz pure tone with a 0.5-ms ITD. Upper panel: The relative number of coincidences per fiber pair as a function of internal delay τ (ms) and CF of the auditory-nerve fibers (Hz). Central panel: The assumed density of internal delays as a function of CF. Lower panel: The average total number of coincidences as a function of internal delay and CF, which is the product of the upper and central panel.

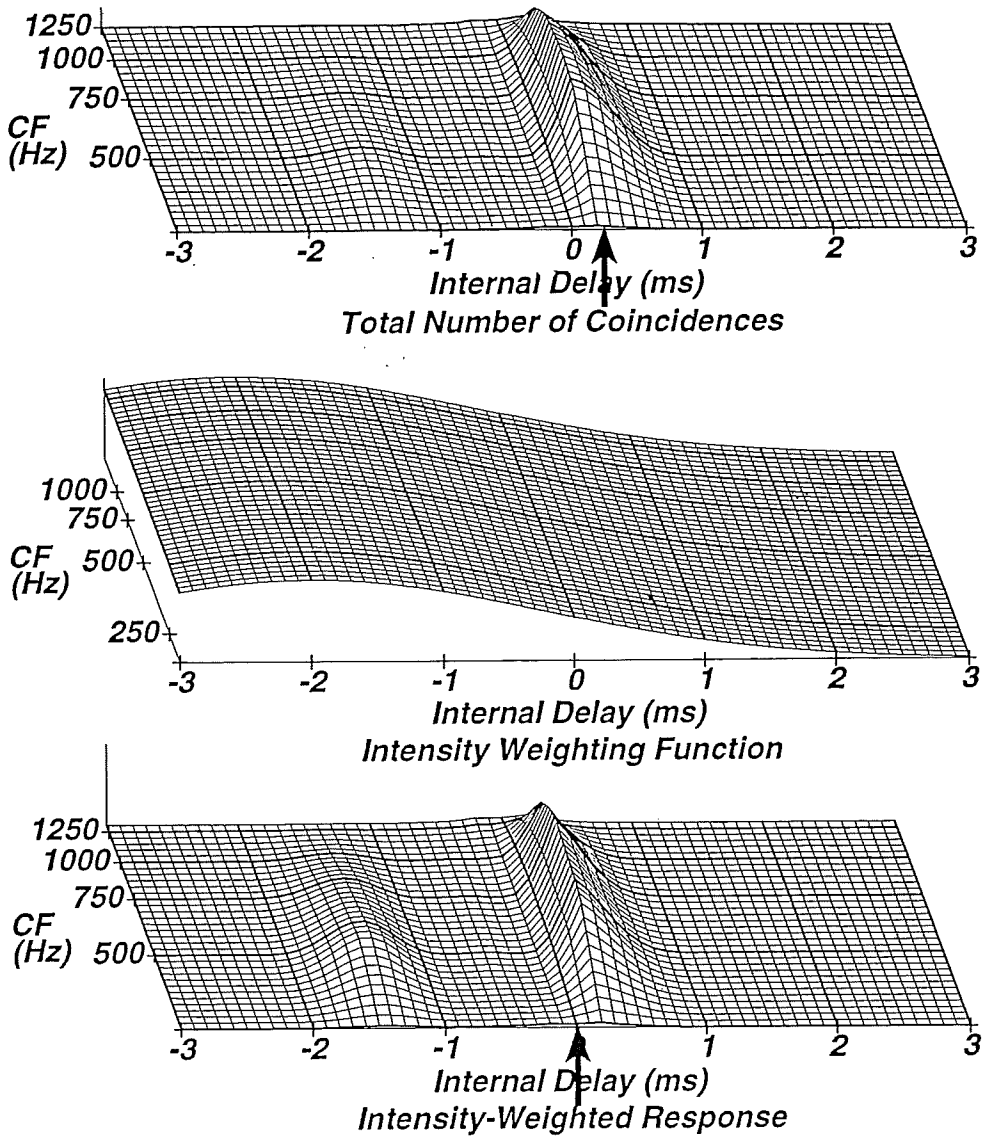

FIG. 5. Response patterns showing how the effects of intensity weighting are incorporated into the model of Stern and Colburn (1978). Upper panel: The average total number of coincidences as a function of internal delay and CF, similar to the lower panel of Fig. 4. Central panel: The Gaussian-shaped intensity weighting function. The function shown corresponds to an IID of about –7 dB. Lower panel: The effective average number of coincidences after intensity weighting, which is the product of the upper and central panel.

panel shows $p(\tau, f)$, the function that describes the density of fiber pairs as a function of internal delay and CF. The lower panel displays the total number of coincidences at each internal delay and CF. That total is the product of the number of counts per fiber pair (upper panel) and the number of fiber pairs (central panel). There is a distinct maximum in the cross-correlation function at a value of internal delay that is close to that of the original interaural delay of the stimulus and extends over a broad range of CFs.

Figure 5 demonstrates the effects of the intensity-weighting mechanism for a set of stimuli that produce time–intensity "trading." The upper panel of Fig. 5 is similar to the lower panel of Fig. 4. It depicts the total number of coincidences in response to a 500-Hz tone with a 0.5-ms ITD, after accounting for the relative

number of fiber pairs at each internal delay. The central panel of Fig. 5 shows the Gaussian-shaped intensity-weighting function, which has a location along the τ axis that depends on the IID. The Gaussian pulse in Fig. 5 is centered at –2.0 ms, which corresponds to an IID of approximately –7 dB. Although the intensity-weighting function is rather broad, its location along the τ axis has considerable effect on the predicted lateral position. The lower panel of Fig. 5 is the product of the two upper panels and shows the effects of intensity weighting on the outputs of the coincidence counters. The vertical arrows in the upper and lower panels indicate the location of the centroid of activity along the τ axis, without and with the intensity weighting. Note that moving the intensity weighting function toward the opposite ear results in a shift of the resulting activity toward that ear. This outcome provides predictions consistent with data obtained in experiments utilizing conflicting ITDs and IIDs, such as those employed in "time–intensity trading" experiments.

B. Extensions by Blauert, Cobben, Lindemann, and Gaik

Blauert and his colleagues made important contributions to correlation-based models of binaural hearing over an extended period of time. Their efforts have been primarily directed toward understanding how the binaural system processes more complex sounds in real rooms and have tended to be computationally oriented. This approach is complementary to that of Colburn and his colleagues, who have focused on explaining "classical" psychoacoustical phenomena using stimuli presented through earphones. In recent years Blauert and his colleagues have been applying knowledge gleaned from fundamental research in binaural hearing toward the development of improved devices that enhance the spatiality of recorded sound, as described by Blauert (Chapter 28, this volume).

One of the most interesting models emerging from Blauert's laboratory is the one proposed by Lindemann (1986a), which may be regarded as an extension and elaboration of an earlier hypothesis of Blauert (1980). Lindemann extended the original Jeffress coincidence-counter model in two ways. He included (1) inhibition of outputs of the coincidence counters when there is activity produced by coincidence counters at adjacent internal delays, and (2) monaural-processing mechanisms at the "edges" of the display of coincidence-counter outputs that become important when the stimulus contains a large IID. Lindemann's inhibitory mechanism produces a "sharpening" of the peaks of the outputs of the coincidence counters along the internal-delay axis.

One of the very interesting properties of the Lindemann model is that it produces a time–intensity trading mechanism at the level of the coincidence-counter outputs. This occurs because the interaction of the inhibitory mechanism and the monaural processing mechanisms causes the locations of peaks of the outputs of the coincidence counters to shift along the internal-delay axis with changes in IID. The net effects of IIDs on the patterns of coincidence-counter outputs in the Lindemann model are not unlike effects produced by the intensity-weighting function used by Stern and Colburn (1978). In a sense, the time–intensity interaction of the Lindemann model is more parsimonious in that

it arises naturally from the fundamental assumptions of the model rather than as the result of the imposition of an arbitrary weighting function.

Gaik (1993) extended the Lindemann mechanism by adding a further weighting to the coincidence-counter outputs that reinforces naturally occurring combinations of ITD and IID. This has the effect of causing physically plausible stimuli to produce coincidence outputs with a single prominent peak that is compact along the internal-delay axis and that is consistent over frequency. Conversely, very unnatural combinations of ITDs and IIDs presented via earphones (which tend to give rise to multiple and/or diffuse perceptual images) produce response patterns with more than one prominent peak along the internal-delay axis.

III. COMPARISONS OF THEORETICAL PREDICTIONS TO EXPERIMENTAL DATA

In this section we describe how the patterns of outputs of the coincidence counters of the Jeffress–Colburn model have been applied to describe some of the phenomena that have been important for researchers in auditory perception. In order to make the presentation easy to follow, the discussion includes both examples of responses of the coincidence counters and predictions obtained using those responses. Simultaneously, we comment on the characteristics and limitations of current models.

A. Subjective lateral position

1. Lateralization of pure tones

Figure 6 compares the predictions of the original position-variable model for the lateral position of 500-Hz tones as a function of ITD and IID (Stern and Colburn, 1978) to data obtained by Domnitz and Colburn (1977). These predictions were obtained by computing the centroid along the internal-delay axis of the intensity-weighted coincidence counts, as exemplified in the lower panel of Fig. 5. The model provides reasonably accurate predictions for a number of fundamental aspects of the lateralization of pure tones based on ongoing ITDs and IIDs. These aspects include (1) the periodicity of lateral position with respect to ITD; (2) the joint dependence of the lateralization of low-frequency pure tones on ITD and IID; and (3) the "cue-reversal phenomenon" wherein the direction of apparent motion of the image reverses at ITDs approaching half the period of the tone. Note that as IID increases, the forms of the curves for both the data and predictions are inconsistent with the latency hypothesis. This is indicated by the fact that the effect of IID is to displace the curves vertically and horizontally, rather than just horizontally.

The model was subsequently modified to allow the position variable to be a function of time (Stern and Bachorski, 1983). This enables it to account for cases where stimuli are presented with slowly varying ITD and/or IID (e.g., Grantham and Wightman, 1978; Licklider *et al.*, 1950). In addition, as noted earlier, it was also necessary to modify the $p(\tau)$ function in order to account for the fact that

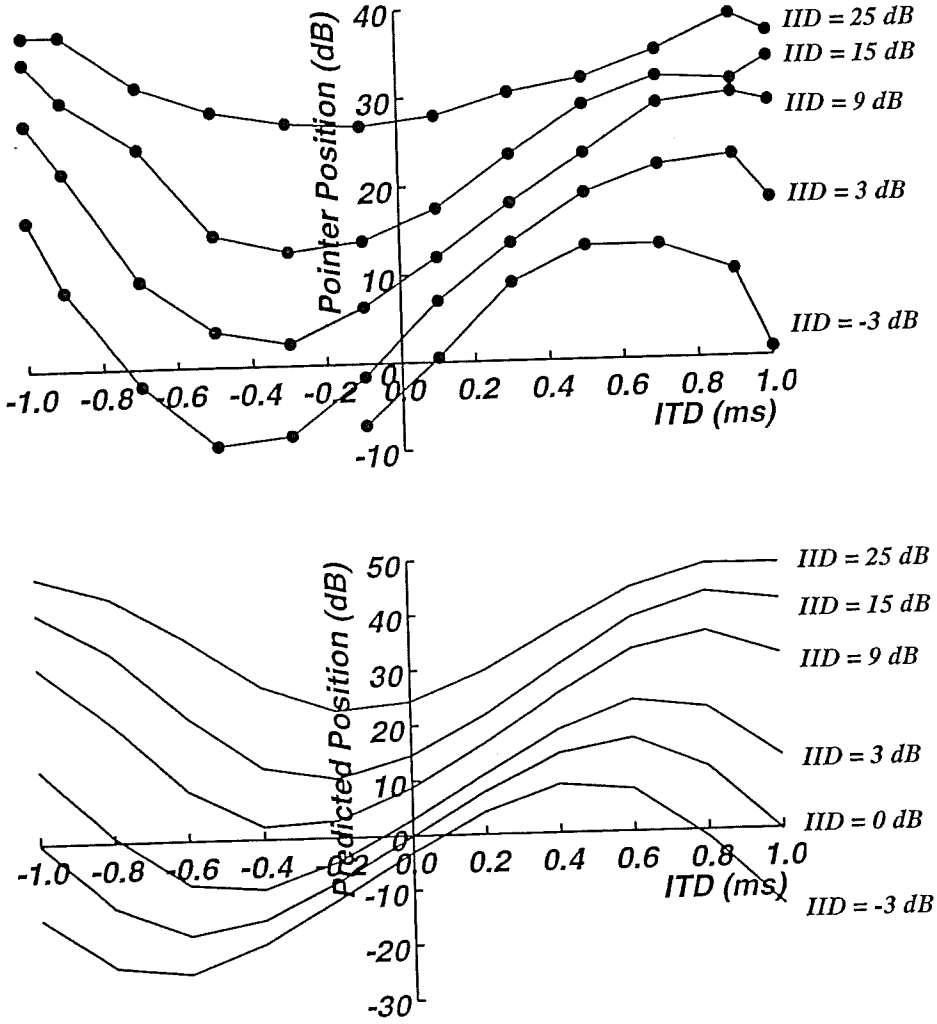

FIG. 6. Upper panel: Experimental lateralization-matching results, showing the IID of a pointer tone required to match the perceived lateral position of a 500-Hz test tone, as a function of the ITD and IID of the test tone (Domnitz and Colburn, 1977). Lower panel: Theoretical predictions for the same stimuli, from Stern and Colburn (1978).

the lateral position of pure tones with fixed ITD is approximately constant up to 1000 Hz (Schiano *et al.*, 1986).

2. Lateralization of low-frequency bandpass noise

In recent years attention has been focussed on the lateralization of spectrally and temporally complex stimuli including bandpass noise and amplitude-modulated tones. The lateralization of bandpass noise and amplitude-modulated tones are treated separately because different issues arise in understanding how they are lateralized.

Figure 7 shows the responses of the coincidence-counting units to bandpass noise presented with a center frequency of 500 Hz and two different bandwidths, 50 Hz (upper panel) and 800 Hz (lower panel). In both cases the stimuli have an ITD of −1.5 ms. The pattern of the responses for the 50-Hz-wide noise looks very

similar to the pattern produced by 500-Hz pure tones presented with the same ITD. For 500-Hz tones, an ITD of −1.5 ms is equivalent to an ITD of +0.5 ms, as shown in the upper panel of Fig. 4. Consistent with this, the stimulus with the 50-Hz bandwidth is lateralized on the "wrong" side of the head, that is, the side receiving the signal that is lagging in time. For larger bandwidths, the intracranial image moves toward the left side of the head (Stern, Zeiberg, and Trahiotis, 1988; Trahiotis and Stern, 1989), indicating the true ITD. With larger bandwidths, as exemplified by the 800-Hz-wide condition in the lower panel of Fig. 7, it is obvious that the ridge at $\tau = -1.5$ corresponds to the true ITD because it is parallel to the CF axis, depicting a consistent stimulus ITD of −1.5 ms for all frequencies.

We have referred to the consistency over frequency of the maxima of the coincidence-count response (that indicates the true ITD) as *straightness*. By independently manipulating ITD, interaural phase difference (IPD), and bandwidth, it became clear that the binaural system weights more heavily the straighter components of the response to bandpass-noise stimuli.

We believe that the straightness-weighting phenomenon results from passing the outputs of the coincidence-counting units through a second level of coincidence-counting units. Each set of inputs to this second layer of temporal processing is assumed to come from first-level coincidence counters representing a range of CFs, but with a common internal delay. The effect of this type of processing is illustrated in Fig. 8, which compares the response of the original model (without

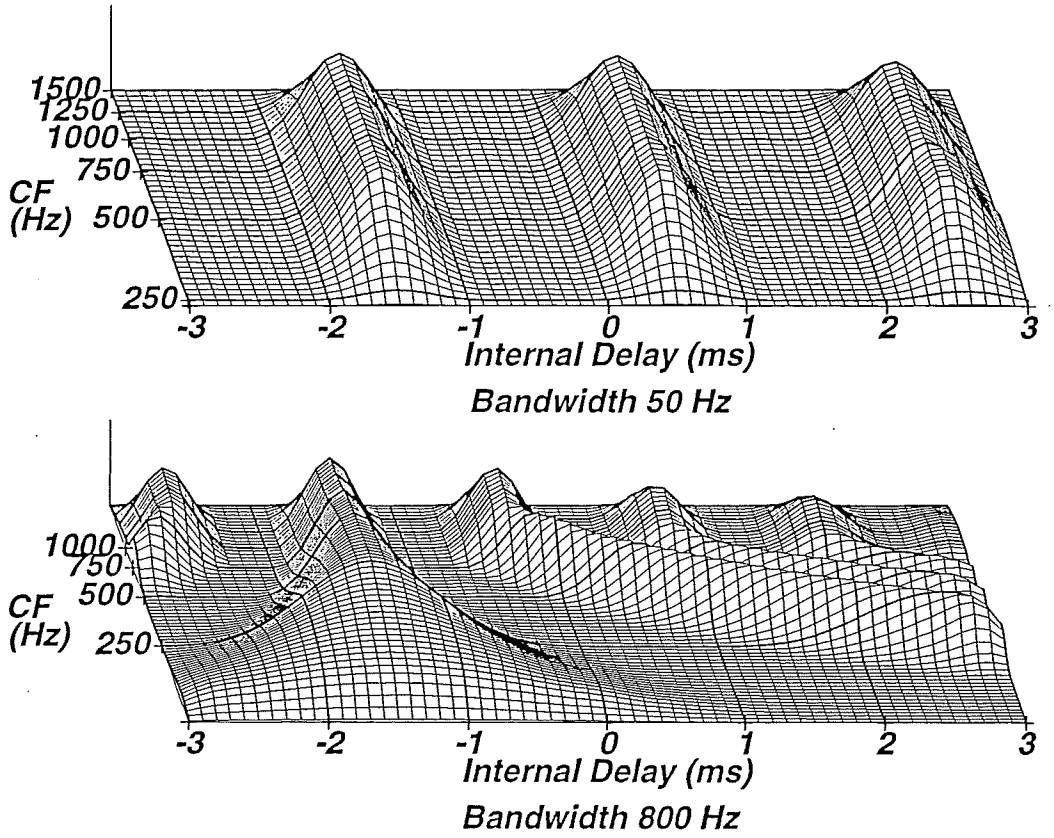

FIG. 7. Response of an ensemble of coincidence-counting units to low-frequency bandpass noise with a center frequency of 500 Hz and an ITD of −1.5 ms. Upper panel: Response to bandpass noise with a bandwidth of 50 Hz. Lower panel: Response to bandpass noise with a bandwidth of 800 Hz.

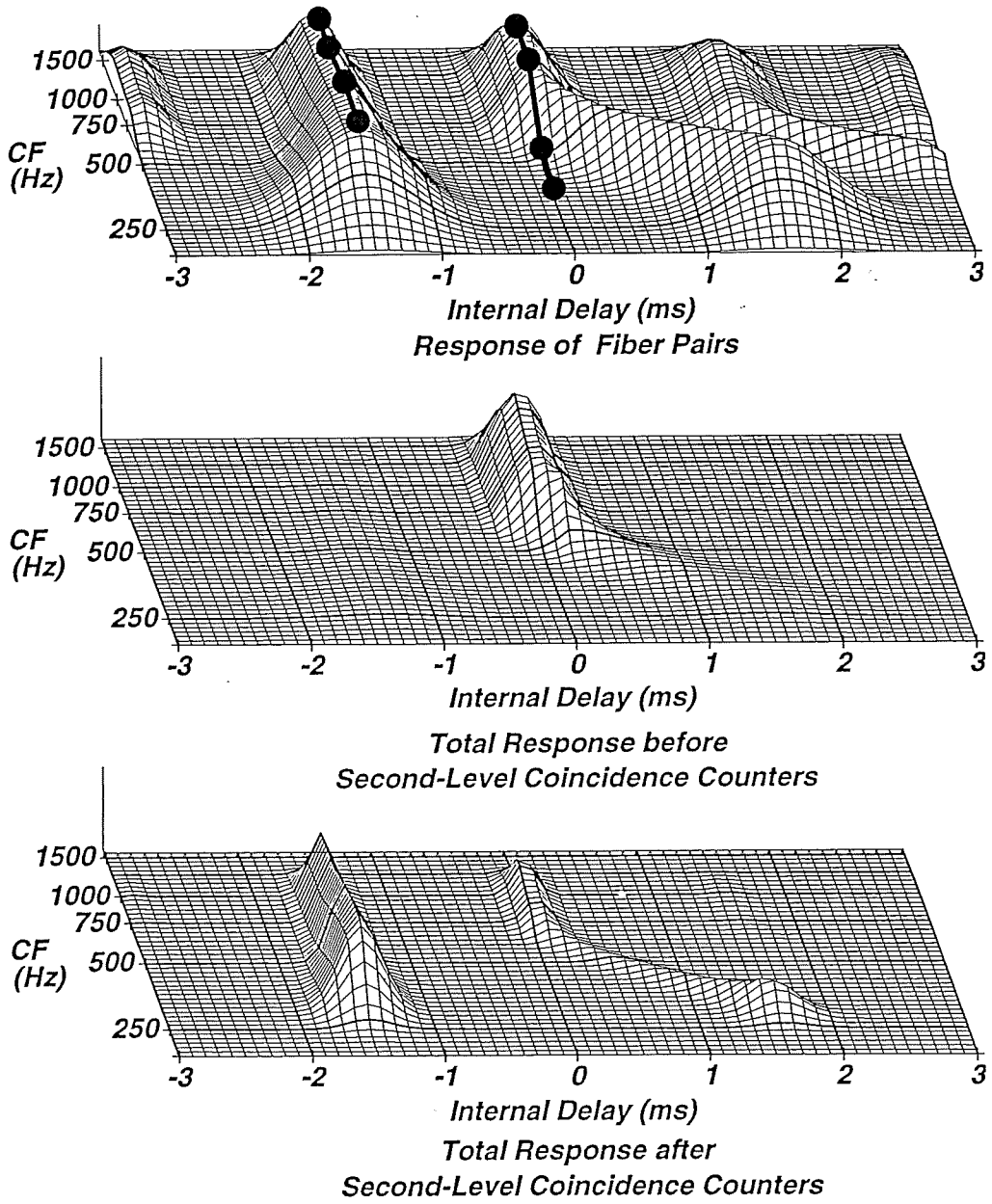

FIG. 8. Effect of the putative secondary level of coincidence-counting units that produce "straightness weighting." Upper panel: The response of an ensemble of binaural coincidence counters to noise with ITD –1.5 ms, center frequency 500 Hz, and bandwidth 400 Hz. Locations of constant internal delay but different CF are identified by filled circles joined by lines. Central panel: Same as upper panel, but incorporating the effects of the relative number of fiber pairs. Lower panel: Same as central panel, but after further processing by the second-level units that compute coincidences over frequency of the outputs of the original coincidence counters with the same internal delay.

any additional straightness weighting) and the response of the extended model. The stimulus in this figure is bandpass noise centered at 500 Hz with an ITD of –1.5 ms and a bandwidth of 400 Hz. The sets of points denoted by the filled circles in the upper panel of Fig. 8 are examples of combinations of CF and internal delay that would comprise inputs to the second-level coincidence counters. The center panel of Fig. 8 shows the effect of weighting by the relative number of fiber pairs, which suppresses the effects of the responses at the true ITD of –1.5

ms. The lower panel of Fig. 8 shows the dramatic effects of applying the second level of coincidences, which provides much greater emphasis to the straight ridge at –1.5 ms. This occurs because, for that ridge, all of the first-level coincidence counters are firing at rates that are at or near their maximum output. In contrast, the ridge closer to the midline (i.e., at an ITD of approximately zero) is attenuated because of the minimal response at characteristic frequencies below approximately 600 Hz at that ITD.

In addition to providing the weighting of straightness needed to describe the lateralization data as in Stern, Zeiberg, and Trahiotis (1988) and Trahiotis and Stern (1989), this manner of combining coincidence information across frequency also sharpens the ridges of the two-dimensional cross-correlation function along the internal-delay axis. For example, the ridges in the lower panel of Fig. 8 exhibit a smaller "width" (along the internal-delay axis) than the corresponding ridges in the upper panel of Fig. 8. The sharpening of the ridges along the internal-delay axis occurs because the rate functions of the outputs of the second-level coincidence counters are approximately proportional to the products of the rate functions of the (first-level) coincidence counters that comprise their inputs. For "straight" ridges, this has the effect of enhancing the peaks and suppressing the "valleys" in the patterns of second-level coincidence output. It is important to note that this sharpening along the internal-delay axis can occur without the explicit lateral-inhibition network proposed by Lindemann (1986a).

Figure 9 demonstrates how straightness weighting is needed to describe the lateralization of low-frequency bandpass noise. The upper panel of Fig. 9 shows the joint dependence of the lateral position of these stimuli on ITD, IPD, and bandwidth, as measured for human subjects by Stern, Zeiberg, and Trahiotis (1988). The combinations of ITD and IPD were selected because they produce maxima of the outputs of the coincidence counters at the same values of internal delay for values of CF near 500 Hz. The lateral position of noise with a 1.5-ms ITD and an IPD of 0° moves from one side of the head to the other as bandwidth increases, for the reasons discussed previously in this section. Other combinations of ITD and IPD show a similar, but weaker, effect. The central and lower panels of Fig. 9 contain predictions for the same set of data, both without and with the second-level coincidence mechanism schematized in Fig. 8 (Stern and Trahiotis, 1992). It can be seen that the second layer of coincidence-counting units is necessary for the model to describe the data. Trahiotis and Stern (1994) recently provided further evidence that a mechanism such as the second-level coincidence detectors is necessary in order to account for the position and character of intracranial images produced by multiple sinusoidally amplitude-modulated tones. That paper also contained a discussion concerning why a simple averaging, across frequency and running time, of the responses of the initial coincidence counters will not suffice.

The position-variable model as extended by Stern and Trahiotis (1992) appears to be able to describe quite well the lateralization of low-frequency bandpass noise as a joint function of ITD, IPD, and bandwidth, given that the signals are presented with equal amplitude to the two ears (e.g., Trahiotis and Stern, 1989). The model does not describe, however, some of the complex effects that occur when IIDs

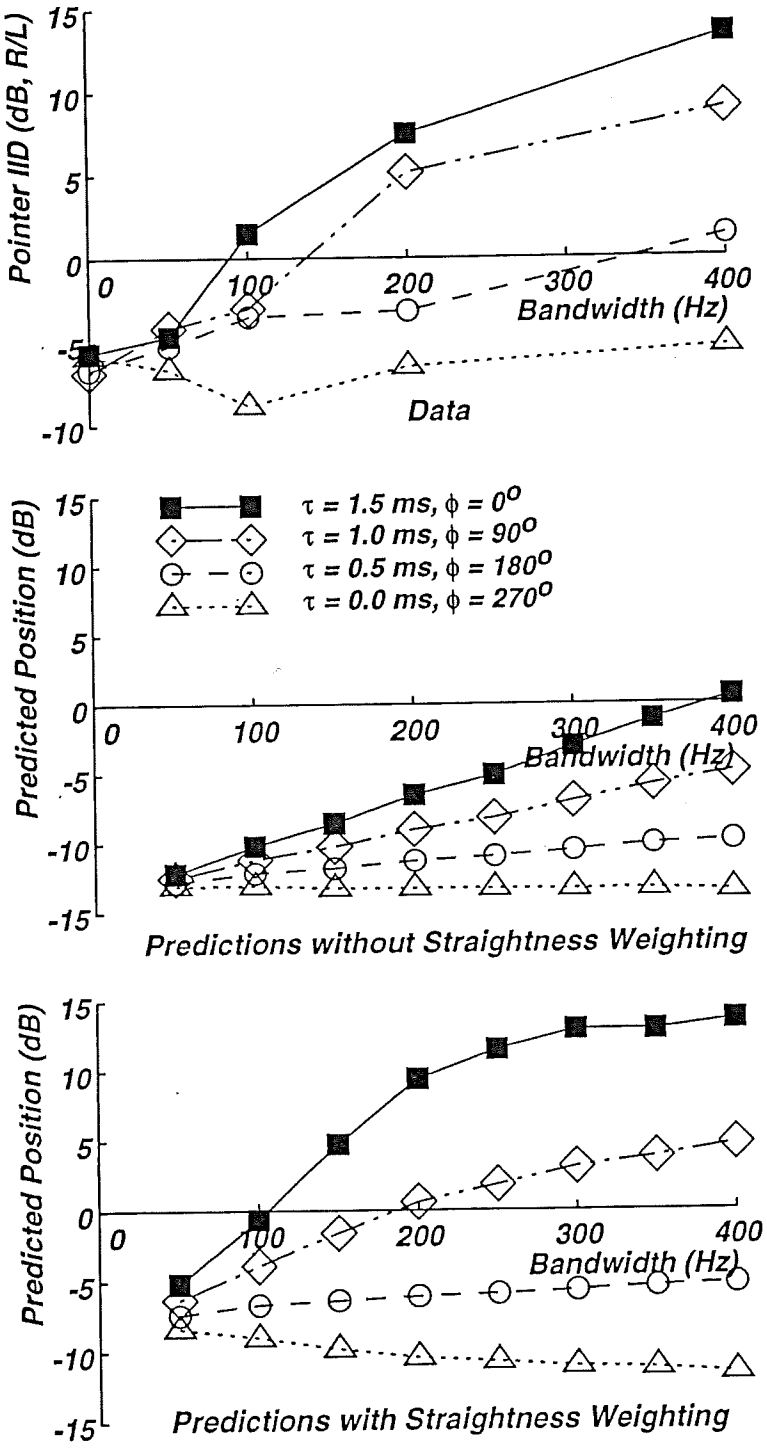

FIG. 9. Upper panel: Pointer IID needed to match the position of bandpass targets with center frequency 500 Hz and several combinations of ITD and IPD, as a function of bandwidth. Central panel: Predictions of the position-variable model without straightness weighting. Lower panel: Predictions of the position-variable model with straightness weighting. (Data from Stern, Zeiberg, and Trahiotis, 1988; predictions from Stern and Trahiotis, 1992.)

are added to these types of stimuli (Buell, Trahiotis, and Bernstein, 1994). For example, the subjective lateral position of bandpass noise presented with an ITD of 0 ms and an IPD of 270° is relatively independent of bandwidth for IIDs ranging from −10 to +10 dB. In contrast, the perceived position of similar stimuli presented with an ITD of 1.5 ms and an IPD of 0° moves toward the ear receiving the signal that leads in time as bandwidth increases from 50 to 400 Hz. To our knowledge the data of Buell *et al.* (1994) cannot be accounted for by any existing model of binaural interaction, despite concerted efforts (Tao, 1992; Tao and Stern, 1992).

3. Lateralization of low-frequency amplitude-modulated tones

It was reemphasized in the mid-1970s that the binaural system can utilize ITDs conveyed by the (low-frequency) envelopes of high-frequency stimuli (e.g., Henning, 1974; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1981). In keeping with the duplex theory, many had believed that ITDs were important for low-frequency stimuli where changes in the fine structure could be utilized. For high-frequency stimuli, IID was considered to be the salient binaural cue. Using sinusoidally amplitude-modulated (SAM) 500-Hz tones, Bernstein and Trahiotis (1985) demonstrated that the lateral position of low-frequency stimuli could also be affected, albeit by a small amount, by the ITD of the envelope of the stimulus.

Figure 10 shows the response of the coincidence-counting units to a 500-Hz tone (presented without amplitude modulation, upper panel) and the response

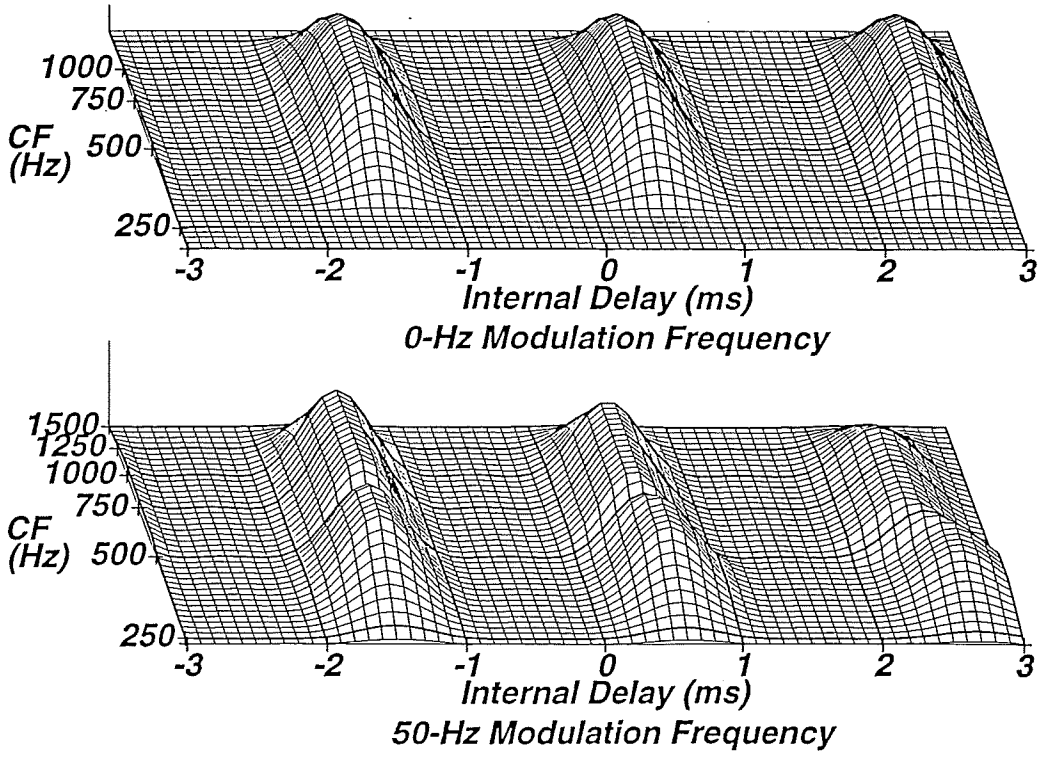

FIG. 10. Patterns of coincidence-counting activity showing the effects of amplitude modulation on low-frequency tones. Upper panel: The response to a 500-Hz tone with a waveform ITD of −1.5 ms. Lower panel: The response to a 500-Hz tone with the same waveform delay and amplitude modulated with a modulation frequency of 50 Hz.

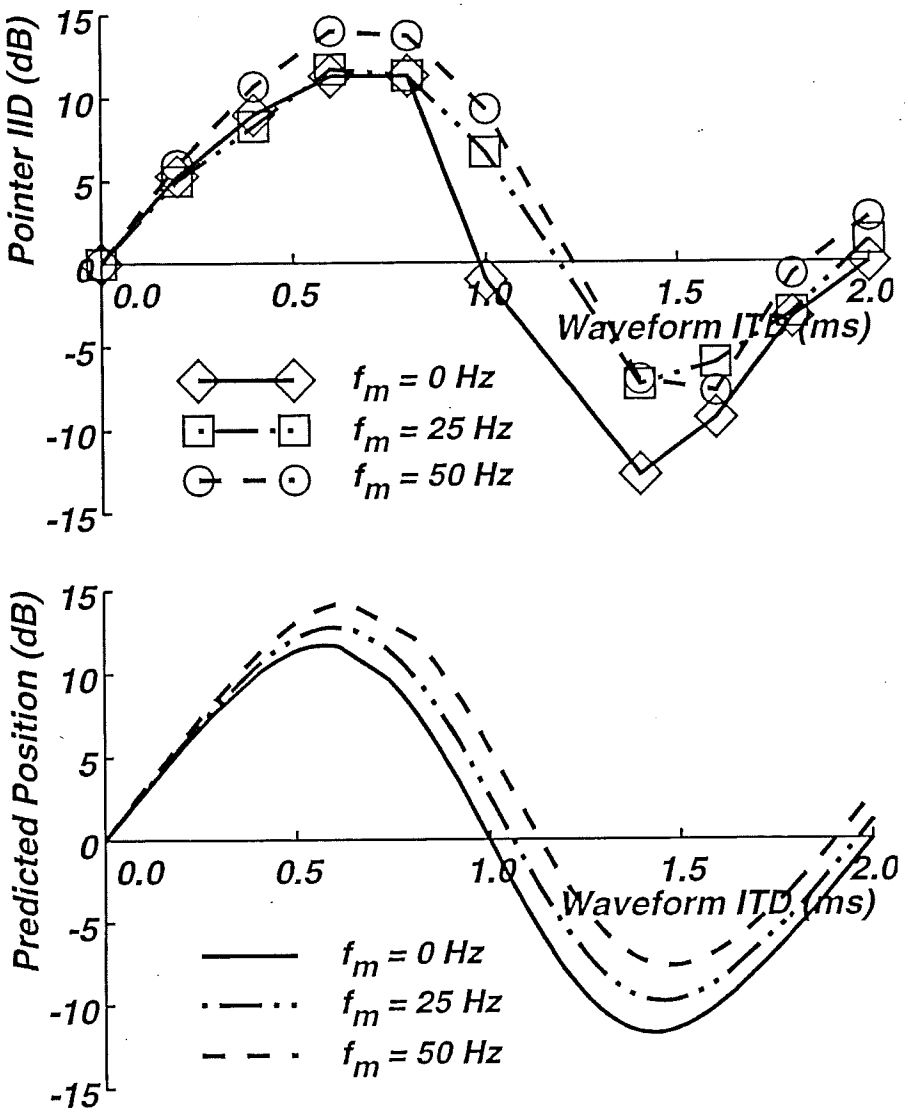

FIG. 11. Upper panel: Data by Bernstein and Trahiotis (1985) describing the subjective lateral position of 500-Hz amplitude modulated tones as a function of waveform ITD and modulation frequency. Lower panel: Predictions of the extended position-variable model (Stern and Shear, 1996) for these data.

to a 500-Hz tone sinusoidally modulated at a rate of 50 Hz (lower panel). The ongoing interaural delay is –1.5 ms in both cases. Figure 11 shows data obtained by Bernstein and Trahiotis (1985) and predictions by Stern and Shear (1996) concerning the joint dependence of the lateral position of SAM tones on modulation frequency and waveform ITD (which happened to be varied over a range of positive values in this particular experiment). The predicted dependence of position on envelope ITD comes about because the ridges in the coincidence-count response to the SAM stimulus are unequal in amplitude. For example, as shown in the lower panel of Fig. 10, the ridge at the true ITD (–1.5 ms in this case) is greater in magnitude than the other ridges. (This can be seen most clearly by comparing the height of the peaks in the lower panel to the constant response at 1500 Hz.) In contrast, the ridges of the response to the pure tone in the upper

panel of Fig. 10 are all of equal magnitude. The extended model also describes other aspects of Bernstein and Trahiotis' data including the dependence of lateral position on pure modulator delay.

4. Lateralization of high-frequency amplitude-modulated tones and bandpass noise

As noted in the preceding section, the extent of laterality of high-frequency binaural stimuli such as SAM tones and bandpass-noise can be affected by the ITD of the envelope. Figure 12 illustrates how such stimuli are represented by the ensemble of coincidence-counting units. These plots were produced without the use of any envelope extraction mechanism save for the lowpass filtering incorporated in the model of auditory-nerve activity. The lowpass filter has a

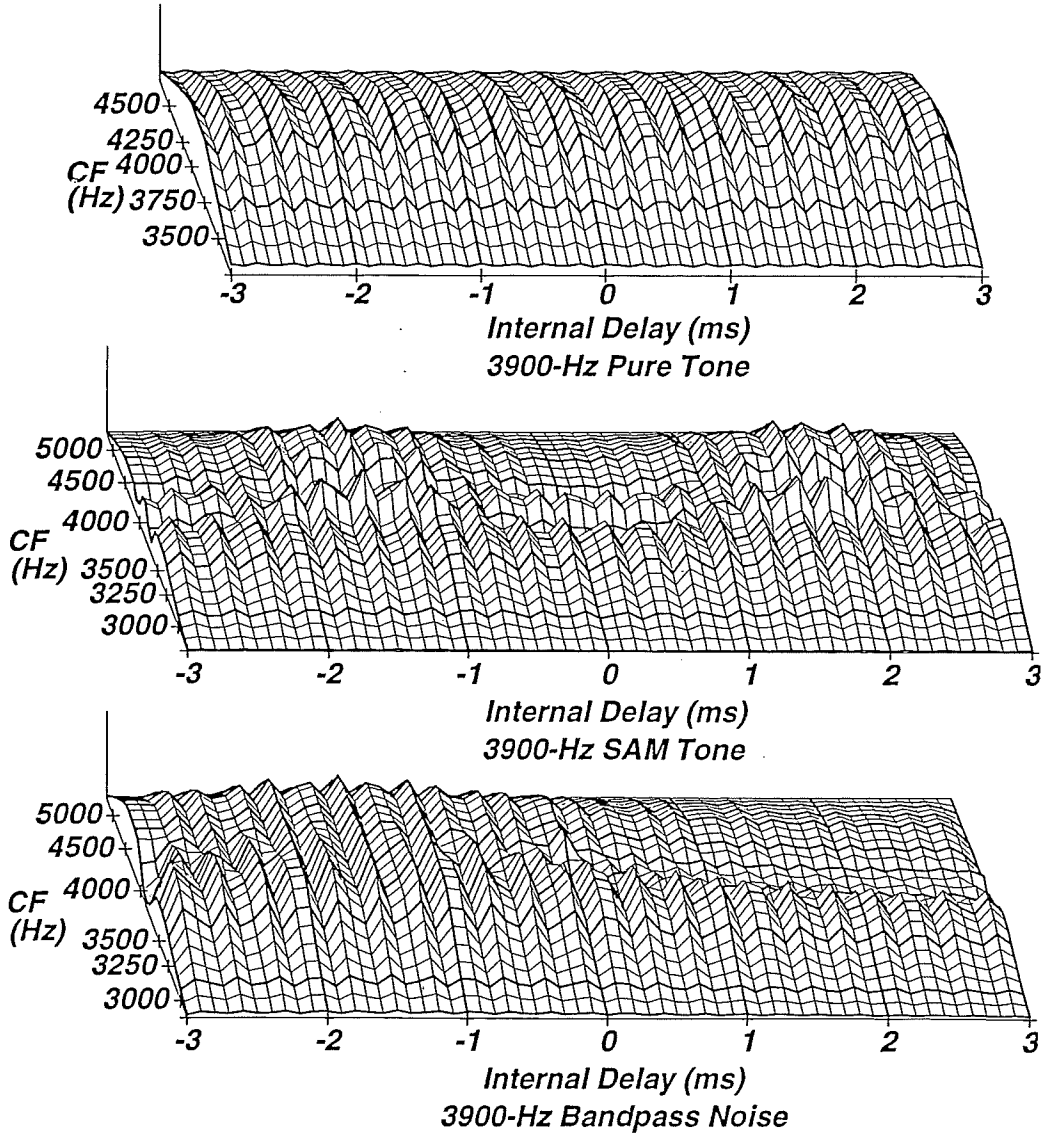

FIG. 12. The response of an ensemble of coincidence-counting units to several types of high-frequency stimuli with an ITD of −1.5 ms. Upper panel: Response to pure tones with a frequency of 3900 Hz. Central panel: Response to SAM tones with a modulation frequency of 300 Hz. Lower panel: Response to bandpass noise with a center frequency of 3900 Hz, bandwidth of 600 Hz.

frequency response that decreases linearly from 1200 to 5200 Hz, as suggested by the physiological data of Johnson (1980). The minor ripples in the plots show the effects of the residual energy at the relatively high carrier frequency after processing by the lowpass filter. The upper panel of Fig. 12 shows the relative number of coincidences observed in response to a pure tone of 3900 Hz. The central panel depicts the response to a SAM tone with a carrier frequency of 3900 Hz and a modulation frequency of 300 Hz. The lower panel of the same figure shows the response to a bandpass noise with a center frequency of 3900 Hz and a bandwidth of 600 Hz. Each stimulus has an ITD of –1.5 ms. Lateralization of the SAM tone and bandpass noise is dominated by the location along the internal-delay axis of the mode of the envelope of the response, which in Fig. 12 can be observed at an internal delay of approximately –1.5 ms. These observations are in accord with the conclusions of Colburn and Esquissaud (1976), who first suggested that cross-correlation-based models could be used to predict high-frequency binaural processing based on only the implicit envelope-extraction properties of the peripheral auditory system.

Models such as the extended position-variable model should, in principle, be able to describe most high-frequency lateralization data based on envelope delays. The upper panel of Fig. 13, for example, shows results of an ITD-discrimination experiment using high-frequency SAM tones (Henning, 1974). The lower panel of the figure depicts the corresponding predictions of the extended position-variable model (Stern, Shear, and Zeppenfeld, 1988). The model predicts the general form of these results by assuming that discrimination performance is mediated by changes in lateralization, as is discussed in Sec. III.B. On the other hand, the model is unable to predict the unexpected observation by Trahiotis and Bernstein (1986) that bandpass noise tends to be lateralized further from the center of the head than SAM tones of similar ITD, carrier frequency, and effective bandwidth. In general, there have been fewer stringent attempts to develop predictions for high-frequency binaural phenomena compared to their low-frequency counterparts.

5. Other lateralization phenomena

Thus far the discussion has concerned stimuli that have been used in "classical" psychoacoustical experiments. Several recent studies have shown that direct application of the cross-correlation-based binaural processing models described in this chapter can describe more complex phenomena as well. For example, Hafter and Shelton (1991) described the lateralization of diotic white noise that was passed through a bandpass filter and subsequently gated by brief rectangular pulses. The gating pulses themselves contained an ITD. Some of their data are shown in the upper panel of Fig. 14, which depicts percentage of "correct" response as a function of the center frequency of the bandpass filter. Some conditions produce significantly less than 50% "correct" response, indicating that the signals were lateralized toward the ear receiving the gating signal that was lagging in time. This apparently paradoxical result occurs because, for reasons discussed in Stern, Zeppenfeld, and Shear (1991), the major mode of the

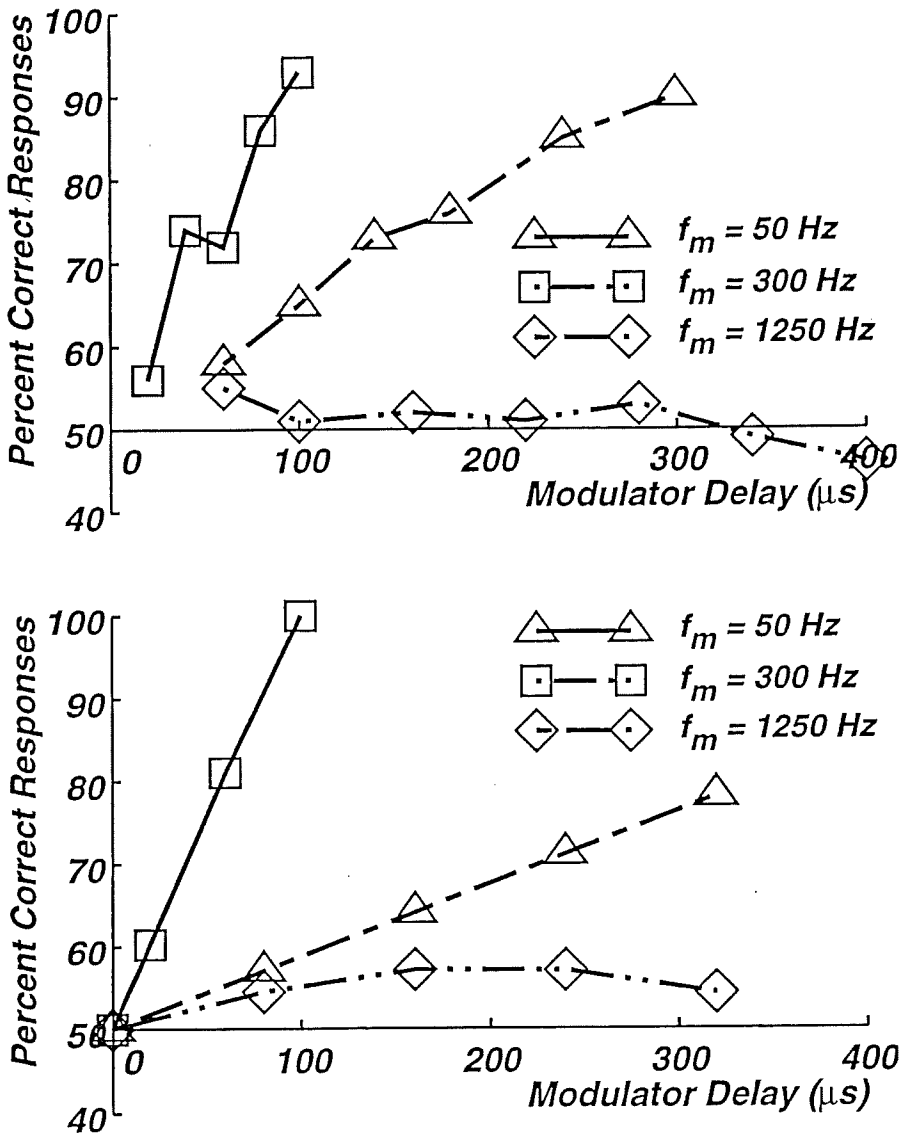

FIG. 13. Upper panel: Data by Henning (1974) describing the subjective lateral position of high-frequency SAM tones as a function of modulator delay and modulation frequency. Lower panel: Predictions of the extended position-variable model (Stern, Shear, and Zeppenfeld, 1988) for these data.

cross-correlation function of the response to these unusual stimuli is on the "wrong" (lagging) side of the internal-delay axis. The lower panel of Fig. 14 shows that the extended position-variable model does quite well in predicting the perception of such stimuli (Stern *et al.*, 1991).

Another interesting phenomenon accounted for by the model is what Bilsen and Raatgever (1973) termed the "dominant region" effect. This name refers to the fact that frequency components in the neighborhood of about 700 Hz are weighed more heavily in the lateralization of broadband noise than are frequency components in spectrally adjacent regions. Data from their experiment along with the corresponding predictions are presented in Fig. 15. The dependent variable, ΔI (dB), reflects the intensity of a narrow portion of the noise relative to the

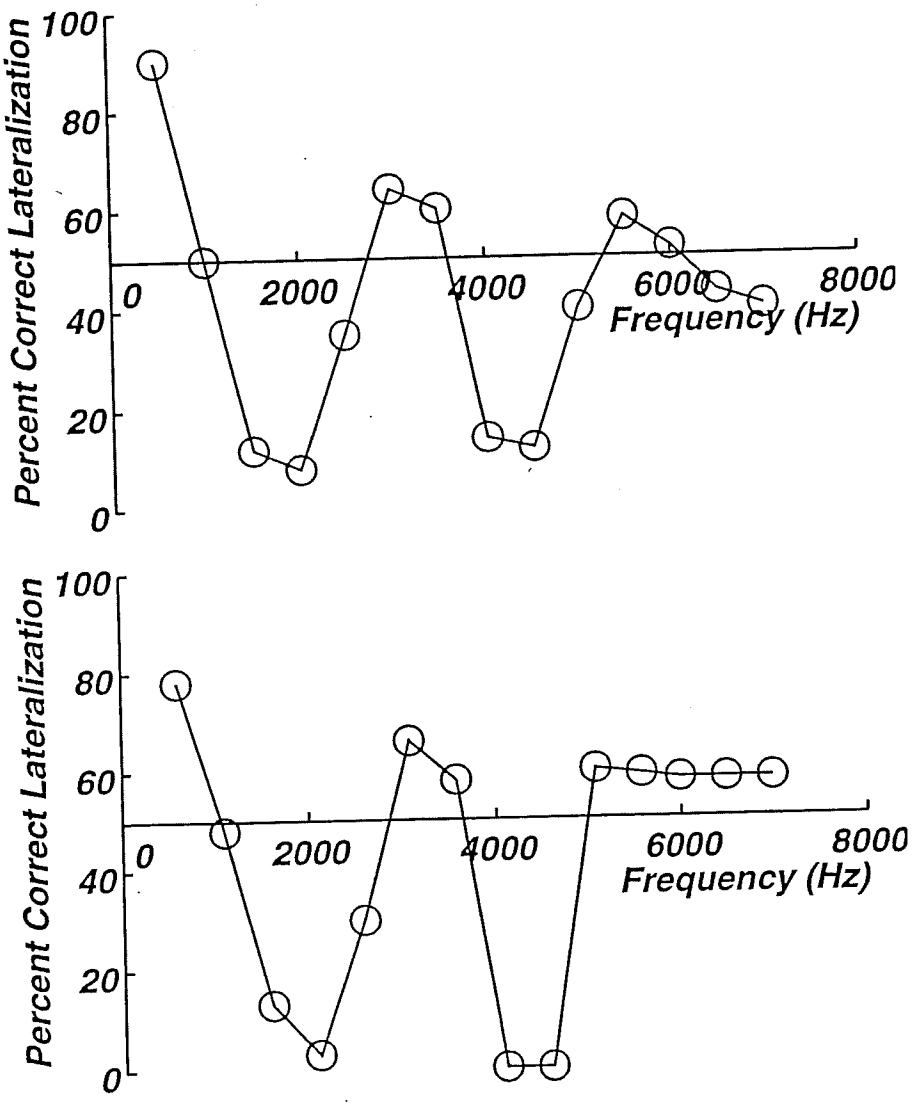

FIG. 14. Upper panel: Lateralization results by Hafter and Shelton (1991) for gated-noise stimuli as a function of center frequency. The ITD of the rectangular gating waveform is equal to 37.5 ms, the repetition period of the entire stimulus is equal to 2 ms, the noise bandwidth is equal to 1000 Hz, and the duration of the gating pulse is 400 ms. Lateralization percentages below 50% are the counterintuitive "illusory" reversals. Lower panel: Predictions of the extended position-variable model for the same stimulus conditions, except that the gating ITD was set equal to 50 ms for computational reasons. (Stern, Zeppenfeld, and Shear, 1991).

intensity of the spectral regions that surround it. The central and surrounding frequency bands were presented with conflicting ITDs and the listener's task was to adjust the level of the central band to maintain a centered image. It can be seen that the predictions of the model (in the curve without the data symbols) provide an excellent fit to the data (Stern, Shear, and Zeppenfeld, 1988; Stern and Shear, 1996). Stern and Shear (1996) have shown that the components of bandpass noise that carry the greatest weight in lateralization are those that produce patterns of activity of the coincidence-counting units that are about as wide as the major central portion of the $p(\tau, f)$ function.

B. Interaural discrimination phenomena related to subjective lateral position

The perceptual cue used by subjects in many interaural discrimination experiments is a change in the subjective lateral position of the stimuli. Models that describe the lateral position of binaural stimuli can be directly applied to discrimination experiments by computing or estimating the variance as well as the mean values of the predicted lateral positions of the stimuli. This is typically done by using optimal decision theory to estimate the best possible discrimination performance (cf. van Trees, 1968). Most of the early binaural models (e.g., Jeffress *et al.*, 1956; Hafter, 1971) implicitly assumed that intracranial position is a linear function of ITD and IID, and that the variance of the position estimate was independent of stimulus ITD and IID. These assumptions are valid if the ITD and IID of the stimuli used in an experiment are sufficiently small (Domnitz and Colburn, 1977). Models incorporating such assumptions can also predict the results of many lateralization-based detection experiments using either tonal targets and maskers or targets and maskers that are obtained by filtering, attenuating, and phase-shifting the same common noise source (e.g., Jeffress and McFadden, 1971; Yost, Nielsen, Tanis, and Bergert, 1974).

Colburn (1973) and Stern and Colburn (1985a, 1985b) provided predictions for interaural discrimination experiments using expressions for the variance of predicted position that were derived from the Poisson variability inherent in the auditory-nerve model describing the response to the stimuli. Colburn (1973) based his predictions on the amount of information in the ensemble of coincidence-counting units (without making any assumptions about the perceptual cue used by the subjects), and he predicted the dependence of just-noticeable

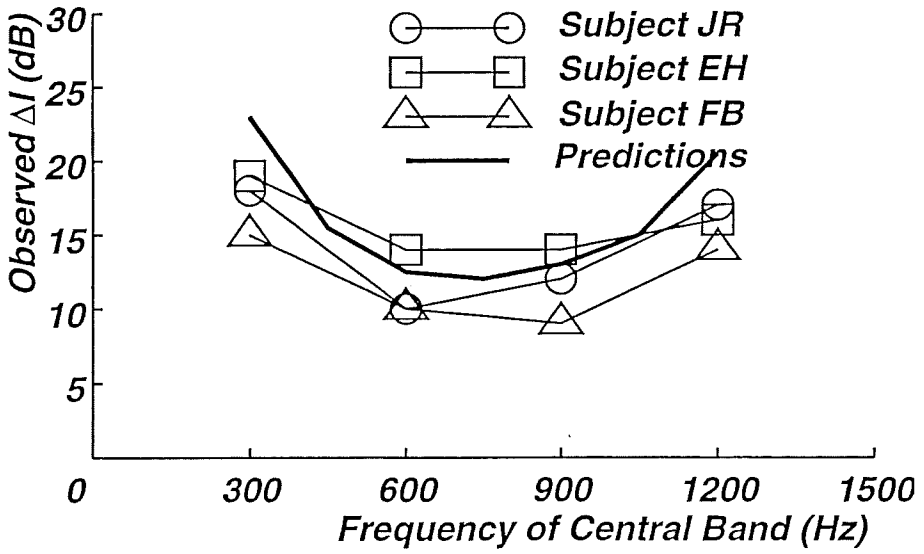

FIG. 15. Comparison of data by Bilsen and Raatgever (1973) and predictions by Stern and Shear (1996) for experiments whose results imply the existence of a "dominant frequency region" for binaural lateralization. The increment in overall intensity, ΔI, that is needed for a mid-frequency region of critical bandwidth to dominate the lateralization mechanism when frequency components in the two flanking bands are presented with a conflicting ITD is plotted as a function of frequency.

differences (JNDs) in ITD and IID on baseline ITD, IID, and overall level (Hershkowitz and Durlach, 1969). Stern and Colburn (1985b) derived an analytical expression for the variance of the predicted position variable $\hat{P}$.

As an example, Fig. 16 compares measurements of interaural JNDs in ITD by Domnitz and Colburn (1977) with predictions by Stern and Colburn (1985a, 1985b). The predictions were obtained by calculating the mean and variance of

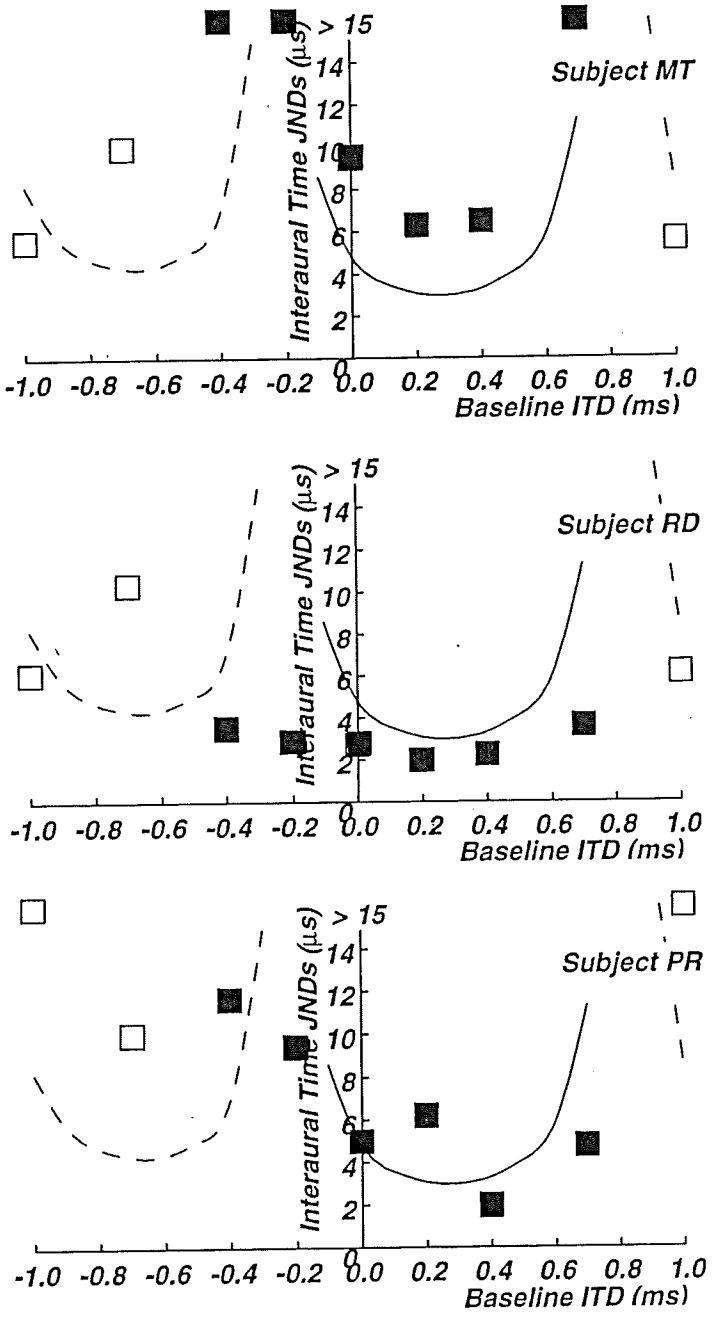

FIG. 16. Data by Domnitz and Colburn (1977) and predictions of the position-variable model (Stern and Colburn, 1985b) for interaural time JNDs as a function of baseline ITD. The stimuli were presented with an IID of +20 dB. Individual data points are plotted for each of three subjects. Filled symbols and solid curves represent conditions for which the cues are "normal"; open symbols and broken curves represent conditions for which the cues are "reversed."

$\hat{P}$ for the stimuli of each experiment and determining the value of ITD needed to produce unit value for the sensitivity index d'. The stimuli were presented with an IID of +20 dB. Filled symbols and solid curves represent conditions for which the cues are "normal" (i.e., moving in the same direction, as in time JND experiments with baseline ITDs and IIDs of small magnitude); open symbols and broken curves represent conditions for which the cues are "reversed" (i.e., moving in the opposite direction). The model correctly predicts the general increase of the magnitude of JNDs of ITD as baseline IID increases and the asymmetry with respect to baseline ITD for stimuli presented with nonzero IID. It also correctly predicts the reversal in direction reported by subjects when the baseline ITD is near 1 ms. Although the quantitative fit of predictions to data for interaural JNDs of IID is not as good, the model correctly predicts a lack of cue reversals, a weaker dependence on baseline ITD (compared to JNDs of ITD), and an overall increase of the magnitude of the JND with increasing IID.

In general, the simple position-variable model is unable to account for discrimination data in which subjects are likely to be making use of additional cues besides the lateral position of a single dominant time–intensity traded image (Stern and Colburn, 1985a, 1985b). For example, in Fig. 16, predicted interaural time JNDs are much larger than most of the observed data for baseline ITDs near the "cue-reversal" points.

A second example of this phenomenon is provided in Fig. 17, which compares data by Jeffress and McFadden (1971) for detection thresholds and "lateralization thresholds" to the corresponding theoretical predictions (Stern and Colburn, 1985b). The targets and maskers were derived from the same narrowband noise with a center frequency of 500 Hz and a bandwidth of 50 Hz. In the detection experiment subjects indicated whether or not the target stimulus was perceived to be present, as in traditional masking studies. In the lateralization-threshold experiments, the target was presented on every trial, but the signals to the two ears were randomly interchanged. Subjects in the lateralization-threshold experiments indicated the side of the head toward which the target-masker complex was perceived. Lateralization thresholds (square symbols) and detection thresholds (circular symbols) are plotted in Fig. 17 as a function of target-to-masker phase angle. Predictions (smooth curve) were obtained by adjusting model parameters to describe the relative salience of cues from ITDs and IIDs for each subject individually (Stern and Colburn, 1985b). The predictions provide a very good fit to the lateralization-threshold data, despite the sharply differing ability of the two subjects to make use of ITDs and IIDs. Nevertheless, the observed detection thresholds are much lower than predicted, implying that subjects are making use of attributes of the stimuli besides lateral position. Subjects perform similarly better than predicted in discrimination results concerning time–intensity tradability (e.g., Hafter and Carrier, 1972; Gilliom and Sorkin, 1972). Again, this occurs because the theoretical predictions are based only on the dominant time–intensity traded image of the stimuli, whereas the data reflect the use of more than one image component or some other additional cue.

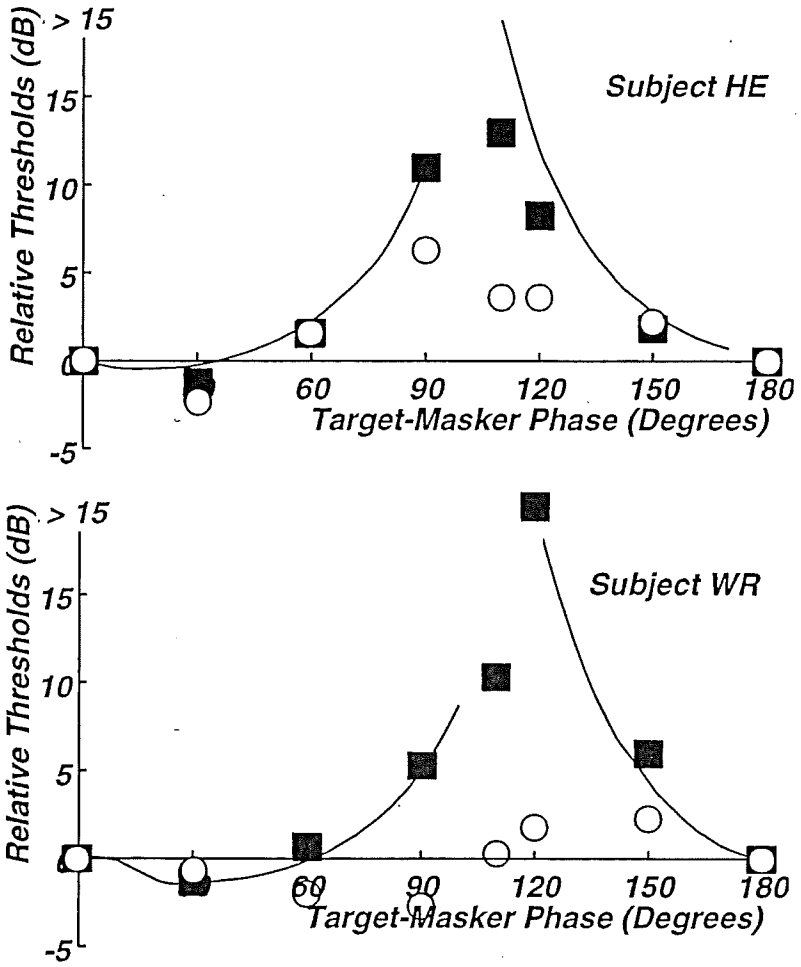

FIG. 17. Comparisons of predictions and data for two experiments by Jeffress and McFadden (1971). The target and masker are coherent narrowband-noise waveforms with center frequency 500 Hz and bandwidth 50 Hz. Lateralization thresholds (square symbols) and detection thresholds (circular symbols) are plotted as a function of target-to-masker phase angle. Predictions (smooth curves) were obtained by adjusting model parameters to describe the relative salience of cues from ITDs and IIDs for each subject individually (Stern and Colburn, 1985b).

C. Binaural masking-level differences

The binaural masking-level difference (BMLD) is an extremely well known and robust binaural phenomenon. A large number of classical measurements of BMLDs are summarized in Durlach and Colburn (1978), and several more recent results are described by Kohlrausch and Fassel (Chapter 9, this volume). Figure 18 illustrates how the ensemble of coincidence-counting units responds to typical stimuli used in classical BMLD experiments. The figure shows the patterns of activity that result when a 500-Hz tonal target and a broadband masking noise are presented in the N_0S_π (masker interaurally in phase, target interaurally out of phase) and N_0S_0 (masker and target both interaurally in phase) configurations. The plots in Fig. 18 include the effects of the relative number of fiber pairs, as specified by the function $p(\tau, f)$. Note that when the N_0 masker is presented alone (Fig. 18, lower panel), the ridge of maxima at zero internal delay has approximately constant amplitude over a broad range of frequencies. The addition

of an in-phase (S_0) target to the masker at a target-to-masker intensity ratio of –20 dB has virtually no effect on the pattern of coincidence-counting activity, because the interaural time differences of the combined target and masker are unchanged (Fig. 18, central panel). On the other hand, the addition of the 500-Hz out-of-phase (S_π) target to the in-phase masker cancels masker components at that frequency, causing a "dimple" to appear in the central ridge for CFs near the target frequency (Fig. 18, upper panel). The target in the N_0S_π configuration is

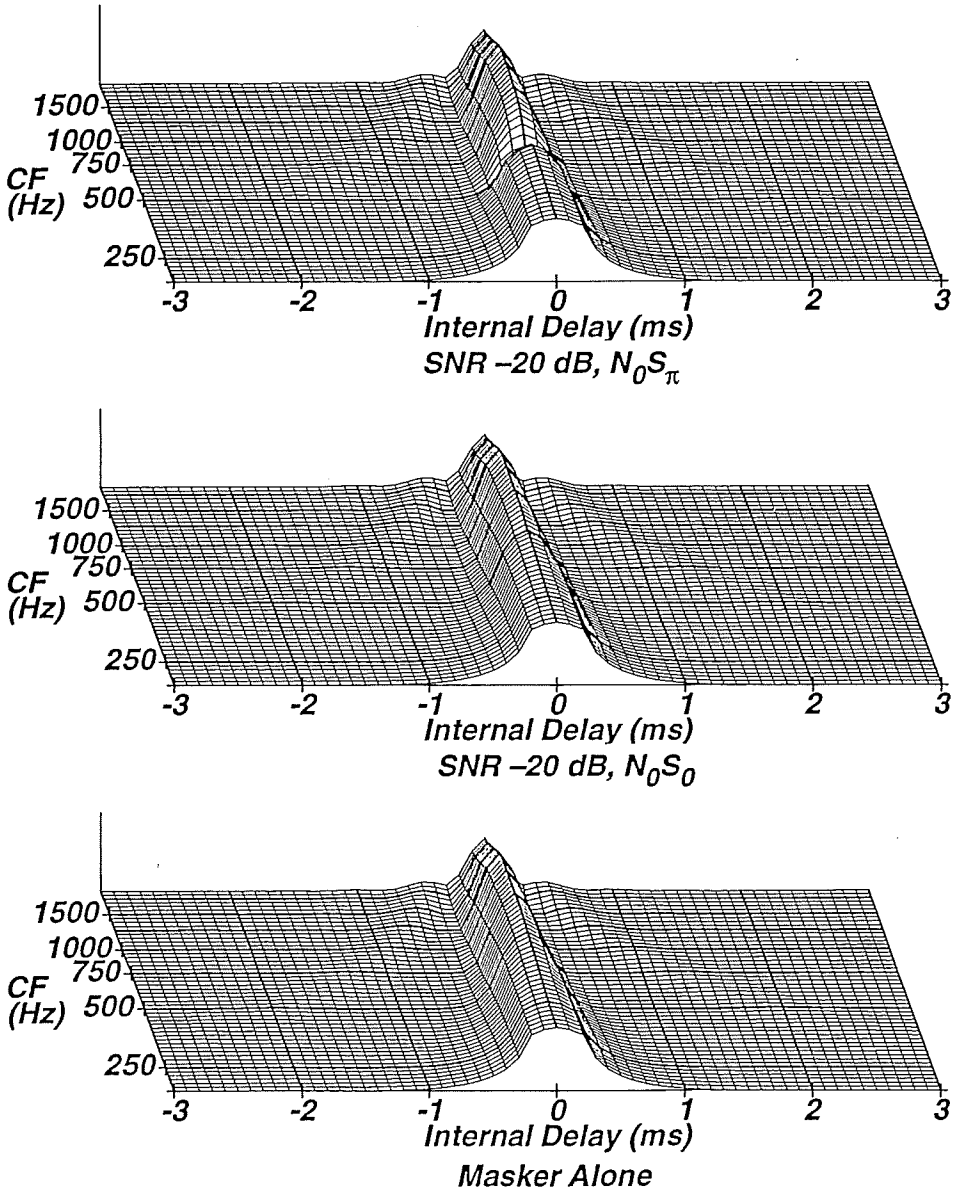

FIG. 18. Patterns of coincidence-counting activity showing the response to stimuli used in N_0S_π and N_0S_0 binaural masking level difference experiments. The target is presented at 500 Hz, either interaurally in phase or out of phase, as indicated, and the masker is broadband diotic noise. These plots include the effects of the relative number of fiber pairs, as specified by the function.

easily detected at –20 dB SNR because the pattern of responses in the upper panel of Fig. 18 is easily discriminated from that in the lower panel. The N_0S_0 stimulus is not detected because the response of the binaural system is largely unaffected by whether the target is present or absent (compare central and lower panels of Fig. 18).

Colburn (1977) was able to describe virtually all of the "classical" data obtained in experiments measuring BMLDs on the basis of the predicted outputs of the coincidence counters. His predictions were developed using the simplifying assumption that experimental performance is limited by the variability of the auditory-nerve response to the signals, as opposed to the intrinsic variability of the masker components. This assumption has since been shown to be invalid for some stimuli by Siegel and Colburn (1983). More recently, Gilkey and his colleagues (e.g., Gilkey, Robinson, and Hanna, 1985; Hanna and Robinson, 1985; Isabelle and Colburn, 1991) presented a number of results using "frozen-noise" maskers in which the actual variability of the masker component of the stimulus can be experimentally controlled. To date no binaural model has been able to account for differences of detectability associated with the individual masker waveforms used in these studies.

Although predictions of lateral position, interaural discrimination, and binaural detection are all obtained by considering the patterns of outputs of the interaural coincidence-counting units, we believe that binaural detection phenomena are mediated by a different type of reading of the information from the display of coincidence-counting units (compared to that used for subjective lateral position and interaural discrimination). Specifically, the subjective lateral position of binaural stimuli and the ability to perform certain interaural discrimination tasks based on changes in lateral position both appear to depend on the locations of the ridges of the cross-correlation function along the τ axis. In contrast, successful predictions for binaural detection tasks can be obtained by quantifying the decrease in amplitude of these ridges at the target frequency produced by the addition of the target to the masker.

IV. CONCLUSIONS

We have provided several examples illustrating how many of the fundamental data concerning binaural hearing can be predicted or explained within a unified theoretical framework. Predictions are based on the internal patterns resulting from cross-correlation of the neural responses to the stimuli by the peripheral auditory system. In our view, recent extensions of basic models by Jeffress (1948) and Colburn (1973) are quite successful in accounting for a wide variety of phenomena. We expect that further advances in signal processing and digital computation will allow an even wider range of stimuli to be considered. These advances, the recent trend toward unifying data obtained in laboratory environments with data obtained in more realistic settings, and the development of practical applications that exploit our theoretical insights make it likely that the next decade will prove to be even more fruitful.

ACKNOWLEDGMENTS

Preparation of this manuscript has been supported by National Science Foundation grant IBN 90-22080 to Richard Stern and by National Institutes of Health grant DC-00234 and Air Force Office of Scientific Research grant 89-0030 to Constantine Trahiotis. Preparation of the figures has been facilitated by the efforts of Carl Block, Wonseok Lee, Steve Palm, Glenn Shear, Sammy Tao, Xaohong Xu, Andreas Yankopolus, and Torsten Zeppenfeld.

REFERENCES

Bachorski, S. J. (1983). "Dynamic cues in binaural perception," MS thesis, Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA.

Bernstein, L. R., and Trahiotis, C. (1985). "Lateralization of low-frequency complex waveforms: The use of envelope-based temporal disparities," J. Acoust. Soc. Am. 77, 1868–1880.

Bernstein, L. R., and Trahiotis, C. (1994). "Spectral interference in a binaural detection task: Effects of masker bandwidth and temporal fringe," J. Acoust. Soc. Am. 94, 735–742.

Bilsen, F. A., and Raatgever, J. (1973). "Spectral dominance in binaural lateralization," Acustica 28, 131–132.

Blauert, J. (1980). "Modelling of interaural time and intensity difference discrimination," in *Psychophysical, Physiological, and Behavioral Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft University Press, Delft), pp. 421–424.

Blauert, J., and Cobben, W. (1978). "Some consideration of binaural cross-correlation analysis," Acustica 39, 96–103.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (1994). "Lateralization of bands of noise as a function of combinations of interaural intensitive differences, interaural temporal differences, and bandwidth," J. Acoust. Soc. Am. 95, 1482–1489.

Carney, L. H. (1993). "A model for the responses of low-frequency auditory nerve fibers in cat," J. Acoust. Soc. Am. 93, 401–417.

Carr, C. E., and Konishi, M. (1990). "A circuit for detection of interaural time differences in the brain stem of the barn owl," J. Neurosci. 10, 3227–3246.

Colburn, H. S. (1969). "Some Physiological Limitations on Binaural Performance," doctoral dissertation, MIT, Cambridge, MA.

Colburn, H. S. (1973). "Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination," J. Acoust. Soc. Am. 54, 1458–1470.

Colburn, H. S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise," J. Acoust. Soc. Am. 61, 525–533.

Colburn, H. S. (1995). "Computational models of binaural processing," in *Auditory Computation*, edited by H. Hawkins and T. McMullin (Springer-Verlag, New York), pp. 332–400.

Colburn, H. S., and Durlach, N. I. (1978). "Models of binaural interaction," in *Handbook of Perception, Volume IV, Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic Press, New York), pp. 467–518.

Colburn, H. S., and Esquissaud, P. (1976). "An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli," J. Acoust. Soc. Am. 59, S23(A).

David, E. E., Guttman, N., and van Bergeijk, W. A. (1958). "On the mechanism of binaural fusion," J. Acoust. Soc. Am. 30, 801–802.

Deatherage, B. H., and Hirsh, I. J. (1959). "Auditory localization of clicks," J. Acoust. Soc. Am. 31, 486–492.

Domnitz, R. H., and Colburn, H. S. (1976). "Analysis of binaural detection models for dependence on interaural target parameters," J. Acoust. Soc. Am. 59, 598–601.

Domnitz, R. H., and Colburn, H. S. (1977). "Lateral position and interaural discrimination," J. Acoust. Soc. Am. 61, 1586–1598.

Duifhuis, H. (1973). "Consequences of peripheral frequency selectivity for nonsimultaneous masking," J. Acoust. Soc. Am. 54, 1471–1488.

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

Durlach, N. I. (1972). "Binaural signal detection: Equalization and cancellation theory," in *Foundations of Modern Auditory Theory, Volume II*, edited by J.V. Tobias (Academic Press, New York), pp. 369–462.

Durlach, N. I., and Colburn, H. S. (1978). "Binaural phenomena," in *Handbook of Perception, Volume IV, Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic Press, New York), pp. 365–466.

Gabriel, K. J. (1983). "Binaural Interaction in Hearing Impaired Listeners," doctoral dissertation, MIT, Cambridge, MA.

Gaik, W. (1993). "Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling," J. Acoust. Soc. Am. 94, 98–110.

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.
Gilliom, J. D., and Sorkin, R. D. (1972). "Discrimination of interaural time and intensity," J. Acoust. Soc. Am. 52, 1635–1644.
Grantham, D. W. (1984). "Discrimination of dynamic interaural intensity differences," J. Acoust. Soc. Am. 76, 71–76.
Grantham, D. W., and Wightman, F. L. (1978). "Detectability of varying interaural temporal differences," J. Acoust. Soc. Am. 63, 511–523.
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 Carrier, S. C. (1970). "Masking-level difference obtained with a pulsed tonal masker," J. Acoust. Soc. Am. 47, 1041–1047.
Hafter, E. R., and Carrier, S. C. (1972). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. 51, 1852–1862.
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 Shelton, B. R. (1991). "Counterintuitive reversals in lateralization using rectangularly-modulated noise," J. Acoust. Soc. Am. 90, 1901–1907.
Hanna, T. E., and Robinson, D. E. (1985). "Phase effects for a sine wave masked by reproducible noise," J. Acoust. Soc. Am. 77, 1129–1140.
Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. 55, 84–90.
Hershkowitz, R. M., and Durlach, N. I. (1969). "Interaural time and amplitude jnds for a 500–Hz tone," J. Acoust. Soc. Am. 46, 1464–1467.
Hirsh, I. J. (1948). "The influence of interaural phase on interaural summation and inhibition," J. Acoust. Soc. Am. 29, 536–544.
Isabelle, S. K., and Colburn, H. S. (1991). "Detection of tones in reproducible narrow-band noise," J. Acoust. Soc. Am. 89, 352–359.
Jeffress, L. A. (1948). "A place theory of sound localization," J. Comp. Physiol. Psychol. 41, 35–39.
Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Wood, C. L. III. (1956). "Masking of tonal signals," J. Acoust. Soc. Am. 28, 416–426.
Jeffress, L.A., and McFadden, D. (1971). "Differences of interaural phase and level in detection and lateralization," J. Acoust. Soc. Am. 49, 1169–1179.
Johnson, D. H. (1980). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. 68, 1115–1122.
Kuwada, S., Stanford, T. R., and Batra, R. (1987). "Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit: Effects of changing frequency," J. Neurophysiol. 57, 1338–1360.
Licklider, J. C. R. (1948). "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. 20, 150–159.
Licklider, J. C. R. (1959). "Three auditory theories," in *Psychology: A Study of a Science*, edited by S. Koch (McGraw-Hill, New York), pp. 41–144.
Licklider, J. C. R., Webster, J. C., and Hedlun, J. M. (1950). "On the frequency limits of binaural beats," J. Acoust. Soc. Am. 22, 468–473.
Lindemann, W. (1986a). "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," J. Acoust. Soc. Am. 80, 1608–1622.
Lindemann, W. (1986b). "Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wavefront," J. Acoust. Soc. Am. 80, 1623–1630.
McFadden, D., and Pasanen, E. G. (1976). "Lateralization at high frequencies based on interaural time differences," J. Acoust. Soc. Am. 59, 634–639.
Meddis, R., Hewitt, M. J., and Shackleton, T. M. (1990). "Implementation details of a computational model of the inner hair-cell/auditory-nerve synapse," J. Acoust. Soc. Am. 87, 1813–1816.
Moushegian, G., and Jeffress, L. A. (1959). "Role of interaural time and intensity differences in the lateralization of low-frequency tones," J. Acoust. Soc. Am. 31, 1441–1445.
Nuetzel, J. M., and Hafter, E. R. (1981). "Discrimination of interaural delays in complex waveforms: Spectral effects," J. Acoust. Soc. Am. 69, 1112–1118.
Osman, E. (1971). "A correlation model of binaural masking level differences," J. Acoust. Soc. Am. 50, 1491–1511.
Payton, K. L. (1988). "Vowel processing by a model of the auditory periphery: A comparison to eighth-nerve responses," J. Acoust. Soc. Am. 83, 145–162.
Rayleigh, Lord (J.W. Strutt, 3rd Baron of Rayleigh) (1907). "On our perception of sound direction," Philos. Mag. 13, 214–232.
Rose, J. E., Gross, N. B., Geisler, C. D., and Hind, J. E. (1966). "Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source," J. Neurophysiol. 29, 288–314.

Sayers, B. M. (1964). "Acoustic-image lateralization judgments with binaural tones," J. Acoust. Soc. Am. 36, 923–926.
Sayers, B. M., and Cherry, E. C. (1957). "Mechanism of binaural fusion in the hearing of speech," J. Acoust. Soc. Am. 29, 973–987.
Schiano, J. L., Trahiotis, C., and Bernstein, L. R. (1986). "Lateralization of low-frequency tones and narrow bands of noise," J. Acoust. Soc. Am. 79, 1563–1570.
Schroeder, M. R. (1977). "New viewpoints in binaural interactions," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic Press, London), pp. 455–467.
Shackleton, T. M., Meddis, R., and Hewitt, M. J. (1992). "Across frequency integration in a model of lateralization," J. Acoust. Soc. Am. 91, 2276–2279(L).
Shamma, S. A., Shen, N., and Gopalaswamy, P. (1989). "Binaural processing without neural delays," J. Acoust. Soc. Am. 86, 987–1006.
Shear, G. D. (1987). "Modeling the Dependence of Auditory Lateralization on Frequency and Bandwidth," MS thesis, Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA.
Siebert, W. M. (1970). "Frequency discrimination in the auditory system: Place or periodicity mechanisms," Proc. IEEE 58, 723–730.
Siegel, R. A., and Colburn, H. S. (1983). "Internal and external noise in binaural detection," Hear. Res. 11, 117–123.
Smith, P. H., Joris, P. X., and Yin, T. C. T. (1993). "Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: Evidence for delay lines to the medial superior olive," J. Comput. Neurol. 331, 245–260.
Stern, R. M., Jr., and Bachorski, S. J. (1983). "Dynamic cues in binaural perception," in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer, Berlin), pp. 209–215.
Stern, R. M., Jr., and Colburn, H. S. (1978). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. 64, 127–140.
Stern, R. M., and Colburn, H. S. (1985a). "Lateral-position-based models of interaural discrimination," J. Acoust. Soc. Am. 77, 753–755.
Stern, R. M., and Colburn, H. S. (1985b). "Subjective Lateral Position and Interaural Discrimination," J. Acoust. Soc. Am., AIP Document No. PAPS JASMA-77–753–29, pp. 1–29.
Stern, R. M., and Shear, G. D. (1996). "Lateralization and detection of low frequency binaural stimuli: Effects of distribution of internal delay," J. Acoust. Soc. Am. (in press).
Stern, R. M., Shear, G. D., and Zeppenfeld, T. (1988). "High-frequency predictions of the position-variable model," J. Acoust. Soc. Am. 84, S60 (A).
Stern, R. M., and Trahiotis, C. (1992). "The role of consistency of interaural timing over frequency in binaural lateralization," in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon Press, Oxford), pp. 547–554.
Stern, R. M., and Trahiotis, C. (1995). "Models of Binaural Interaction," in *Handbook of Perception and Cognition, Volume 6: Hearing*, edited by B. C. J. Moore (Academic Press, New York), pp. 347–386.
Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (1988). " Lateralization of complex binaural stimuli: A weighted image model," J. Acoust. Soc. Am. 84, 156–165.
Stern, R. M., Zeppenfeld, T., and Shear, G. D. (1991). "Lateralization of rectangularly-modulated noise: explanations for counterintuitive reversals," J. Acoust. Soc. Am. 90, 1901–1907.
Tao, S. H. (1992). "Additive versus multiplicative combination of differences of interaural time and intensity," MS thesis, Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA.
Tao, S. H., and Stern, R. M. (1992). "Additive versus multiplicative combination of differences of interaural time and intensity," J. Acoust. Soc. Am. 91, 2414(A).
Trahiotis, C., and Bernstein, L. R. (1986). "Lateralization of bands of noise and sinusoidally amplitude-modulated tones: effects of spectral locus and bandwidth," J. Acoust. Soc. Am. 79, 1950–1957.
Trahiotis, C., and Stern, R. M. (1989). "Lateralization of bands of noise: effects of bandwidth and differences of interaural time and intensity," J. Acoust. Soc. Am. 86, 1285–1293.
Trahiotis, C., and Stern, R. M. (1994). "Across-frequency interaction in lateralization of complex binaural stimuli," J. Acoust. Soc. Am. 96, 3804–3806(L).
van Bergeijk, W. A. (1962). "Variation on a theme of von Bekesy: A model of binaural interaction," J. Acoust. Soc. Am. 34, 1431–1437.
van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York).
Whitworth, R. H., and Jeffress, L. A. (1961). "Time versus intensity in the localization of tones," J. Acoust. Soc. Am. 33, 925–929.
Yost, W. A. (1981). "Lateral position of sinusoids presented with intensive and temporal differences," J. Acoust. Soc. Am. 70, 397–409.
Yost, W. A., Nielsen, D. W., Tanis, D. C., and Bergert, B. (1974). "Tone-on-tone binaural masking with an antiphasic masker," Percept. Psychophys. 15, 233–237.
Young, S. R., and Rubel, E. W. (1983). "Frequency-specific projections of individual neurons in chick brainstem auditory nuclei," J. Neurosci. 3, 1373–1378.