THE ROLE OF CONSISTENCY OF INTERAURAL TIMING OVER FREQUENCY IN BINAURAL LATERALIZATION

Richard M. Stern
Department of Electrical and Computer Engineering
and Biomedical Engineering Program
Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

Constantine Trahiotis
Center for Neurological Sciences, Surgical Research Center,
and Department of Surgery (Otolaryngology)
University of Connecticut Health Center
Farmington, Connecticut 06032, USA

ABSTRACT

Recent experimental results indicate that consistency of interaural timing information plays an important role in the lateralization of broadband binaural stimuli. Specifically, it appears that the binaural system lateralizes stimuli from the centroids of their cross-correlation functions after peripheral bandpass filtering, and it weights more heavily the contributions of peaks in these functions that occur at internal delays that are consistent over a range of frequencies. In this paper we extend the *position-variable model* to incorporate a mechanism that effects this weighting, which we refer to as *straightness weighting*. Straightness weighting is introduced by assuming a second stage of processing which records coincidences of activity from outputs of primary coincidence-counting units with the same characteristic internal delay, but with different characteristic frequencies. We demonstrate that the extended position-variable model can describe the role that straightness plays in the lateralization of bandpass-noise stimuli.

KEYWORDS

Binaural hearing; "centrality"; cross-correlation; interaural temporal differences; lateralization; "straightness"; theories of hearing.

INTRODUCTION

Our understanding of phenomena that contribute to binaural lateralization and localization has increased greatly over the last few years. Classical theories of binaural interaction focussed upon the effects of interaural temporal differences (ITD), interaural phase differences (IPD), interaural intensitive differences (IID), and on the ability to locate sounds in the presence of echoes. More recent studies have evaluated the potency of these cues and their interaction as a function frequency. Many of these phenomena are discussed in excellent reviews by Durlach and Colburn (1978), Yost and Hafter (1987), and Zurek (1987). In this paper, we consider another important aspect of auditory lateralization: the consistency, across frequency, of interaural timing information. The purpose of this report is to describe how this attribute can be quantified in the context of cross-correlation-based models of binaural interaction which incorporate auditory-nerve response as inputs to a central processor, using a mechanism we refer to as *straightness weighting*.

We will begin by reviewing some of the perceptual data that motivated us to consider consistency of timing information as an important attribute of binaural image formation. We then discuss how current cross-correlation-based theories of binaural interaction fail to predict these experimental data. Finally, we will describe a new mechanism, based on the interaction of *coincident* responses from interaural-time-difference detectors across a range of characteristic frequencies, that is able to describe these data. This mechanism appears to be physiologically plausible.

LATERALIZATION OF BANDPASS NOISE: CENTRALITY AND STRAIGHTNESS

The importance of consistency of interaural timing information can be illustrated by reviewing data concerning lateral displacements of acoustic images obtained at our two laboratories. The curve with the solid symbols in Fig. 1 (from Stern *et al.*, 1988) describes the perceived lateral position of bandpass noise presented with a center frequency of 500 Hz, and an ITD of 1.5 ms (three quarters of the period of the 500-Hz signal). The noise is perceived toward the left side of the head (the side which receives the signal that is actually lagging) when the bandwidth is very narrow. But, as the bandwidth of the stimulus is increased, the acoustic image moves far across the midline toward the right side of the head (the side receiving the leading signal).

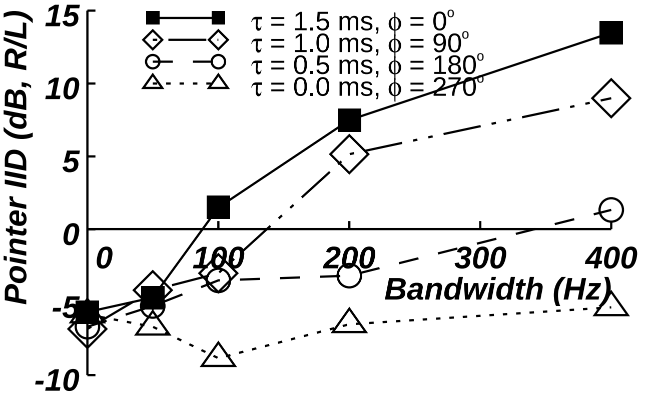

Fig. 1. Pointer IID needed to match the position of bandpass targets with center frequency 500 Hz and several combinations of ITD and IPD, as a function of bandwidth. (From Stern *et al.*, 1988.)

This phenomenon can be understood intuitively by considering some of the attributes of the stimuli. When the noise bandwidth is 50 Hz, the signal approximates a pure tone of 500 Hz with an ITD of +1.5 ms. This is physically equivalent to a 500-Hz tone presented with an ITD of -.5 ms, *i.e.* one which would normally be lateralized toward the left side of the head. As the bandwidth increases to 400 Hz, the stimulus becomes more "noise-like", and it is lateralized toward the right side, which is the signal that is actually leading in time. Let us consider the lateralization of each of these stimuli in more detail in the context of current models of binaural information processing.

Most modern theories of binaural interaction (*e.g.* Stern and Colburn, 1978; Blauert and Cobben, 1978; Lindemann, 1986) include at least two stages: a model of the response of auditory-nerve fibers to the stimuli and a model depicting how interaural time differences are encoded and extracted.

Our auditory-nerve model is based on the work of Siebert (1968), Colburn (1973), and their colleagues. It includes bandpass filtering, nonlinear rectification, lowpass filtering, and a mechanism which generates sample functions of Poisson processes with instantaneous rates that are proportional to the output of the lowpass filter.

Our model for interaural time comparison is based on the classical hypothesis of Jeffress (1948), which was later quantified and modified by Colburn (1973). This model assumes a matrix of coincidence-counting units that each receive impulses from two auditory-nerve fibers of the same characteristic frequency, one from each ear. Each unit includes a fixed internal interaural delay on one side, and coincident neural responses from the two input fibers are recorded after the internal delay. The coincident outputs can be characterized by two parameters: the characteristic frequency of the two fibers, and the value of the internal delay. The position-variable model (Stern and Colburn, 1978), as well as others (*e.g.* Lindemann, 1986), provides predictions concerning the lateral position of binaural stimuli by computing the centroid along the internal-delay axis of the total number of coincidence counts.

Lateralization of Narrowband Noise: Centrality

Figure 2 illustrates the relative response of these coincidence-counting units to bandpass noise with a center frequency of 500 Hz and a bandwidth of 50 Hz presented with an ITD of 1.5 ms. The horizontal axis corresponds to the fixed internal delay of the coincidence-counting units, and the oblique axis refers to the characteristic frequency of those units. The upper panel represents the relative number of coincidences that would be recorded per unit time by units with a given internal delay and characteristic frequency. As can be seen, a maximum appears at the internal delay of +1.5 ms, the stimulus ITD. Because the signal is tone-like in nature, this pattern repeats along the internal delay-axis with a period of 2 ms, producing a second row of maxima at -0.5 ms.

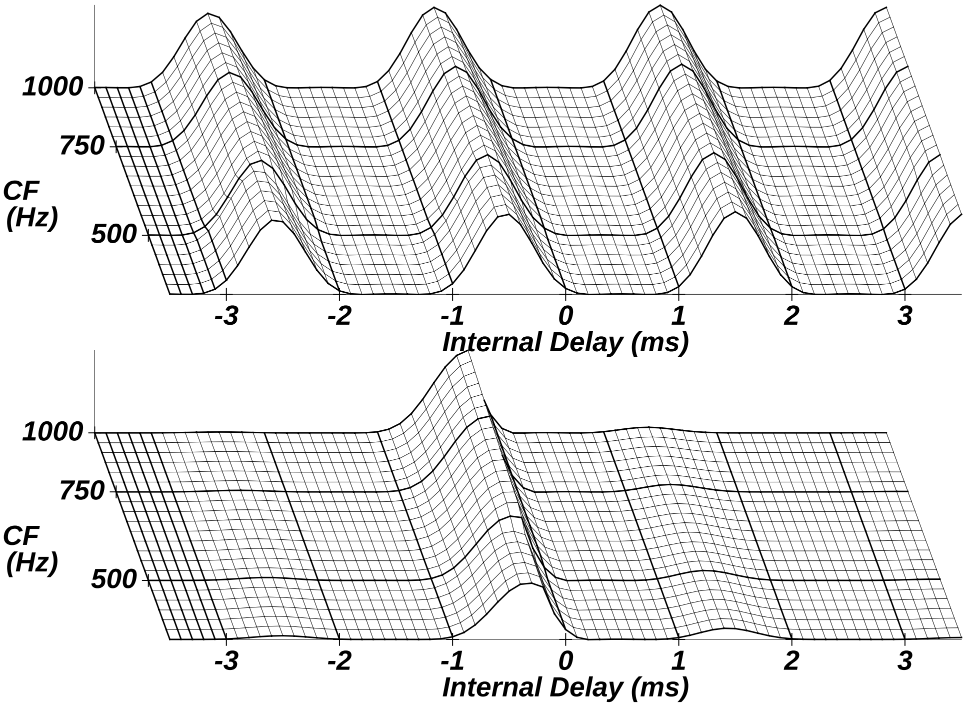

Fig. 2. Cross-correlation patterns showing the response of an ensemble of binaural fiber pairs to noise with center frequency 500 Hz and bandwidth 50 Hz. The horizontal axis indicates the internal interaural delay of the fiber pairs (in ms), and the oblique axis indicates the characteristic frequency of the auditory-nerve fibers (in Hz). The upper panel shows the average number of coincidences per fiber pair; the lower panel panel shows that average multiplied by the relative number of fiber pairs, producing the total number of coincidences over all fibers pairs.

The curves in the upper panel of Fig. 2 do not provide the whole story, however, because it is believed, on the basis of psychophysical and physiological evidence that there are more units with internal delays of smaller magnitude than with larger magnitude (*e.g.* Colburn, 1977; Shear, 1987; Kuwada, *et al.*, 1987). The curves in the lower panel represent the relative number of coincidences weighted by a function approximating the relative number of fiber pairs. The resulting function is an estimate of the *total* number of coincidences observed for all units in the matrix. The specific function utilized for the distribution of internal delays was developed by Colburn (1973, 1977) and refined by Shear (1987), and it appears to describe the lateralization of many low-frequency stimuli reasonably well (Stern and Shear, 1991).

We refer to this weighting by the relative number of fiber pairs as *centrality*, because it emphasizes those coincidence-counter outputs that are closest to the midline (0-ms delay). In this fashion, the position-variable model correctly predicts lateralization toward the apparently-lagging side of the head when the bandwidth of the noise is 50 Hz.

Lateralization of Broadband Noise: Straightness

Figure 3 describes the relative response of the coincidence counters to bandpass noise with center frequency 500 Hz and ITD 1.5 ms, but with a bandwidth of 400 Hz. Once more, maxima can be seen at 1.5 ms for all values of characteristic frequency, and the resulting pattern appears to be straight and parallel to the characteristic-frequency axis. The secondary ridges at interaural delays other than 1.5 ms are no longer parallel to the ridge at 1.5 ms because the stimulus has a wider bandwidth than the peripheral filters. As a consequence, these ridges are separated along the internal-delay axis by the reciprocal of the characteristic frequency of the peripheral bandpass filters, and not by the reciprocal of the 500-Hz center frequency of the stimulus.

The lower panel of Fig. 3 displays the total number of coincidences, once again taking into account the relative number of fiber pairs with a given internal delay. Because there are not many coincidence-counting units with an internal delay of 1.5 ms the straight ridge at that internal delay is greatly attenuated relative to the ridge that runs between zero and -0.5 ms. However, listeners actually lateralize according to the true stimulus ITD, +1.5 ms, which is indicated by the *straightness* of the ridge at +1.5 ms in the upper panel of Fig. 3.

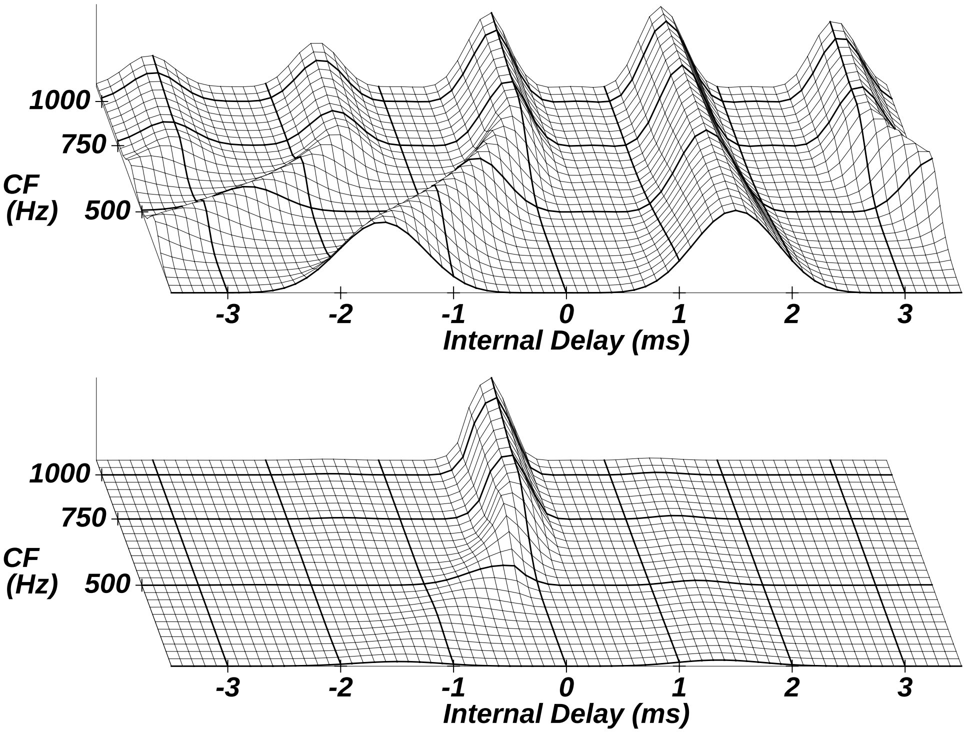

Fig. 3. Cross-correlation patterns, as in Fig. 2, showing the response of an ensemble of binaural fiber pairs to noise with center frequency 500 Hz and bandwidth 400 Hz.

Because of the small number of coincidence counters that are assumed to be present in the region of about 1.5 ms, the original position-variable model fails to describe the lateralization data shown in Fig. 1. As will be shown, the predicted position moves from left to right as bandwidth increases, but not nearly enough to depict the data. This means that a simple averaging of coincidences across frequency is not sufficient to predict the lateralization data, and that we need somehow to weight explicitly the "straight" trajectories more heavily in order to predict the observed lateral position.

The need to incorporate straightness in the lateralization process was confirmed by measuring the lateral position of stimuli containing combinations of ITDs and IPDs that were selected to maintain fixed centrality while straightness was manipulated as a parameter. The curves in Fig. 4 show the locations of peaks of the cross-correlation functions of bandpass-filtered stimuli, again plotted as a function of internal delay and characteristic frequency for these combinations of ITD and IPD. As discussed in Stern *et al.* (1988), these four combinations of ITD and IPD provide equivalent centrality but progressively decreasing straightness at +1.5 ms as the stimulus parameters are changed from 1.5 ms and 0 degrees to 0 ms and 270 degrees. The lateralization data in Fig. 1 (taken from that study) demonstrate that as the amount of straightness in the trajectory at +1.5 ms decreases, the amount by which the image is "pulled" over to the right-hand side of the head for larger bandwidths decreases as well.

In general, we have shown (Stern, *et al.*, 1988; Trahiotis and Stern, 1989) that the actual location of a perceived image results from a competition between straightness and centrality. If all other factors

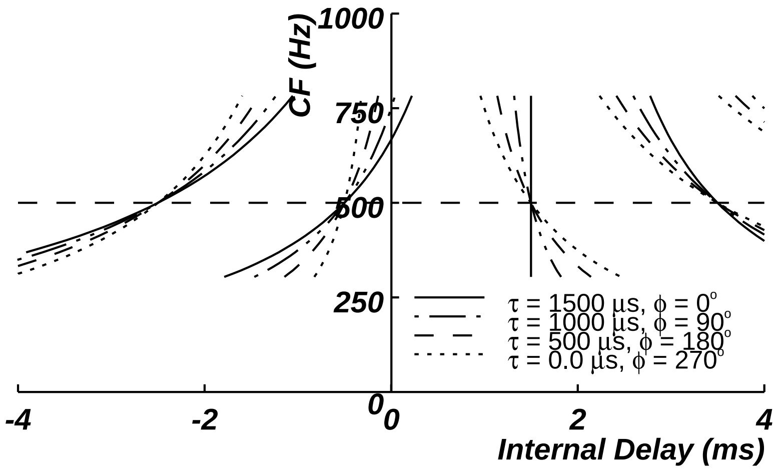

Fig. 4. Loci of the peaks of the cross-correlation functions for broadband noise after peripheral auditory filtering for the several combinations of ITD τ and IPD φ used for the data shown in Fig. 1. The vertical axis indicates characteristic frequency and the horizontal axis indicates the argument of the cross-correlation function. The combinations of ITD and IPD were selected to produce maxima at 1.5 ms and 500 Hz in each case. (From Stern *et al.*, 1978.)

are equal, we believe that images will be lateralized toward the side of the head with the straightest ridge of coincidences. On the other hand, if all ridges are equally straight, lateralization would be dominated by centrality. This makes sense intuitively: real sound sources produce ITDs that are consistent over frequency. By weighting more heavily those values of internal delay in the model that appear to be responding to stimulus ITDs that are consistent over frequency, we are weighting more heavily the response to possible sound sources that are more physically plausible.

A COINCIDENCE-BASED MODEL THAT PREDICTS STRAIGHTNESS WEIGHTING

We now consider how we can implement the straightness-weighting factor more suitably while retaining the general form of the position-variable model (Stern and Colburn, 1978).

As stated earlier, we found that simply summing the contributions of coincidence counters over frequency fails to describe adequately how straightness affects lateralization. We considered several strategies in order to derive a more effective weighting of straight modes of the cross-correlation function. Two of the unsuccessful methods that we tried include increasing the expansiveness of the nonlinear rectifier in the model of peripheral auditory-nerve activity and, separately, the use of the derivative with respect to frequency of the two-dimensional cross-correlation function as a measure of straightness. (These algorithms, and other intuitively-plausible possibilities that failed to predict the lateralization data are described by Shear, 1987.)

More recently, we have developed predictions from a straightness-weighting algorithm that appears to be much more promising. We continue to assume the Jeffress/Colburn model for the extraction of interaural timing information. In addition, we also assume a *second* level of coincidence-counting units that take as inputs the outputs from a small number of first-level coincidence-counters. Each set of inputs is assumed to come from first-level coincidence counters representing a range of characteristic frequencies, but with a common internal delay. The sets of points denoted by the filled circles in the upper panel of Fig. 5 are examples of combinations of characteristic frequency and internal delay that would comprise inputs to the second-level coincidence counters. The lower panel of Fig. 5 shows the dramatic effects of applying this second level of coincidence weighting on cross-correlation functions representing neural responses to bandpass noise with a 1.5-ms ITD. Comparing the lower panels of Figs. 3 and 5, it is seen that this second level of coincidences provides much greater emphasis to the straight trajectory at 1.5 ms. This occurs because, for that trajectory, all of the first-level coincidence counters are firing at a level that is at or near their maxima. In contrast, the ridge closer to the midline is attenuated because of the minimal response for characteristic frequencies below approximately 600 Hz. In addition, this manner of weighting straightness also sharpens the ridges of the two-dimensional cross-correlation function along the internal-delay axis without an explicit mechanism for inhibition such as those postulated

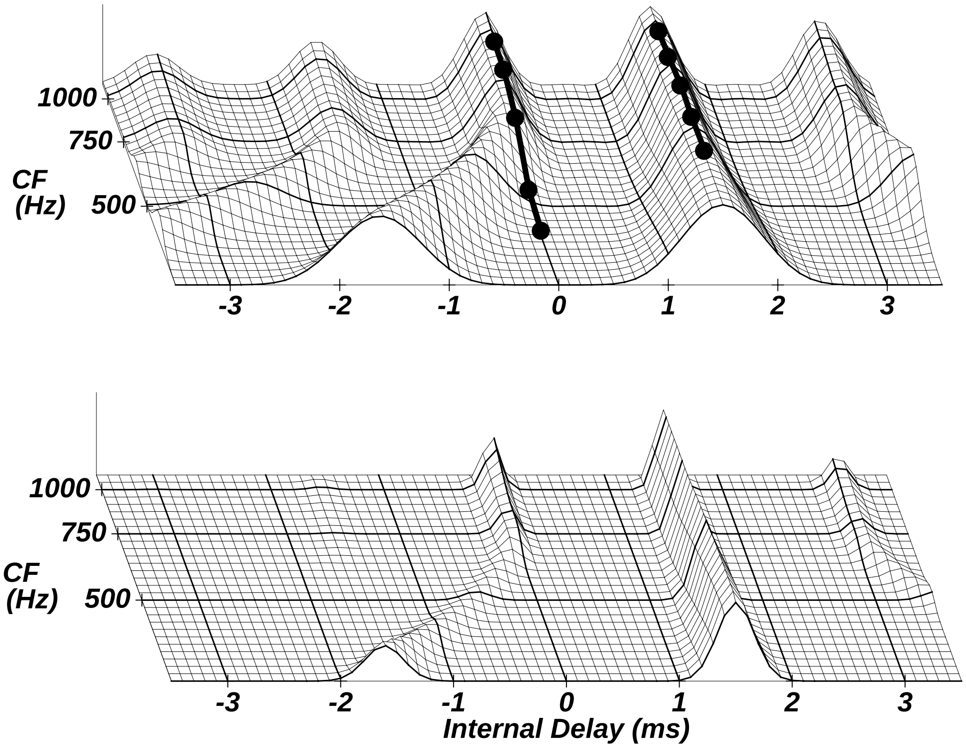

Fig. 5. Upper panel: cross-correlation patterns, as in Fig. 2, showing the response of an ensemble of binaural fiber pairs to noise with center frequency 500 Hz and bandwidth 400 Hz. Locations of constant internal delay but different characteristic frequency are identified by filled circles joined by lines. Lower panel: the total number of coincidences observed by second-level units which compute coincidences over frequency of the outputs of the original coincidence counters.

by Blauert and Cobben (1978) and Lindemann (1986).

Figure 6 contains predictions for the complete set of lateralization data described in Fig. 1, both with (left panel) and without (right panel) the second-level mechanism that provides straightness weighting by a second set of coincidence counters. Comparing the predictions in the right panel of Fig. 6 to the data of Fig. 1, it can be seen that this weighting function results in predicted images that are, in fact, pulled over to the right side of the head as bandwidth increases in the same fashion that the data.

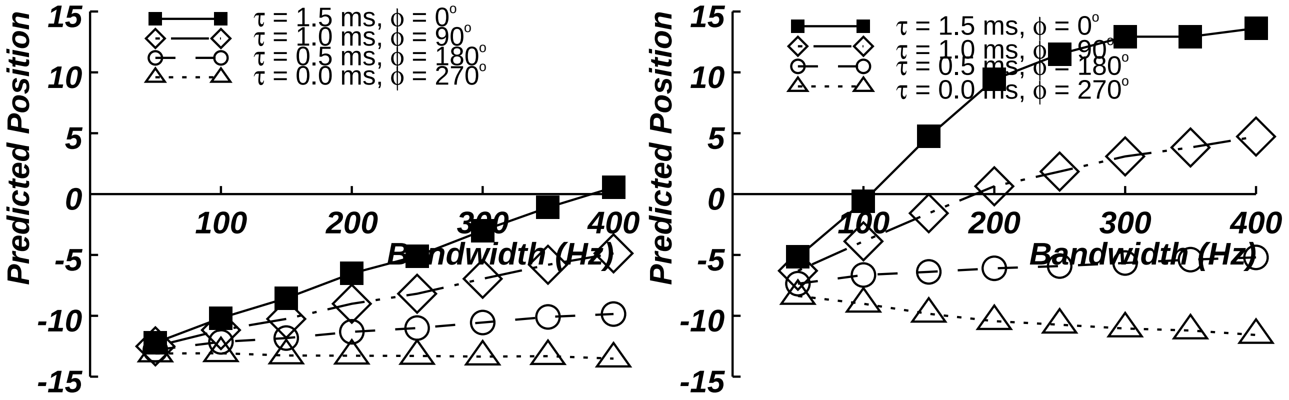

Fig. 6. Predictions of the position-variable model for the data of Fig. 1 without (left panel) and with (right panel) straightness weighting.)

Support for such a second-level coincidence mechanism is provided by physiological findings of Takahashi and Konishi (1986). They compared the responses of interaural-time-sensitive units in the inferior colliculus in the barn owl to a single tone at the best frequency, to a tone of a second frequency, and to the two tones presented simultaneously. Takahashi and Konishi found that units tuned to the same interaural delay over a range of stimulus frequencies produced responses to the simultaneous presentation of the two tones that were greater than the sum of the responses to the each of the tones presented in isolation. In other words, the response to a tone presented at the best frequency with the "best" ITD is facilitated by the presentation of a second tone with that same ITD. This is exactly the type of response that would be predicted by mechanisms like the second-level coincidences across frequency that we propose.

SUMMARY

In summary, we believe that the grouping and emphasis of attributes of binaural signals that are consistent over frequency is a very important component in the formation, identification, and lateralization of spatial images. Lateralization data strongly indicate that "straight" trajectories in the internal delay-characteristic frequency plane greatly affect where we hear sounds. We have argued that the position-variable model (as well as all others), must explicitly incorporate a mechanism for straightness weighting in order to describe the observed psychophysical data. This weighting can be accomplished by postulating an additional set of coincidences that integrate information, across frequency, from the classical first-level interaural coincidence-counting units in the Jeffress/Colburn model.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant BNS-87099349), the National Institutes of Health (Grant DC 00234), and the Air Force Office of Scientific Research (Grant AFOSR 89-003). We also thank G. Shear, S. Tao, and X. Xu for their help in implementing the predictions of the model.

REFERENCES

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

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. Amer.*, **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. Amer.*, **61**, 525-533.

Durlach, N. I., and Colburn, H. S. (1978). Binaural perception. In *Handbook of Perception, Vol. 4. Hearing* (E. C. Carterette and M. P. Friedman (Eds.), pp. 365-466, Academic Press, NY.

Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.*, **41**, 35-39.

Kuwada, S., Stanford, T. R., and Batra, R. (1987). Interaural phase-sensitive units in the inferior colliculus of the unanaesthetized rabbit: effects of changing frequency. *J. Neurophysiol.*, **57**, 1338-1360.

Lindemann, W. (1986). Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *J. Acoust. Soc. Amer.*, **80**, 1608-1622.

Shear, G. D. (1987). *Modeling the Dependence of Auditory Lateralization on Frequency and Bandwidth*. Master's thesis, Elec. and Comp. Eng. Dept., CMU.

Siebert, W. M. (1968). Stimulus transformations in the peripheral auditory system. In *Recognizing Patterns* (P. Kolers and M. Eden, eds.). MIT Press, Cambridge, Mass.

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. Amer.*, **64**, 127-140.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (1988). Lateralization of complex binaural stimuli: A weighted image model. *J. Acoust. Soc. Amer.*, **84**, 156-165.

Stern, R. M., and Shear, G. D. (1991). Lateralization of broadband and complex stimuli: Predictions of the position-variable model. *J. Acoust. Soc. Amer.*, (in preparation).

Takahashi, T. A., and Konishi, M. (1986). Selectivity for interaural time differences in the owl's midbrain. *J. Neurosci.*, **6**, 3413-3422.

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. Amer.*, **86**, 1285-1293.

Yost, W. A., and Hafter, E. R. (1987). Lateralization. In *Directional Hearing* (W. A. Yost G. Gourevitch, Eds.), pp. 49-84. Springer-Verlag.

Zurek, P. (1987). The precedence effect. In *Directional Hearing* (W. A. Yost G. Gourevitch, Eds.), pp. 85-105. Springer-Verlag.