

Lateralization of two-transient stimuli

PATRICK M. ZUREK

Massachusetts Institute of Technology, Cambridge, Massachusetts

and

KOUROSH SABERI

University of California, Irvine, California

In this study of the precedence effect in binaural hearing, subjects adjusted the interaural delay of a wideband acoustic pointer to match the perceived intracranial position of transient test stimuli presented over headphones. The test stimuli had leading and lagging components (either brief noise bursts or clicks), each with its own interaural delay. In some test conditions, the leading and lagging stimuli were coherent copies of one another, whereas in others, they were independent samples of noise. The duration of the stimuli and the delay from the leading component to the lagging component were also varied. All the stimulus conditions showed a moderate or strong precedence effect (i.e., covariation of perceived lateral position of the composite two-transient stimulus with the interaural delay of the leading component). Predictions of the lateralization data are presented for variants of models based on temporal weighting and/or bandpass correlation. In one model variant, the binaural stimuli are temporally weighted to emphasize the onset and then subjected to bandpass correlation analysis. In another variant, it is assumed that the onset mechanism provides a rough estimate of the initial interaural delay that guides a slower and more focused bandpass correlation analysis. The accuracies of these two model's predictions were equivalent and superior to those of models that either represent leading and lagging cues equally (bandpass correlation with no onset effect) or do not represent lagging cues at all (a complete precedence effect). The results of these analyses show the need for both a strong onset effect and for bandpass correlation analysis and suggest two modeling approaches for achieving that goal.

Many studies in binaural hearing have shown that the interaural cues of earlier-arriving stimulus components dominate later ones, at least when the stimulus possesses transients and the components fuse into a single image. Such findings are usually associated with the precedence effect in sound localization (Wallach, Newman, & Rosenzweig, 1949; see Zurek, 1987, for a review). The precedence effect has been the topic of extensive physiological (Cranford & Oberholtzer, 1976; Yin & Litovsky, 1995), developmental (Clifton, Morrongiello, & Dowd, 1984; Morrongiello, Clifton, & Kulig, 1982), applied (Blauert, 1989; Muncey, Nickson, & Dubout, 1953; Snow, 1954), and clinical (Ashmead, Hill, & Taylor, 1989; Ashmead et al., 1998; Hochster & Kelly, 1981) study within the hearing sciences.

Despite this extensive study, the mechanisms that underlie the precedence effect are not well understood. All theoretical conceptions of the precedence effect have naturally involved a mechanism that emphasizes interaural cues at transient onsets. However, proposals for the site

of that mechanism have ranged from adaptation of hair cell response (Hartung & Trahiotis, 2001), to inhibition along mid-level delay line neurons (Lindemann, 1986a, 1986b), to central temporal gating and inhibition mechanisms (Harris, Flanagan, & Watson, 1963; Zurek, 1987). Furthermore, such mechanisms have not been firmly incorporated into the cross-correlation models that account well for a wide range of binaural phenomena observed with long-duration stimuli having steady interaural differences (Colburn, 1996; Colburn & Durlach, 1978). Only in a few studies has there been an attempt to integrate the temporal-weighting mechanisms motivated by the precedence effect into cross-correlation models (Hartung & Trahiotis, 2001; Lindemann, 1986a, 1986b), and very limited sets of data were examined.

The present study expands the effort to integrate temporal-weighting and cross-correlation mechanisms by examining the predictions of various temporal-weighting/cross-correlation models for an extensive set of lateralization data obtained with stimuli composed of two binaural transients. Measurements of subjective lateral position were obtained for two-transient stimuli that varied in temporal and spectral parameters and in the degree of coherence between the leading and the lagging transients. These stimulus manipulations produce a wide variation in the parameters relevant to both binaural temporal-weighting and cross-correlation mechanisms.

This work was supported by Grants R01 DC00100 and R29 DC03648 from the National Institute on Deafness and Other Communication Disorders. The authors thank Steve Colburn and Richard Freyman for their comments on an earlier draft of this paper. Correspondence concerning this article should be addressed to P. M. Zurek, Sensimetrics Corporation, 48 Grove St., Somerville, MA 02144 (e-mail: pat@sens.com).

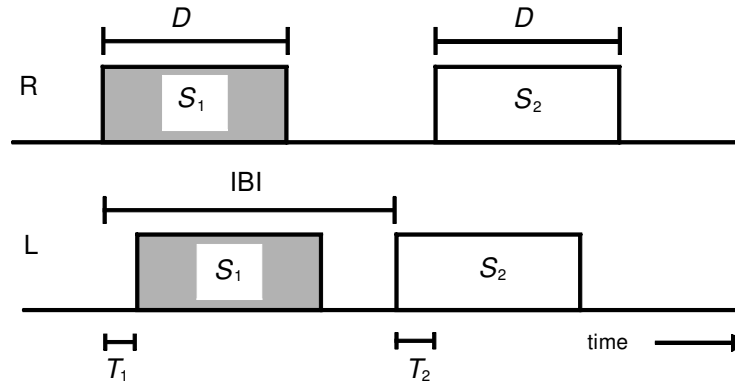


Figure 1. Schematic diagram of the two-transient stimuli.

METHOD

Stimuli

Figure 1 shows a schematic diagram of the stimuli, which in all cases were composed of two binaural transients, S_1 and S_2 . In most conditions, S_1 and S_2 were noise bursts, whereas in a few other conditions, they were brief clicks. In all cases, the left and the right components of S_1 were identical between the ears except for the interaural delay T_1 . Likewise, the left and the right components of S_2 were identical between the ears except for the interaural delay T_2 . When the stimuli were composed of noise bursts, S_2 was, in some conditions, a coherent copy of S_1 , and in other conditions, it was an incoherent (independent) sample of noise. The onset of S_1 preceded that of S_2 by the interburst interval (IBI). The IBI thus was determined from the leading edge of the earlier component of the first binaural stimulus to the leading edge of the earlier component of the second stimulus. The common duration of the components was D .

Table 1 summarizes the 14 test conditions used in this study. Most conditions employed noise bursts, and most of these used incoherent leading and lagging bursts. Conditions 7–10 used incoherent noise bursts that were either low-pass filtered at 1.2 kHz or high-pass filtered at 2.5 kHz. For the wideband conditions, stimuli were low-pass filtered at 20 kHz. In all cases, the slopes of the filters outside of the passband were greater than 60 dB/oct. The duration, D , for the filtered stimulus conditions is the duration of the white noise prior to filtering. Conditions 11 and 12 used 100- μ sec rectangular clicks. A subset of the conditions was tested on three occasions; the remainder were tested only once, owing to time constraints and because repetitions produced consistent results. A control condition was also tested in which the stimulus contained only a single 1-msec wideband noise burst. This condition assessed the subjects' precision in matching the pointer to itself.

These 14 test conditions were selected to provide a wide range of stimuli for which the precedence effect has been reported. The precedence effect with transients is elicited to varying degrees with IBIs ranging from 0.3 to 3 msec, with maximal effect at IBI \approx 1 msec. IBIs less than 1 msec were used to sample the *summing localization* range (Blauert, 1983). Duration D ranged from 100 μ sec (clicks) to 3 msec for noise bursts. Filtered stimuli were used to sample known variations in the precedence effect owing to spectral content (Saber, 1996; Shinn-Cunningham, Zurek, Durlach, & Clifton, 1995).

Within each condition, 40 (four sets of 10) stimuli were generated with four sets of frozen-noise tokens and 10 combinations of T_1 and T_2 . These 10 (T_1, T_2) combinations were (400, -400), (200, -200), (100, -100), (200, 0), (0, 200), (-400, 400), (-200, 200), (-100, 100), (-200, 0), and (0, -200). Note that Stimuli 6–10 in a set of 10 are the left–right reversals of Stimuli 1–5. Positive val-

ues indicate that the right stimulus led the left. New tokens were generated for each condition.

For each condition in which the stimulus was composed of noise, four sets of noise tokens were employed. Stimuli 1–10 were all generated by appropriately time-shifting the S_1 and S_2 tokens to achieve the desired T_1 and T_2 combinations. The second 10 stimuli were generated in the same way with a new set of S_1 and S_2 tokens, and similarly for the third and fourth sets of 10 stimuli. This resulted in four replications of the same stimulus conditions, but with different tokens of noise. The use of four different tokens within each condition was intended to assess the variation in performance with noise stimuli having nominally identical parameters but different waveforms.

Stimuli were generated at a sampling rate of 40 kHz, were low-pass filtered, and were presented to subjects through TDH-49 earphones in circumaural cushions. All the stimuli were scaled to have the same absolute peak electrical amplitude of 1 v. As a result, the spectral levels of the noises depended on their bandwidth. For wideband and high-pass-filtered noise, the spectrum level was 55 dB SPL; for low-pass-filtered noise, it was 68 dB SPL. Clicks were presented at a peak amplitude of about 6 Pa.

Procedure

The perceived intracranial lateral position of a test stimulus was measured with an acoustic pointer. A test stimulus was repeated at a rate of about 1/sec. Using a keypad, the subject could switch be-

Table 1
Conditions and Parameter Values for the 14 Test Conditions

Condition	D (msec)	IBI (msec)	C/I	Signal	Replicates
1	1.0	0.3	I	WBN	3
2	1.0	0.6	I	WBN	3
3	1.0	1.0	I	WBN	3
4	1.0	3.0	I	WBN	1
5	3.0	1.0	I	WBN	3
6	3.0	3.0	I	WBN	1
7	1.0	1.0	I	LPN	3
8	1.0	3.0	I	LPN	1
9	1.0	1.0	I	HPN	3
10	1.0	3.0	I	HPN	3
11	0.1	1.0	C	Clicks	1
12	0.1	3.0	C	Clicks	1
13	1.0	1.0	C	WBN	1
14	3.0	3.0	C	WBN	1

Note— D , duration; IBI, interburst interval; C/I, coherent/incoherent; WBN, wideband noise; LPN, low-pass-filtered noise; HPN, high-pass-filtered noise; Replicates, number of replications.

tween this repeating test stimulus and a pointer, which was presented with a different repetition rate to distinguish it from the test. The pointer was a single 1-msec burst of white noise with an interaural delay that was under the subject's control. The subject adjusted the pointer's interaural delay to match its lateral position to that of the test stimulus, terminating the trial with a keypress. The range of pointer delays available to the subject was -600 to $600 \mu\text{sec}$; the smallest step size was $25 \mu\text{sec}$. Simulations showed that the variance added from this step size was negligible ($\sigma = 12.5 \mu\text{sec}$), as compared with the pointer adjustment variance.

A run consisted of presentation of the 40 stimuli in a condition, in random order. The order in which conditions were tested was also randomized. Approximately 15 min was required for a subject to make adjustments to all 40 stimuli in a condition. The subjects typically completed six runs in a daily listening session. Those conditions that were replicated were done so on different visits to the lab. Three young adults served as subjects. None had a history or complaints of hearing problems. They were paid for their time.

RESULTS

Example results, for Condition 3, are shown in Figure 2, where pointer adjustments averaged over replications for the 3 subjects are plotted for the 40 stimuli. The patterns of T_1 and T_2 variation across the stimuli are illustrated by the solid and the dashed lines, respectively. The extent to which pointer adjustments are closer to T_1 than to T_2 is a measure of the precedence effect. Differences in the apparent strength of the precedence effect, using different noise tokens with identical stimulus parameters, are seen by comparing responses to stimuli that differ by multiples of 10. For example, Stimuli 1, 11, 21, and 31 all had $T_1 = 400 \mu\text{sec}$ and $T_2 = -400 \mu\text{sec}$, but different noise tokens.

The subjects' reliability in adjusting the pointer was reasonably good. Across all replicated conditions, the replication standard deviation ranged from 51 to $101 \mu\text{sec}$ and averaged 79 , 67 , and $77 \mu\text{sec}$ for the 3 subjects. When replication standard deviations were plotted against mean pointer adjustment, there was no visually apparent dependence of response variability on mean response. The standard deviation among the 3 subjects' mean responses was $64 \mu\text{sec}$. Results from the control condition revealed that the subjects adjusted the pointer to itself with good accuracy and consistency. A scatterplot of responses versus stimulus interaural delay was fit with a line with slope = 0.99 and with $r^2 = .92$.

The entire set of pointer adjustments was tested for left/right asymmetry by adding the adjustments made to left/right reversed, but otherwise identical, test stimuli. Since half of the stimuli in a 10-stimulus set were left-right reversals of the other half, a true lack of asymmetry would yield a distribution of these sums with a mean close to zero. None of the 3 subjects' means differed significantly from zero, suggesting that there was relatively little or no gross asymmetry in the lateral position of these test stimuli.

The strength of the precedence effect can be summarized with a metric that describes the apparent relative weighting of the leading and lagging interaural delays

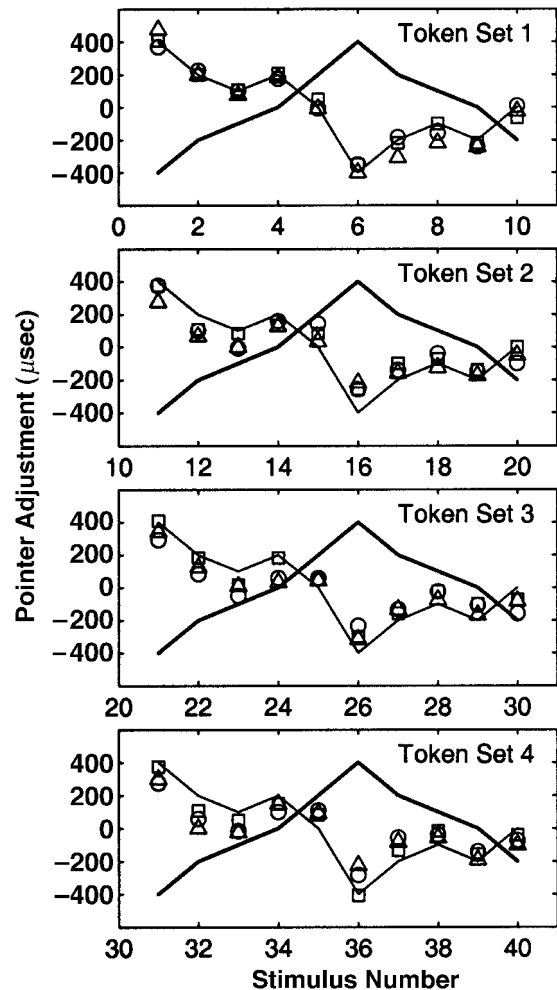


Figure 2. Example pointer adjustments from Condition 3. Data points are the averages over three replications for each of the 3 subjects. The thin line is the interaural delay of the leading transient, T_1 , and the bold line that of the lagging transient, T_2 .

(Shinn-Cunningham, Zurek, & Durlach, 1993). This metric is found by assuming that a pointer adjustment, α , is the weighted sum of the two interaural delays,

$$\alpha = cT_1 + (1-c)T_2, \quad (1)$$

resulting in the weight c being determined by

$$c = \frac{\alpha - T_2}{T_1 - T_2}. \quad (2)$$

Note that c is the relative distance of the pointer adjustment between T_1 and T_2 . A value of $c = 1$ corresponds to a pointer adjustment coincident with T_1 , whereas $c = 0$ corresponds to a pointer adjustment coincident with T_2 . A value of $c = .5$ suggests equal weighting of the two interaural delays.

Values of c for the 14 test conditions are shown in the two panels of Figure 3. Mean c (shown by the filled sym-

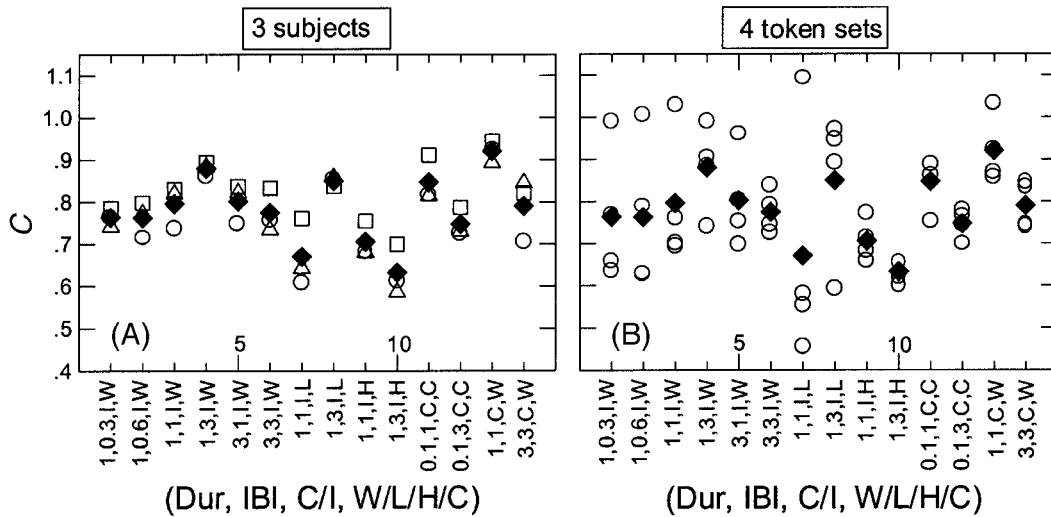


Figure 3. (A) Values of c averaged over replications and the 40 stimuli in the 14 test conditions. Open symbols distinguish the 3 subjects; the closed symbols denote the means. (B) Values of c averaged over replications, subjects, and the 10 stimuli in a token set. Open symbols distinguish the four token sets; the closed symbols denote the means. In both panels, the abscissa shows conditions, with parameter values summarized in the labels. The first entry in each label is duration (Dur) in milliseconds, the second is interburst interval (IBI) in milliseconds, the third is coherent versus incoherent (C/I), and the fourth is the type of signal (W, wideband noise; L, low-pass-filtered noise; H, high-pass-filtered noise; C, click).

bols in either panel) varied over the 14 conditions from .64 to .92, with an overall mean of .78. Two of the lowest c s were obtained in Conditions 9 and 10, which used noise high-pass filtered at 2.5 kHz, a result that will be discussed further below. A relatively low c was also obtained in one of the low-pass-filtered conditions (7), but not in the other (8).

There was no clear superiority in c for either coherent or incoherent stimuli. Although the mean c for Conditions 11–14 (which used coherent stimuli) was larger (.83) than that for the remaining conditions (.76), the variance that was due to tokens was too large (see below) to make such a difference significant.

It is worth noting that a substantial dominance of the initial time delay ($c = .77$) was seen even with an IBI as short as 300 μ sec. Previous studies suggested the existence of a range of *summing localization* (Blauert, 1983), a time interval over which interaural cues contribute equally regardless of their temporal order.

As is suggested by the example data in Figure 2, the subjects were fairly consistent in their adjustments and, hence, in the resulting values of c . The open symbols in Figure 3A show c for the 3 subjects and 14 conditions averaged over the 40 stimuli in each condition. The average standard deviation among the subjects' c values was .05, indicating close agreement among the subjects. A difference among listeners is suggested, however, by the generally higher c s for one subject (symbolized by squares) than for another (circles).

There was substantial variability owing to noise tokens, as can be seen in the example raw data of Figure 2. Adjustments to Stimulus 23, for example, with $T_1 =$

100 μ sec and $T_2 = -100 \mu$ sec, averaged slightly less than 0 μ sec, whereas those to Stimulus 3, with the same values of T_1 and T_2 , were close to T_1 . Such differences among token sets are summarized for all the conditions in Figure 3B, which plots c separately for the four 10-stimulus subsets that used the same stimulus tokens, along with the average c for each condition. These data points represent averages of c across subjects and the 10 (T_1 , T_2) combinations in a set of tokens. The size of the variance in c owing to tokens appears to be related to the type of stimulus used. The smallest variance owing to tokens was obtained with the high-pass conditions (9, 10) and the conditions using coherent stimuli (12–14), whereas the largest variance was seen with the incoherent-noise conditions, especially the low-pass-filtered conditions (7 and 8) and the 1-msec duration conditions (1–4). This trend is roughly consistent with the increased relative variance in the overall power of noise samples as the time-bandwidth product is reduced.

THEORETICAL ANALYSIS

The present data, which were obtained with a wide range of parameter variation for transient stimuli, provide an opportunity to explore modifications for temporal weighting to current black-box models of binaural hearing. Four versions of a bandpass correlation model of binaural hearing are examined here. The four versions will be described after the basic elements of the model, which are common to all versions, have been described.

The basic bandpass correlation model implemented here incorporates the features of cross-correlation mod-

Table 2
The q Weighting Function, From Stern, Zeiberg, and Trahiotis (1988), at the Third-Octave Center Frequencies

f (Hz)	q	f (Hz)	q
100	0.024	400	0.584
125	0.037	500	0.861
157	0.059	630	1.0
200	0.104	800	0.800
250	0.184	1000	0.472
315	0.333	≥ 1250	0.310

els that have evolved over the years to account for a variety of steady-state lateralization data (e.g., Colburn, 1996; Licklider, 1959; Sayers & Cherry, 1957; Stern & Colburn, 1978). The core of the processing is the calculation of cross-correlation functions between counterpart left and right bandpass signals. An estimate of lateral position for a wideband stimulus is derived from the pattern of correlation versus frequency and delay, after weighting in both the delay and the frequency dimensions to emphasize dominant regions.

The present model processes the signals in the following way to arrive at a single wideband estimate of lateral position.

1. The entire $S_1 + S_2$ stimuli in the L and R channels are passed through identical 1/3-octave filter banks, with center frequencies ranging from 100 to 8000 Hz. This filter bank provides a crude simulation of auditory filters.

2. In bands with a center frequency below 1500 Hz, the bandpass L and R signals are cross-correlated. In this way, the correlation is based on stimulus fine structure in these bands.

3. In bands with a center frequency above 1500 Hz, the envelopes of the bandpass L and R signals are extracted (by taking the magnitude of the Hilbert transform), and these are cross-correlated. Cross-correlation in these bands is thus made to reflect interaural delays in the bandpass envelopes, not in the fine structure, in agreement with psychophysical data.

4. A weighting function is applied to each bandpass cross-correlation function along the correlation delay dimension. This function, p , applies uniform maximal weighting over the range between -300 and $+300$ μ sec, with exponential tapering beyond those limits (Stern & Colburn, 1978). A rectangular window is also applied to the cross-correlation functions from -1 to $+1$ msec to eliminate contributions outside of that range. These weightings reflect the relatively more acute sensitivity near zero delay and the irrelevance of delays beyond 1 msec.

5. The bandpass cross-correlation functions are weighted by a frequency-weighting function, q , as de-

scribed by Stern, Zeiberg, and Trahiotis (1988) and by a $1/\text{frequency}$ factor to remove the effect of increasing bandwidth. The q weighting function, which is given in Table 2, roughly reflects the frequency dependence of human sensitivity to interaural delay cues.

6. The cross-correlation functions are *partially normalized* prior to the wideband estimation of interaural delay. This partial normalization is best explained if we have an expression for the processing in terms of the operations on the discrete-time signals.

Let the subscript i indicate the i th third-octave band, and L_i and R_i the left and right bandpass-filtered signals in band i . Further, define \tilde{L}_i and \tilde{R}_i to be the left and the right signals that are cross-correlated in band i (i.e., the bandpass signals at low frequencies and the envelopes of the bandpass signals at high frequencies). Then, if we let n represent running discrete time and k be the discrete-time correlation delay variable, the weighted bandpass cross-correlation function, $r_i(k)$, is shown at the bottom of the page as Equation 3, where all the sums are taken over running time and where $p(k)$ is the correlation-delay weighting, f_i is the center frequency, and q_i is the weighting from Stern et al. (1988) for the i th third-octave band.

The power-normalization factors (in the radicals) were chosen, in conjunction with the $1/f$ factor, in an experimental process in order to solve the problem of how bandpass correlations contribute to a wideband estimate of interaural delay. Ordinarily, when one speaks of the "correlation" between two signals, one refers to a power-normalized quantity that can vary between -1 and 1 . The quantity in square brackets on the right-hand side of Equation 3 is such a normalized sample correlation. However, it would be clearly unreasonable to use normalized correlations for combining information across frequency bands, because there could be no effect of stimulus level. Bands with low-level signals would make contributions equal to those with high-level signals.

At the other extreme, one could use the cross-product of the signals with no normalization, which would be given by the numerator of the square-bracket term with neither of the normalization factors in radicals present. With this approach, each band would make a contribution proportional to its signal power. This approach is attractive, but it has the problem with broadband signals, such as those used here, that it gives undue weight to the high-frequency region where most of the power resides. Psychophysical studies (e.g., Wightman & Kistler, 1992) have shown that when interaural delay cues conflict across bands, the low-frequency cues dominate.

The 4th-root power-normalization term in Equation 3 is part way between the two extremes of full normalization

$$r_i(k) = \left[\frac{\sum \tilde{L}_i(n) \tilde{R}_i(n+k)}{\sqrt{\sum \tilde{L}_i^2(n) \sum \tilde{R}_i^2(n)}} \right] \sqrt[4]{\sum L_i^2(n) \sum R_i^2(n) p(k) q_i / f_i} \quad (3)$$

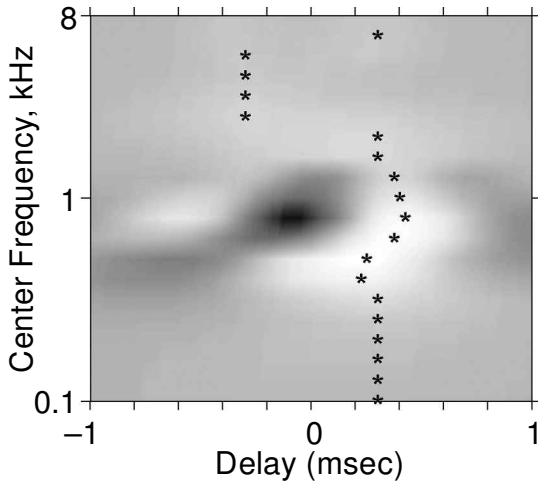


Figure 4. An example showing the bandpass correlation model's cross-correlation functions and maxima (using Stimulus 1 from Condition 3: $T_1 = 400 \mu\text{sec}$ and $T_2 = -400 \mu\text{sec}$). The cross-correlation function in any band is represented by a horizontal strip, with gray value indicating the degree of correlation (with white indicating high correlation). The maximum of the cross-correlation function in each of the 20 bands is indicated by an asterisk.

and no normalization. The combination of this normalization and the $1/f$ and q factors was found in pilot studies to give an acceptable compromise between the sensitivity to high frequencies and the sensitivity to low-level bands.

7. The global maximum of the weighted cross-correlation function in each band is estimated, and a weighted least-squares line of constant delay is fit to these maxima. The weight for each frequency band is proportional to the magnitude of the peak of the cross-correlation function, a choice that was made on intuitive grounds. The best-fitting delay is the model's estimate of the lateral position that would be measured by interaural pointer delay. Quantitatively, if in band i the maximum value of $r_i(k)$ is r_i^* and it is achieved at k_i^* , the wideband estimate, k^* , of interaural discrete-time shift is

$$k^* = \frac{\sum_i r_i^* k_i^*}{\sum_i r_i^*}. \quad (4)$$

The estimated interaural delay was gotten by dividing k^* by the sample frequency.

Figure 4 shows an example display in the frequency delay plane of weighted bandpass cross-correlation functions, along with the maximum in each band, using Stimulus 1 of Condition 3, which has $T_1 = 400 \mu\text{sec}$ and $T_2 = -400 \mu\text{sec}$. The model's estimate of interaural delay for this stimulus was $259 \mu\text{sec}$; the average adjustment was $416 \mu\text{sec}$. It should be noted that the lack of distinct peaks in both the low- and the high-frequency regions in this figure results primarily from the frequency weightings applied. A secondary factor that deemphasizes the peaks in the low- and high-frequency regions is that the

bandpass signals at low frequencies and the bandpass envelopes at high frequencies have relatively broad cross-correlation functions.

Four versions of the cross-correlation model were implemented.

BPC. This version is the basic bandpass correlation model just described.

BPC+TW. This is a modified version of the BPC model that includes a temporal weighting of the L and R input signals. The form of this temporal weighting is shown in Figure 5. The weighting function has a value of unity from time = 0 to T_{ON} and a value that declines exponentially with time constant k thereafter. In applying this weighting function to the stimuli, the start of the weighting function was aligned with the start of the stimulus (the first nonzero sample) to each ear. The weighted signals were then presented to the bandpass correlation model as just described. Note that a single weighting function is applied to each wideband signal, and that there is no band-dependent temporal weighting. Furthermore, because only brief-duration stimuli are considered here, the shape of the weighting function is specified only over a few milliseconds; a more complete temporal weighting model would have to be specified in order to predict results with longer stimuli. The choice of the parameters T_{ON} and k is described below.

BPC/RG. This model operates just as the BPC model does, but with knowledge of the approximate range (or rough guide, RG) containing the leading time delay T_1 . This knowledge was instilled computationally by zeroing-out the values of the bandpass cross-correlation functions outside of the range $T_1 \pm 300 \mu\text{sec}$. Maxima and lateral position were thereafter estimated as in the BPC model.

T_1 . This is an artificial model for a complete precedence effect, which generates a prediction equal to T_1 for each stimulus. For the present stimulus set, this model can be considered a version of the BPC+TW model, with $T_{ON} = 300 \mu\text{sec}$ and $k = 0$. With these parameters, no part of the lagging component of any of the stimuli used here would survive the temporal windowing, and the windowed stimulus would be a transient with a simple delay of T_1 .

Figures 6–8 present the measured pointer adjustments plotted against the predictions from the four models for all 14 test conditions. In these figures, the model is given

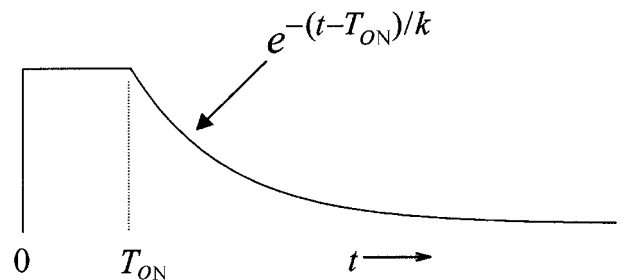


Figure 5. Diagram of the form of the temporal weighting function used in the BC+TW model.

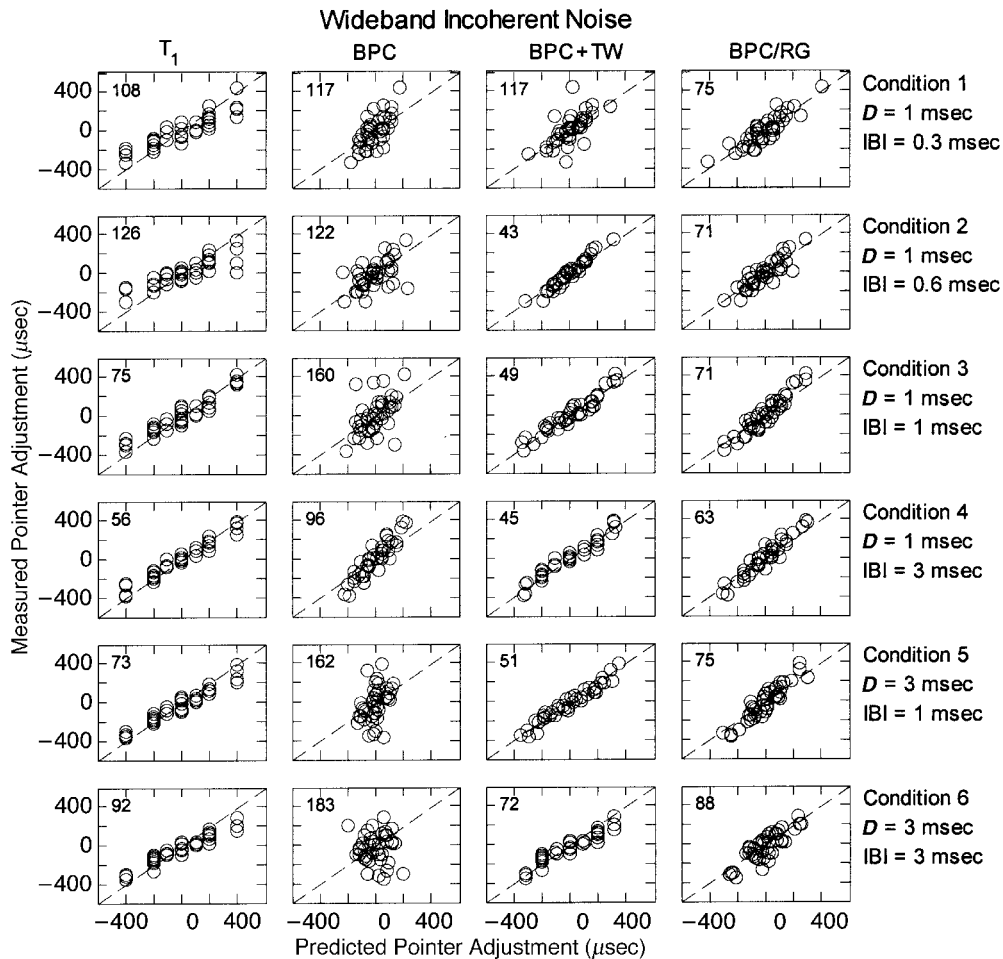


Figure 6. Scatterplots of mean pointer adjustment (vertical axes) plotted against predicted adjustments for four different models, indicated at the top of each column. The six stimulus conditions shown here all used wideband incoherent noise. The line is the major diagonal. The number in the upper left corner of each plot is the root-mean square error in predicting the data.

at the top of each column of panels, whereas different test conditions form the rows. The 40 data points in each panel represent the 40 stimuli in a condition (four sets of 10 tokens), where the y-coordinate of each point is the average pointer adjustment over subjects and replications and the x-coordinate is generated by passing the stimulus through the model. The number in the upper left corner of each panel is the root-mean square (RMS) error (in μsec) in predicting the measured adjustments.

The dominance of the initial onset, which was seen in Figure 3 in terms of stimulus-averaged values of c greater than .5 for each condition, is seen here in the generally good predictability by the T_1 model of the lateral position data for each stimulus (leftmost columns in Figures 6–8). Over all fourteen conditions, the RMS error for the T_1 model was 98.0 μsec . There was a trend in Conditions 1–4 showing best agreement with IBI = 1 or 3 msec and clearly poorer agreement with IBI = 0.3 or 0.6 msec.

This result may reflect summing localization at short IBIs, which would degrade the prediction of the data from T_1 alone. Such comparisons across conditions must be made cautiously, however, because different noise tokens were used in each condition and differences in power levels between S_1 and S_2 tokens can play a role.

Conditions 9 and 10 (Figure 7), which used noise high-pass filtered at 2.5 kHz, present an interesting result. Although the data plotted against T_1 clearly defined linear relationships, because nearly all of the responses were contained within the range from -200 to $200 \mu\text{sec}$, the slopes of those lines were less than unity. This result is likely related to the finding that the extent of lateralization is much smaller for stimuli restricted to the high-frequency region (above about 2 kHz) than for stimuli containing low-frequency energy (Trahiotis & Bernstein, 1986). Since the pointer that the subjects adjusted was always wideband regardless of the spectral content

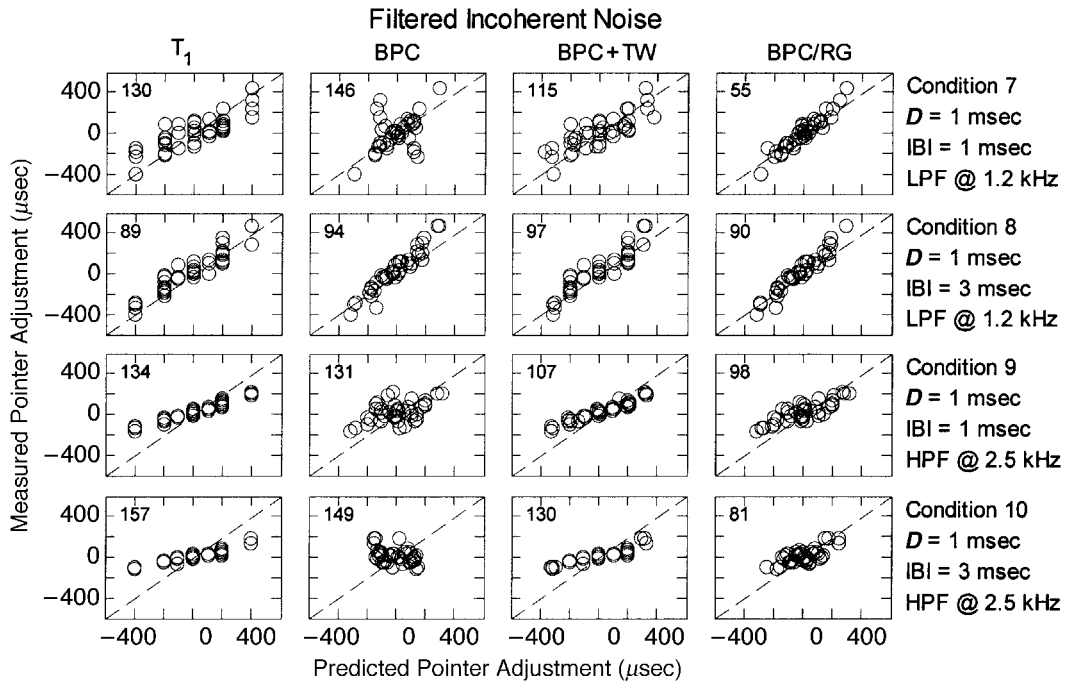


Figure 7. Scatterplots of mean pointer adjustment (vertical axes) plotted against predicted adjustments for four different models, as indicated at the top of each column. The four stimulus conditions shown here used low-pass- or high-pass-filtered incoherent noise. The line is the major diagonal. The number in the upper left corner of each plot is the root-mean square error in predicting the data.

of the test stimulus, the conditions existed for this difference in laterality to be observed. It is expected then, that the values of c for Conditions 9 and 10—.70 and .63, respectively—are reduced from their true values by this measurement artifact. This problem affects the predictions of the other models too.

Predictions by the T_1 model for the four conditions that used coherent stimuli, shown in Figure 8, were very good to excellent.

The clear failure of the BPC model to predict the measured data was expected. By averaging over the entire stimulus, there was no emphasis of initial cues, as would be required to generate predictions reflecting the dominance of initial interaural cues. Visual inspection reveals that, except in two cases, the measured data were roughly independent of the BPC predictions. The average RMS error over all 14 conditions was 148 μ sec.

In Condition 4 (Figure 6) and Condition 8 (Figure 7) the BPC model predicted the data with a fair degree of accuracy. Since there was no mechanism within the BPC model to produce this effect in the long run, it must be expected that the chance selection of noise tokens in those cases resulted in sets containing strong stimulus cues favoring the initial delay. Analysis of the leading and lagging tokens in these two conditions showed that, for three of the four sets of tokens in each condition, the level of the leading burst [spectrally weighted by the importance function $q(f)$, mentioned above] was at least 5.5 dB larger than that of the lagging burst.

The BPC+TW model was the only one given free parameters in this study. To find best-fitting values of the T_{ON} and k parameters, an error measure was defined to be the overall RMS deviation of the predicted from the mean pointer adjustments for all 14 test conditions. A search grid was defined by values of T_{ON} from 0 to 2,200 μ sec in 200- μ sec steps and of k from 0.125 to 8 msec in steps of a factor of two. The resulting error surface (Figure 9) shows a minimum RMS error of 81.7 μ sec at $T_{ON} = 600$ μ sec and $k = 1$ msec. The error contours in Figure 9 show the sensitivity of the error to variations in the two parameters. They show a local error “valley” in the region $600 < T_{ON} < 1,200$ μ sec and $k < 2$ msec outlining where other near-optimal parameter choices might lie.

The predictions of the BPC+TW model, generated using the minimum-error parameter values, are displayed in the third columns of Figures 6–8. In almost all the individual conditions, the RMS error for this model is comparable to or smaller than the smaller of the RMS errors for the T_1 and BPC models. In a few conditions (e.g., Conditions 2–5 in Figure 6), the predictive improvement provided by the combined BPC+TW model is marked.

The final set of predictions comes from the BPC/RG model, which was given a 600- μ sec-wide restriction (the RG) for bandpass correlation analysis centered on T_1 . This model also made very good predictions. In fact, the overall RMS error from all 14 conditions of 79.3 μ sec was slightly less than the 81.7- μ sec error of the minimum-error BPC+TW model. Although in most conditions

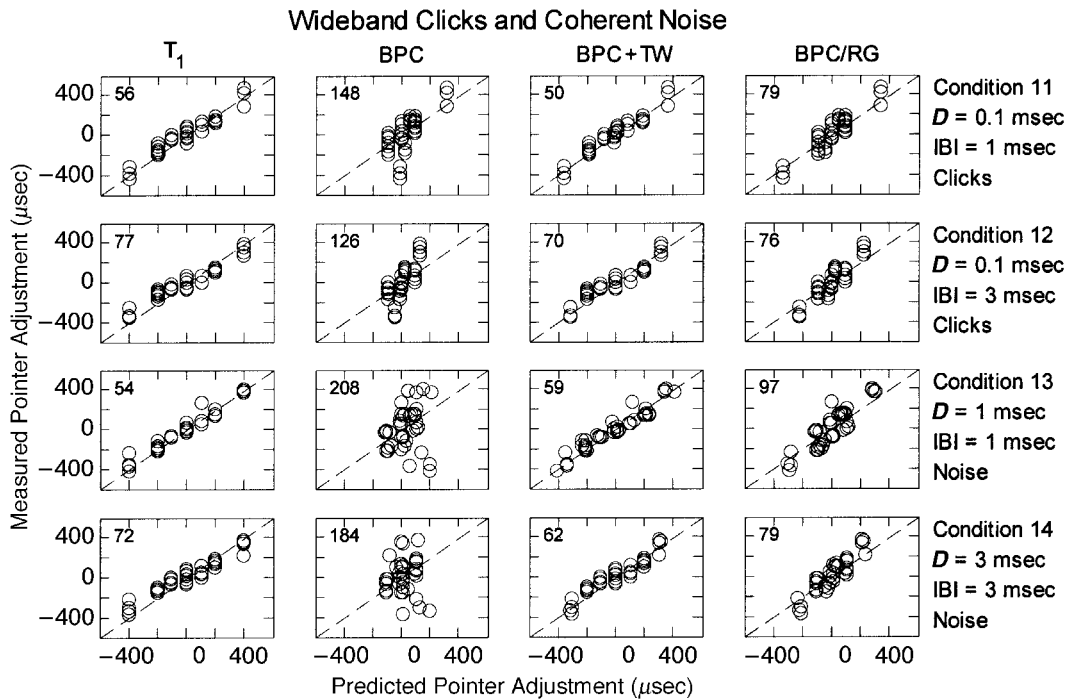


Figure 8. Scatterplots of mean pointer adjustment (vertical axes) plotted against predicted adjustments for four different models, as indicated at the top of each column. The four stimulus conditions shown here used wideband coherent noise or clicks. The line is the major diagonal. The number in the upper left corner of each plot is the root-mean square error in predicting the data.

shown in Figures 6–8, the RMS error of the BPC/RG model was larger than that of the BPC+TW model, the BPC/RG model corrected the few large errors made by the BPC+TW model, such as those in Conditions 1, 7, and 10.

DISCUSSION

The present data and theoretical analysis indicate the benefit of incorporating some form of temporal weight-

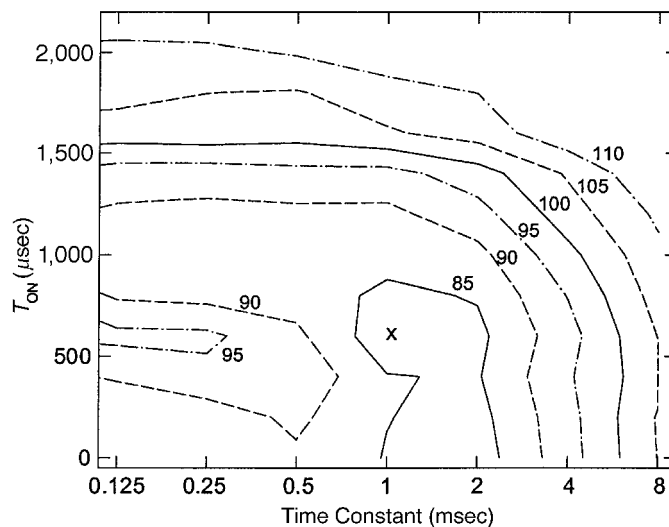


Figure 9. Contours of constant overall root-mean square (RMS) error in the error surface resulting from the two-parameter search in the BPC+TW model. The minimum point, at which the RMS error was $81.7 \mu\text{sec}$, is marked by an x. The RMS error level of each contour is indicated.

ing into bandpass correlation models of binaural hearing. This need is seen most easily with transient stimuli, such as those used here. However, onset dominance is not limited to short-duration sounds. The precedence effect was originally discovered with continuous speech and music (Haas, 1951; Snow, 1954; Wallach et al., 1949) and has been studied in several laboratory investigations with long-duration noise and click trains (Freyman, Zurek, Balakrishnan, & Chiang, 1997; Hafter, Buell, & Richards, 1988; Saberi & Perrott, 1995). The modeling effort begun here should eventually be extended to accommodate long-duration signals.

The form of the temporal weighting used in the present BPC+TW model suggests a weighting that occurs prior to binaural interaction. This order would be consistent with Hartung and Trahiotis's (2001) recent work showing that a basic BPC-type model preceded by left and right peripheral auditory models can provide—through the temporal adaptation of hair cell response that enhances neural representation of onsets—accurate predictions of precedence effect results with transient stimuli. However, the basic temporal weighting as used here could also be achieved by a model in which the temporal emphasis is either embedded in the binaural processing (Lindemann, 1986a) or results from a feed-forward weighting driven by impulsive stimulation (Zurek, 1987). The present analysis says little about the site of the temporal-weighting mechanism.

The findings in this study that (1) measured values of c were less than one, (2) the best-fitting value of k was greater than the minimum value possible, and (3) the RMS error for predictions of the T_1 model was greater than zero are all different indications that the precedence effect does not completely suppress cues following onset. Cues from the lagging transient clearly make a measurable contribution to lateral position. That the precedence effect is not complete has been known for a long while, at least since Wallach et al. (1949) assessed the relative contribution of the interaural delays of leading and lagging clicks for centering the composite image. Haas (1951), using sound localization in a free field, had the subjects counter the effect of an inter-loudspeaker delay with an inter-loudspeaker level difference. In both of these studies, the lagging sound had measurable influence. Like the present study, Shinn-Cunningham et al. (1993; Shinn-Cunningham et al., 1995) measured c in a variety of stimulus conditions to be clearly less than one. Tollin and Henning's (1999) recent demonstrations showing contributions from lagging clicks are consistent with this body of results.

The BPC/RG model was implemented here to explore another way—aside from direct temporal weighting—that onset dominance could manifest itself. The idea behind this model is that the initial interaural cues at the transient onset of a signal might be crudely but quickly estimated to be contained within a relatively large region. The result of that estimation process would restrict the range of directions to be inspected more thoroughly and slowly, using bandpass correlation analysis.

The BPC/RG model was originally suggested by onset effects with long-duration stimuli. Studies by Saberi and Perrott (1995) and Freyman et al. (1997) both used long-duration trains of transients with ambiguous BPC displays, in which multiple patterns of correlation peaks could simultaneously be seen as being roughly consistent delays across frequency. The finding of those studies was that the lateral position of such stimuli is determined completely by the interaural cue carried by the initial transient. If the initial transient led to the left, the entire long-duration, ambiguous train was heard to the left; if the initial transient led to the right, the entire train was heard to the right. It is as though the initial onset establishes a perceptual set for viewing the correlation display.

Such ambiguous BPC displays also characterize two-component transients when the leading and the lagging components are coherent. Figure 10 shows an example of such a display, using Stimulus 1 of Condition 13. It is a complex pattern, unlike that produced by a single source, and one that contains no information as to the side of the leading interaural delay. The BPC/RG model was developed to reflect the notion that the initial onset might be used to form only a crude estimate of interaural delay that emphasizes the range examined subsequently by correlation analysis. Of course, the same process can be applied to any stimulus, including the incoherent noise stimuli used here, whose bandpass correlation patterns are also unnatural for a single source but are not symmetric and not necessarily ambiguous.

The BPC+TW and BPC/RG models offer contrasting views of how the binaural system comes to place such strong apparent emphasis on the interaural cues of a transient onset. In the *temporal-weighting* view, the abrupt onset of a sound triggers a momentary partial suppression of subsequent binaural information (or, equiva-

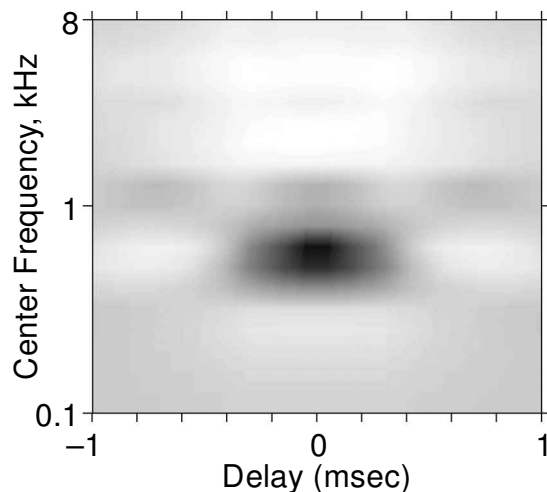


Figure 10. An example showing the BPC model's display with coherent noise and equal but opposite interaural delays (Stimulus 1 from Condition 13: $T_1 = 400 \mu\text{sec}$ and $T_2 = -400 \mu\text{sec}$). White indicates high correlation.

lently, weights the initial cues more heavily), with the combined temporally weighted information subjected to correlation analysis on a longer time scale than is the variation in temporal weighting. This line of thinking suggests a hard-wired mechanism that could plausibly be located at any of several sites in the auditory system.

In the RG view, the initial estimate from interaural onset cues is crude and is used to restrict, or bias, the range over which correlation analysis is performed on the two *unweighted* input signals. This RG would operate at the level of binaural analysis or higher. It suggests that, with a stimulus with ambiguous cues, if the bias is switched, perceived lateral position would also switch. This more central and perceptual characterization of the precedence effect, operating in conjunction with bandpass correlation analysis, might be capable of explaining not only lateralization data with transients, but also the findings of slow-acting and labile phenomena (Clifton, 1987; Freyman, Clifton, & Litovsky, 1991). It may also provide connections to related onset effects, such as the “restarting” of the binaural system with triggering events (Freyman et al., 1997; Hafter et al., 1988), the Franssen effect (Franssen, 1960; Hartmann & Rakerd, 1989), and the lateralization of sharply gated tones with ambiguous fine-structure cues (Abel & Kunov, 1983; Kunov & Abel, 1981). Although lateralization data with transients can be accommodated by either temporal weighting or the RG, the slow-acting and/or long-lasting onset effects just listed may require a perceptual biasing mechanism like that offered in the latter view.

REFERENCES

- ABEL, S. K., & KUNOV, H. (1983). Lateralization based on interaural phase differences: Effects of frequency, amplitude, duration, and shape of rise/decay. *Journal of the Acoustical Society of America*, **73**, 955-960.
- ASHMEAD, D. H., HILL, E. W., & TAYLOR, C. R. (1989). Obstacle perception by congenitally blind children. *Perception & Psychophysics*, **46**, 425-433.
- ASHMEAD, D. H., WALL, R. S., EATON, S. B., EBINGER, K. A., SNOOK-HILL, M. M., GUTH, D. A., & YANG, X. F. (1998). Echolocation reconsidered: Using spatial variations in the ambient sound field to guide locomotion. *Journal of Visual Impairment & Blindness*, **92**, 615-632.
- BLAUERT, J. (1983). *Spatial hearing*. Cambridge, MA: MIT Press.
- BLAUERT, J. (1989). Binaural technology: Fundamentals and applications [Abstract]. *Journal of the Acoustical Society of America*, **86**, S66.
- CLIFTON, R. K. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America*, **82**, 1834-1835.
- CLIFTON, R. K., MORRONGIELLO, B. A., & DOWD, J. M. (1984). A developmental look at an auditory illusion: The precedence effect. *Developmental Psychobiology*, **17**, 519-536.
- COLBURN, H. S. (1996). Computational models of binaural processing. In H. Hawkins, T. McMullen, A. Popper, & R. Fay (Eds.), *Auditory computation*. New York: Springer-Verlag.
- COLBURN, H. S., & DURLACH, N. I. (1978). Models of binaural interaction. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. IV. Hearing* (pp. 467-518). New York: Academic Press.
- CRANFORD, J., & OBERHOLTZER, M. (1976). Role of the neocortex in binaural hearing in the cat: II. The “precedence effect” in sound localization. *Brain Research*, **111**, 225-239.
- FRANSSSEN, N. V. (1960). *Some considerations on the mechanism of directional hearing*. Unpublished doctoral thesis, Delft, The Netherlands: Technische Hogeschool.
- FREYMAN, R. L., CLIFTON, R. K., & LITOVSKY, R. Y. (1991). Dynamic processes in the precedence effect. *Journal of the Acoustical Society of America*, **90**, 874-884.
- FREYMAN, R. L., ZUREK, P. M., BALAKRISHNAN, U., & CHIANG, Y.-C. (1997). Onset dominance in lateralization. *Journal of the Acoustical Society of America*, **101**, 1649-1659.
- HAAS, H. (1951). Über den Einfluss eines Einfachechos auf die Hörbarkeit von Sprache. *Acustica*, **1**, 49-58. [English translation in H. Haas (1972). The influence of a single echo on the audibility of speech. *Journal of the Audio Engineering Society*, **20**, 146-159.]
- HAFTER, E. R., BUELL, T. N., & RICHARDS, V. M. (1988). Onset-coding in lateralization: Its form, site, and function. In G. M. Edelman, W. E. Gail, & W. M. Cowan (Eds.), *Auditory function* (pp. 647-676). New York: Wiley.
- HARRIS, G. G., FLANAGAN, J. L., & WATSON, B. J. (1963). Binaural interaction of a click with a click pair. *Journal of the Acoustical Society of America*, **35**, 672-678.
- HARTMANN, W. M., & RAKERD, B. (1989). Localization of sound in rooms: IV. The Franssen effect. *Journal of the Acoustical Society of America*, **86**, 1366-1373.
- HARTUNG, K., & TRAHOTIS, C. (2001). Peripheral auditory processing and investigations of the “precedence effect” which utilize successive transient stimuli. *Journal of the Acoustical Society of America*, **110**, 1505-1513.
- HOCHSTER, M. E., & KELLY, J. B. (1981). The precedence effect and sound localization by children with temporal lobe epilepsy. *Neuropsychologia*, **19**, 49-55.
- KUNOV, H., & ABEL, S. M. (1981). Effects of rise-decay time on the lateralization of interaurally delayed 1-kHz tones. *Journal of the Acoustical Society of America*, **69**, 769-773.
- LICKLIDER, J. C. R. (1959). Three auditory theories. In E. S. Koch (Ed.), *Psychology: A study of a science* (pp. 41-144). New York: McGraw-Hill.
- LINDEMANN, W. (1986a). Extension of a binaural cross-correlation model by contralateral inhibition: I. Simulation of lateralization for stationary signals. *Journal of the Acoustical Society of America*, **80**, 1608-1622.
- LINDEMANN, W. (1986b). Extension of a binaural cross-correlation model by contralateral inhibition: II. The law of the first wavefront. *Journal of the Acoustical Society of America*, **80**, 1623-1630.
- MORRONGIELLO, B. A., CLIFTON, R. K., & KULIG, J. W. (1982). Newborn cardiac and behavioral orienting responses to sound under varying precedence-effect conditions. *Infant Behavior & Development*, **5**, 249-259.
- MUNCEY, R. W., NICKSON, A. F. B., & DUBOUT, P. (1953). The acceptability of speech and music with a single artificial echo. *Acustica*, **3**, 168-173.
- SABERLI, K. (1996). Observer weighting of interaural delays in filtered impulses. *Perception & Psychophysics*, **58**, 1037-1046.
- SABERLI, K., & PERROTT, D. R. (1995). Lateralization of click-trains with opposing onset and ongoing interaural delays. *Acustica*, **81**, 272-275.
- SAYERS, B. M., & CHERRY, E. C. (1957). Mechanism of binaural fusion in the hearing of speech. *Journal of the Acoustical Society of America*, **29**, 973-987.
- SHINN-CUNNINGHAM, B. G., ZUREK, P. M., & DURLACH, N. I. (1993). Adjustment and discrimination measurements of the precedence effect. *Journal of the Acoustical Society of America*, **93**, 2923-2932.
- SHINN-CUNNINGHAM, B. G., ZUREK, P. M., DURLACH, N. I., & CLIFTON, R. K. (1995). Cross-frequency interactions in the precedence effect. *Journal of the Acoustical Society of America*, **98**, 164-171.
- SNOW, W. B. (1954). Basic principles of stereophonic sound. *Journal of the Society of Motion Picture & Television Engineers*, **61**, 567-589.
- STERN, R. M., & COLBURN, H. S. (1978). Theory of binaural interaction based on auditory-nerve data: IV. A model for subjective lateral position. *Journal of the Acoustical Society of America*, **64**, 127-140.
- STERN, R. M., ZEIBERG, A. S., & TRAHOTIS, C. (1988). Lateralization of complex binaural stimuli: A weighted image model. *Journal of the Acoustical Society of America*, **84**, 156-165.
- TOLLIN, D. J., & HENNING, G. B. (1999). Some aspects of the lateral-

- ization of echoed sound in man: II. The role of stimulus spectrum. *Journal of the Acoustical Society of America*, **105**, 838-849.
- TRAHIOTIS, C., & BERNSTEIN, L. R. (1986). Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth. *Journal of the Acoustical Society of America*, **79**, 1950-1957.
- WALLACH, H., NEWMAN, E. B., & ROSENZWEIG, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology*, **52**, 315-336.
- WIGHTMAN, F. L., & KISTLER, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America*, **91**, 1648-1661.
- YIN, T. C. T., & LIPOVSKY, R. Y. (1995). Physiological studies of the precedence effect in the inferior colliculus of the cat. In G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, & H. Oeckinghaus (Eds.), *Advances in hearing research* (pp. 314-323). Singapore: World Scientific.
- ZUREK, P. M. (1987). The precedence effect. In W. A. Yost & G. Gourevitch (Eds.), *Directional hearing* (pp.85-105). New York: Springer-Verlag.

(Manuscript received December 15, 2000;
revision accepted for publication June 6, 2002.)