Psychological Review
2004, Vol. 111, No. 1, 52–66

0033-295X/04/$12.00 DOI: 10.1037/0033-295X.111.1.52

A Detection-Theoretic Model of Echo Inhibition

Kourosh Saberi and Agavni Petrosyan
University of California, Irvine

A detection-theoretic analysis of the auditory localization of dual-impulse stimuli is described, and a model for the processing of spatial cues in the echo pulse is developed. Although for over 50 years "echo suppression" has been the topic of intense theoretical and empirical study within the hearing sciences, only a rudimentary understanding of its mechanisms has emerged. In this article, psychometric functions and results from matching studies are used in developing a model that specifies the perceived position of the echo pulse as a normal deviate, with an expectation that is a logistic function of the echo delay and a variance that depends on interaural time difference. Loss of information in the echo event is quantified as a decline in the efficiency with which the binaural system receives information from the lag impulse.

Spatiotemporal analysis of multiple events is fundamental to all sensory systems. In visual metacontrast, for example, the successive rapid occurrence of two spatially segregated events triggers a suppression of the information in the earlier event, a phenomenon speculated to play a role in visual motion (Alpern, 1953; Bachmann, 1994; Francis, 2000). Motor and tactile systems, as well, exhibit similar spatiotemporal interactions as reported in studies of vibrotactile temporal-order discrimination or two-point sequential tactile localization where the perceived position of one tactile event is affected by a second temporally and spatially separated stimulus (Geldard & Sherrick, 1986). Cognitive systems likewise require selective processing of sequential cues from different spatial origins through temporal filters or quantal shifts in spatial attention (Shiffrin, 1988; Sperling, Reeves, Blaser, Lu, & Weichselgartner, 2001). Parallel to these, the auditory system analyzes space and time cues in domains as diverse as speech reception in a multisource environment (i.e., the well-known cocktail-party effect; Yost, 1997) to localization of complex sounds in reverberant surroundings where dozens of echoes reach the listener within a few milliseconds.

This latter ability of localizing an acoustic target in the presence of interfering echoes, some of which are more intense than the original sound (because of vectorial summation), has been the basis of a central puzzle in the hearing sciences for over 50 years since its initial formulation in two independently published studies in the United States and Germany (Haas, 1949/1972; Wallach, Newman, & Rosenzweig, 1949). When two auditory events, a sound and its echo, occur in rapid succession, the perceived spatial position of the second event is dominated by the location of the first (i.e., contrary to visual metacontrast). This dominance of the "echo" information by the first-arriving acoustic wavefront is known as the precedence or Haas effect in sound localization (Haas, 1949/1972; Wallach et al., 1949) and is speculated to play a critical role in the avoidance of spatial ambiguities in reverberant environments (Zurek, 1980). Aside from its obvious theoretical significance for what it can reveal about the mechanisms of auditory spatiotemporal processing, the study of echo suppression has had substantial impact on applied fields from architectural acoustics and concert-hall design to virtual reality and sound-reinforcement systems (Gardner, 1968; Zurek, 1987).

In spite of decades of theoretical (Hafter, Buell, & Richards, 1988; Lindemann, 1986a, 1986b; Tollin & Henning, 1998a, 1998b; Zurek, 1987), experimental (Freyman, Zurek, Balakrishnan, & Chiang, 1997; Perrott, Marlborough, Merrill, & Strybel, 1989; Saberi & Perrott, 1995; Yost & Soderquist, 1984; Zurek, 1980), physiological (Cranford & Oberholtzer, 1976; McFadden, 1973; Yin & Litovsky, 1995), applied (Blauert, 1989; Muncey, Nickson, & Dubout, 1953), and even clinical (Hochster & Kelly, 1981) research on this topic, the mechanisms that underlie echo suppression are not well understood. Current perspectives consider the precedence effect to incorporate a number of onset-dominance phenomena, some of which may involve peripheral inhibitory mechanisms (Hafter & Buell, 1990; Hafter & Dye, 1983), whereas others seem to implicate a high-order, cortical, and cognitively mediated process (Clifton, 1987; Freyman, Clifton, & Litovsky, 1991). See Blauert and Col (1991) for a discussion of irregularities in defining the precedence effect. For extended reviews see Gardner (1968), Zurek (1987), and Litovsky, Colburn, Yost, and Guzman (1999).

The notion that several phenomena are involved in onset dominance has necessitated multiple approaches to its investigation. Here we describe a psychometric model of signal detection to segregate the limitations in detection of an interaural time cue imposed by linear effects, such as additive internal noise, from the effects of nonlinear internal transforms of the psychophysical scale (Egan, 1965; Laming, 1986). Detection-theory analysis of sensory events is ubiquitous in perception research from visual localization (Allan, 1968; De Valois & De Valois, 1988) to tactile perception (Boyer, Cross, Guyot, & Washington, 1970; Cross, Boyer, & Guyot, 1970), olfactory signal detection (Cain, 1977), attention (Swets, 1984), cross-modal processing (Gescheider, Sager, & Ruffolo, 1975; Haessly, Sirosh, & Miikkulainen, 1995; Mulligan & Shaw, 1980), and kinesthetic discrimination (Cox & Hawkins, 1976). A noteworthy feature of the detection-theory model de-

Kourosh Saberi and Agavni Petrosyan, Department of Cognitive Sciences, University of California, Irvine.

This work was supported by National Institutes of Health Grant DC-03648. We thank Patrick Zurek and Ted Wright for helpful discussions.

Correspondence concerning this article should be addressed to Kourosh Saberi, Department of Cognitive Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100. E-mail: kourosh@uci.edu

scribed here is that it is not domain specific and may thus be adapted to phenomena related to other sensory systems that, in principle, display similar types of spatiotemporal interactions. In addition, we use findings from the current work to develop a measure of onset dominance based on the efficiency (Tanner & Birdsall, 1958) with which the binaural system receives information from the echo pulse of a two-impulse waveform.

A Psychometric Interpretation of an Echo Threshold Shift

For a large class of auditory stimulus dimensions (Dai, 1994; Egan, 1965; Green & Luce, 1975; Green, McKey, & Licklider, 1959; Saberi & Green, 1996, 1997), psychometric functions may be described by

$$d' = \beta \Delta\tau^m, \tag{1}$$

where d' is the index of detectability (Green & Swets, 1966/1988), β is a scalar usually related to the inverse of the internal noise variance, and $\Delta\tau$ is the stimulus scale. In the discrimination experiments described here, $\Delta\tau$ represents a change in an interaural time difference in a two-interval forced-choice (2IFC) task. Each interval contains a direct sound presented without an interaural time difference (i.e., simulating a sound source directly in front of an observer) and an echo. The echo leads to one ear in one interval and to the other ear in the second interval. The exponent m represents a nonlinear internal transform of the measurement scale, $\Delta\tau$. Similar nonlinear transforms have been described by Laming (1986), who has shown that psychometric functions in forced-choice tasks may be fitted with the upper half of the normal integral, with exponents that range from less than unity to as high as 8 (e.g., for modulation detection in vision).

On a logarithmic scale Equation 1 is a linear function:

$$\log d' = \log \beta + m \log \Delta\tau, \tag{2}$$

where $\log \beta$ and m are the intercept and slope, respectively.

Studies of the precedence effect have shown that the spatial information (e.g., τ) in a single pulse is more detectable than that in an echo pulse for short interpulse intervals (IPIs; Gaskell, 1983; Yost & Soderquist, 1984; Zurek, 1980).[1] If psychometric functions that relate discrimination performance to $\Delta\tau$ are linear on double-log coordinates, then an increase in threshold for the echo pulse may be interpreted in one of two ways, either resulting from a parallel shift of the psychometric function caused by a change in $\log \beta$, or a change in the slope of the psychometric function resulting from a nonlinear internal transform, that is, a change in m (Green, 1990, 1995; Jesteadt, Reimer, Schairer, & Nizami, 2001; Saberi, 1995; Saberi & Green, 1997). The first case usually assumes a change in internal noise variance, as expected from a loss-of-neural-sample process (Lindemann, 1986b), whereas the second assumes that the psychophysical discriminator observes a quantity that is related to the stimulus scale by a power transformation (Egan, 1965; Laming, 1986). Both interpretations may result in identical increases in threshold. Clearly, one may also observe a joint effect of $\log \beta$ and m on the psychometric function.

In the current study, we estimate psychometric functions for single pulses, as well as for the echo pulse in a two-pulse stimulus as a function of IPI. Previous studies of the precedence effect have shown that IPI is inversely related to $\Delta\tau$ threshold for IPIs in the range of 1 ms to 12 ms (Hafter et al., 1988; Hafter & Dye, 1983; Saberi, 1996; Zurek, 1980). Here, we measure psychometric functions at IPIs of 3 ms, 6 ms, and 12 ms, which are known to produce highly different thresholds. We also measure, using an acoustic pointing paradigm, effects of IPI on perceived lateral position of the echo pulse, from which the expected value and variance of the distribution of perceived echo positions for a given condition may be estimated. These parameter estimates are used to explore the nature of changes in the scalar β as the primary determinant of threshold elevation, and as a basis for development of a model of onset dominance in lateralization.

Psychometric Functions for Solitary and Echo Pulses

The stimuli used in this study were solitary and dual Gaussian pulses. Such pulses have been extensively used in studies of onset dominance in spatial hearing (Hafter & Buell, 1990; Hafter et al., 1988; Saberi, 1996). Each pulse was the impulse response of a Gaussian filter with a 1-ms effective temporal envelope (between terminal 0 V; see Figure 1); the envelope was centered on the cosine phase of a 6-kHz carrier. A single pulse can be described as

$$X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5[(t-g/2)/\sigma]^2} \cos[\omega(t - g/2)], \tag{3}$$

where $g = 1$ ms and $\sigma = 0.15$ ms. The amplitude and power spectra of each pulse are also Gaussian with widths within 1σ point of the carrier equal to 2800 Hz and 1978 Hz, respectively. Signals were generated in an IBM PC and presented through digital-to-analog converters (TDT DA2) and Sennheiser (HD 450) headphones at a sampling rate of 50 kHz. They were lowpass filtered at 20 kHz (Kemo VBF/24). The level of a continuous train (100/s) was calibrated to 60 dB SPL using a digital rms voltmeter (Fluke 8050A).

Subjects were young undergraduate college students with normal hearing within +10/−10 dB of ISO standard between 125 Hz and 8000 Hz as determined by a Békésy tracking procedure. All had previous experience in sound-localization tasks. They were trained on each task for 2 hr prior to measurement of their psychometric functions.[2]

Each run consisted of 100 trials in a constant-stimulus design.

[1] The current experiments and analysis involve IPIs $\geq$ 1 ms. At IPIs shorter than 1 ms, a categorically different phenomenon is observed, often referred to as "summing localization." In this temporal region (IPI between 0 ms and 1 ms), the interaural information from the echo has a progressively increasing influence as IPI approaches zero. Similar U-shaped inhibitory functions are reported for spatiotemporal tasks in other sensory modalities, such as for visual metacontrast.

[2] Extensive training on the precedence effect has been shown to affect lag-pulse thresholds (Saberi & Antonio, 2003; Saberi & Perrott, 1990). To check for practice effects, we selected several conditions (IPI/τ combinations) across observers and examined the difference between the first and last 250 trials, by calculating a d' for each trial group. A t test on the distribution of d' differences showed no statistically significant deviation from a zero mean. Any effects of practice during the measurement of the psychometric functions, at least for these high-frequency stimuli, the number of hours tested, and the intensities used, were negligible. This assumption of stationarity has also been made in other studies of the precedence effect (Hafter et al., 1983, 1988; Shinn-Cunningham et al., 1993).

At the beginning of every run, a sample pulse with $\tau = 0$ was presented over headphones and repeated at a rate of twice per second. Because slight asymmetric placements of the headphones over the subject's head may cause a corresponding small shift in image position, subjects were instructed to adjust the positioning of the headphones such that the percept of the sample pulse appeared centered in their head (Domnitz, 1973; Hafter, Dye, & Wenzel, 1983).[3] On each run, either the single-pulse or one of three two-pulse conditions were selected on a random basis. When the two-pulse condition was selected, the IPI was fixed within a run at 3 ms, 6 ms, or 12 ms. The $\Delta\tau$ on each trial was randomly selected from a set of predetermined values. These values were different for each subject and were selected according to each observer's performance on the training phase to cover a wide range of performance levels. The single-pulse condition was a control condition with which thresholds for echo conditions were to be compared. For the echo conditions, the stimulus consisted of two pulses, the first representing the lead event and the second representing the lag (echo) event.

On the first interval of each trial of the single-pulse condition, the pulse led to one randomly selected ear by τ, and in the second interval, it led to the other ear by the same τ. The subject's task was to determine if the order of leading τ was left ear then right ear, or vice versa ($\Delta\tau = 2\tau$). Perceptually, this is equivalent to determining if the two intracranial sound images in the two intervals of the trial were heard left then right, or right then left. The subject would then press either a left or a right key to respond (left-key response meant that they perceived the sound orders as right to left). Visual feedback was provided immediately after each trial. Data collection on each subject continued until a minimum of 500 trials were obtained for each point on the psychometric function.

Figure 2 shows psychometric functions from three observers. The parameter is experimental condition, with asterisks representing the single-pulse control condition, and the circles, plus signs, and cross symbols representing IPIs of 12 ms, 6 ms, and 3 ms, respectively. The lines are regression fits in double-log coordinates. Performance of one observer, S2, for the 3-ms IPI was quite poor and at chance at all $\Delta\tau$, and is therefore excluded. Thresholds for a single pulse at $d' = 1$ are within the range reported by Hafter and Dye (1983) for similar high-frequency Gaussian pulses (75–300 μs).

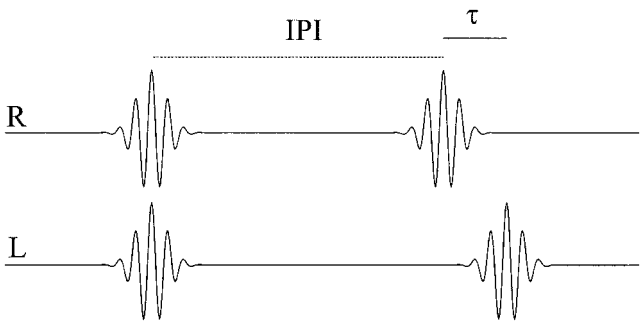

Figure 1. A schematic of the stimuli shows a lead pulse and an echo pulse. The lead pulse, which has a zero interaural time difference, is followed by a silent interpulse interval (IPI), followed by the echo pulse with interaural time difference τ, leading in this diagram to the right (R) ear. This diagram shows the waveform in one interval of the two-interval forced-choice task. The other interval contains an identical waveform, except that τ leads to the left (L) ear (i.e., $-\tau$).

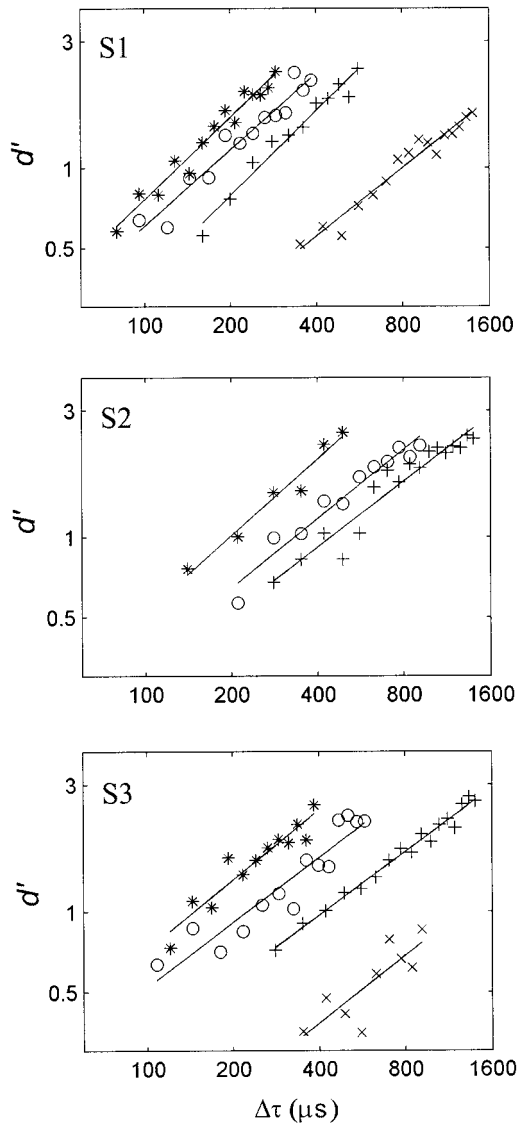

Figure 2. Psychometric functions measured for 3 observers. Asterisks represent the single-pulse condition. ○ = 12-ms interpulse interval (IPI); + = 6-ms IPI; × = 3-ms IPI; S1 = Subject 1; S2 = Subject 2; S3 = Subject 3.

Results from this experiment show that the data are well described by linear functions on double-log coordinates, and that an increase in thresholds from a decrease in IPI is a result of parallel shifts in the psychometric functions. As IPI decreases, the intercept progressively changes to lower values. The largest change in intercept is for S3 from the single-pulse

[3] The centering stimulus was the same as that used for that particular run, and was either a single pulse or a dual pulse with the appropriate IPI, except that the interaural differences were always zero.

to the 3-ms IPI case. The slope parameters for these two functions are nearly identical (0.88 and 0.84). The change in threshold at $d' = 1$, however, would be nearly a factor of 10. Other subjects show similar patterns. All slopes had values near unity across all subjects and IPI, suggesting an absence of nonlinear effects of IPI on the stimulus scale. The slopes are slightly smaller than 1 for most conditions tested, including for the single-pulse control condition, consistent with previous reports of psychometric functions in spatial-hearing tasks (Saberi, 1995; Saberi & Green, 1997).

Position Estimate as a Joint Function of τ and IPI

Results of the previous section suggest that a change in $\Delta\tau$ threshold is related to a change in the intercept of the psychometric function. Significant nonlinear transforms of the internal measurement scale were not observed, that is, slope values were near unity and independent of IPI. A change in the intercept may represent one of two factors. Equation 1 shows that given a constant $\Delta\tau$ and a constant exponent m, a change in d' would correspond to a change in the scalar β, which may be interpreted as either a change in the mean or variance of the distribution of perceived positions. It should be noted that in the current experiments, the target signal is the second pulse of the two-pulse pair and that the first pulse carries no information for resolving the detection task. Figure 3 shows that a reduction of d' to half its value (from the upper diagram to the lower two diagrams) may result either from a change from μ to $\mu/2$ or from σ to 2σ. In this section, we distinguish between these two linear effects on threshold by determining the extent to which changes in IPI affect the mean and variance of the distribution of perceived positions.

Four normal-hearing subjects participated in this experiment. Subjects ran in an Industrial Acoustics Company (Bronx, NY) sound-attenuating chamber in 1-hr sessions, two to three times per week. The test stimulus was either a single pulse or a dual pulse as shown in Figure 2. Stimuli were presented through Sony MDR-V1 headphones at a sampling rate of 40 kHz. The experiment involved a pointing task in which each observer adjusted the perceived intracranial position of an acoustic pointer to match that of a signal (target) pulse. On each trial, one of four conditions was randomly selected. These were the same as those used in the previous section, that is, single and dual pulses with IPIs of 3 ms, 6 ms, and 12 ms. The interaural time difference of the signal pulse, τ_s, was also randomly selected on each trial from a set of predetermined values that was different for each condition. To simplify later analysis and to distinguish between the signal and pointer, the subscript s has been added to τ to denote the signal interaural time difference. For the single-pulse and the 12-ms IPI conditions, τ_s was either -400, -200, -50, 50, 200, or 400 μs, whereas for the 6 ms and 3 ms IPIs, τ_s was -700, -500, -200, 200, 500, or 700

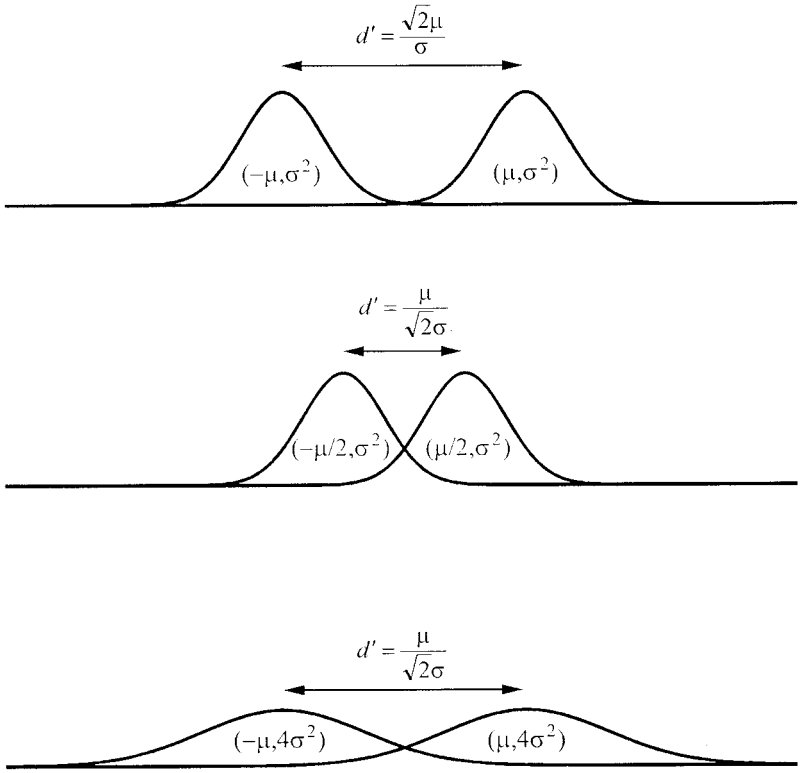

Figure 3. The top diagram is a schematic corresponding to the detection of a change in interaural time difference, from $-\mu$ to μ, of a single-pulse waveform in a two-interval forced-choice task. The lower two diagrams correspond to two explanations for a reduction in d' for the lag pulse. A smaller d' may be interpreted as either a change in the means of the distribution of perceived positions (middle diagram) or an increase in variance (bottom diagram).

μs.[4] A negative time difference denotes a left-ear leading signal and a positive time difference represents a pulse leading to the right ear. It should be noted from Figure 1 that τ_s is the interaural time difference of the single-pulse control, or the interaural time difference of the second pulse of the dual-pulse waveforms.

The pointer was a white Gaussian noiseburst, 1 ms in duration. Broadband Gaussian pulses have been used previously as a global pointer against which the perceived positions of other binaural stimuli are compared (Saberi, Sadralodabai, & Perrott, 1991; Shinn-Cunningham, Zurek, & Durlach, 1993; Zurek, 1980; Zurek & Saberi, 2003).[5] Others have used pointers that contain interaural level differences (Bernstein & Trahiotis, 1985; Stern, Zeiberg, & Trahiotis, 1988; Trahiotis & Bernstein, 1986; Trahiotis & Stern, 1989) instead of time, to require subjects to match perceived location instead of matching a time difference per se. Nonetheless, both types of pointers, based on either an interaural time or level difference, have successfully and extensively been applied to mapping perceived position, and we have here elected to use a time difference cue. On each of 40 trials within a run, a target stimulus was repeated three times at a rate of 3/s, followed by a 0.5-s silent period, followed by the pointer pulse, which was also repeated three times at 3/s. The subject then adjusted the perceived position of the pointer by pressing either a right key or a left key to "move" the pointer image toward the target image by changing the pointer interaural time difference. When the stimulus was a two-pulse waveform and two images were perceived, the target image was defined as that corresponding to the second pulse. Because at short IPIs, observers may have difficulty determining the temporal order of the onset and echo pulses (Stellmack, Dye, & Guzman, 1999), they were instructed to adjust the pointer to the image that occurred second when this was clear and that was farthest from midline. The sequence of three-target followed by three-pointer stimuli was repeated after each pointer adjustment, and the subject was allowed to adjust the pointer interaural time difference as many times as he or she required until satisfied that the pointer and target images had identical perceived loci. Three pairs of keys were available to the subject for adjustment of the pointer interaural time difference either in large (200 μs), medium (100 μs), or small (25 μs) step sizes, toward the right or left side of the interaural axis. Large step sizes were usually used to rapidly bring the pointer image near to the target image, and medium and small step sizes were used for fine adjustments of perceived positions. The range of pointer interaural time differences available to the subject was –1,000 μs to 1,000 μs, and the initial pointer time difference was selected randomly from within this range. After the subject was satisfied that the pointer and target pulses were at the same perceived positions, he or she pressed a separate key that recorded the results of the trial and proceeded to the next trial of the run. On average, each trial was completed within approximately 30 s, and each run within 25 min.

Figure 4 shows results of this experiment for 4 subjects. Each column of panels shows data from one observer and different rows represent the four experimental conditions (single pulse, dual pulse with IPIs of 12 ms, 6 ms, and 3 ms from top to bottom). The abscissa of each panel represents signal interaural time difference τ_s, and the ordinate is the pointer interaural time difference, that is, the position estimate $\tau_{\hat{p}}$. Each circle is one pointing response. The dashed lines have a slope of unity. The solid lines are regression fits.

The slopes of the regression lines decrease with decreasing IPI. Largest slope is observed for the single-pulse control condition. Even an IPI of 12 ms is not sufficient to eliminate the effects of the lead pulse on the perceived position of the lag pulse, as the slopes of the functions in the second row are significantly more shallow than those of the single-pulse condition. At an IPI of 3 ms, the slopes average to near zero, suggesting that subjects perceive the target stimulus near the center of the interaural axis at all τ_s.

It is evident that not only does the mean of the perceived positions approach zero with decreasing IPI but that this change in mean is proportional to the signal interaural time difference, that is, larger shifts occur for larger τ_s. There does not, however, appear to be a change in the variance of position estimates as a function of IPI (see also Litovsky & Macmillan, 1994, who reported a similar finding). For all but 1 subject, this variance is nearly the same for the single-pulse condition (top row) as it is for the 3-ms IPI condition (bottom row). In fact, in two cases, the variance of position estimates is marginally smaller for the 3-ms condition compared with the single-pulse condition (Subjects S4 and S6). *F* tests of the hypothesis that the variance of pointing estimates corresponding to the two-pulse condition is larger than that of the single-pulse conditions were nonsignificant at all IPI: 3 ms, $F(240, 186) = 0.63$, *ns*; 6 ms, $F(250, 186) = 0.68$, *ns*; and 12 ms, $F(336, 186) = 0.47$, *ns*.[6] Figure 5A shows variance measures as a function of τ_s. Thus, one may conclude that the decrease in detection of $\Delta\tau_s$ (previous experiment) is a consequence of changes in mean position judgments as described in the middle diagram of Figure 3, and not its variance (bottom diagram).

Predicted d' From Position Estimates

The preceding analysis shows that the reduction of spatial information in the echo pulse of a two-pulse waveform results from a linear scaling of the mean of the distribution of position judgments as a function of IPI. To compare the relation between position judgments from Experiment 2 and the psychometric functions obtained from Experiment 1, we derived predicted d' from pointing estimates (Green & Swets, 1966/1988; van Trees, 1968; see also Colburn, 1973; Stern & Colburn, 1978) by calculating at

[4] For two observers, pointing judgments were also collected at $\tau_s = +600/-600$ μs in the single-pulse control condition (see Figure 4). However, these were not included in the averaged scores reported because other observers did not run this τ value.

[5] Broadband noise pointers have also been used in other studies of the precedence effect (Shinn-Cunningham et al., 1993; Zurek & Saberi, 2003). We used this type of stimulus as a global referent against which the perceived position of any other stimulus may be calibrated. As we show, the position judgments for a single high-frequency pulse is matched by the broadband pointer as slightly offset from τ_s, consistent with Zurek and Saberi (2003). Nonetheless, we prefer to use this broadband pulse as a general metric against which the perceived position of other stimuli may be referenced.

[6] For estimating the variances, the signs of the pointing estimates corresponding to the negative τ_s were reversed, and these estimates were pooled with those corresponding to the positive τ_s across subjects at a given IPI.

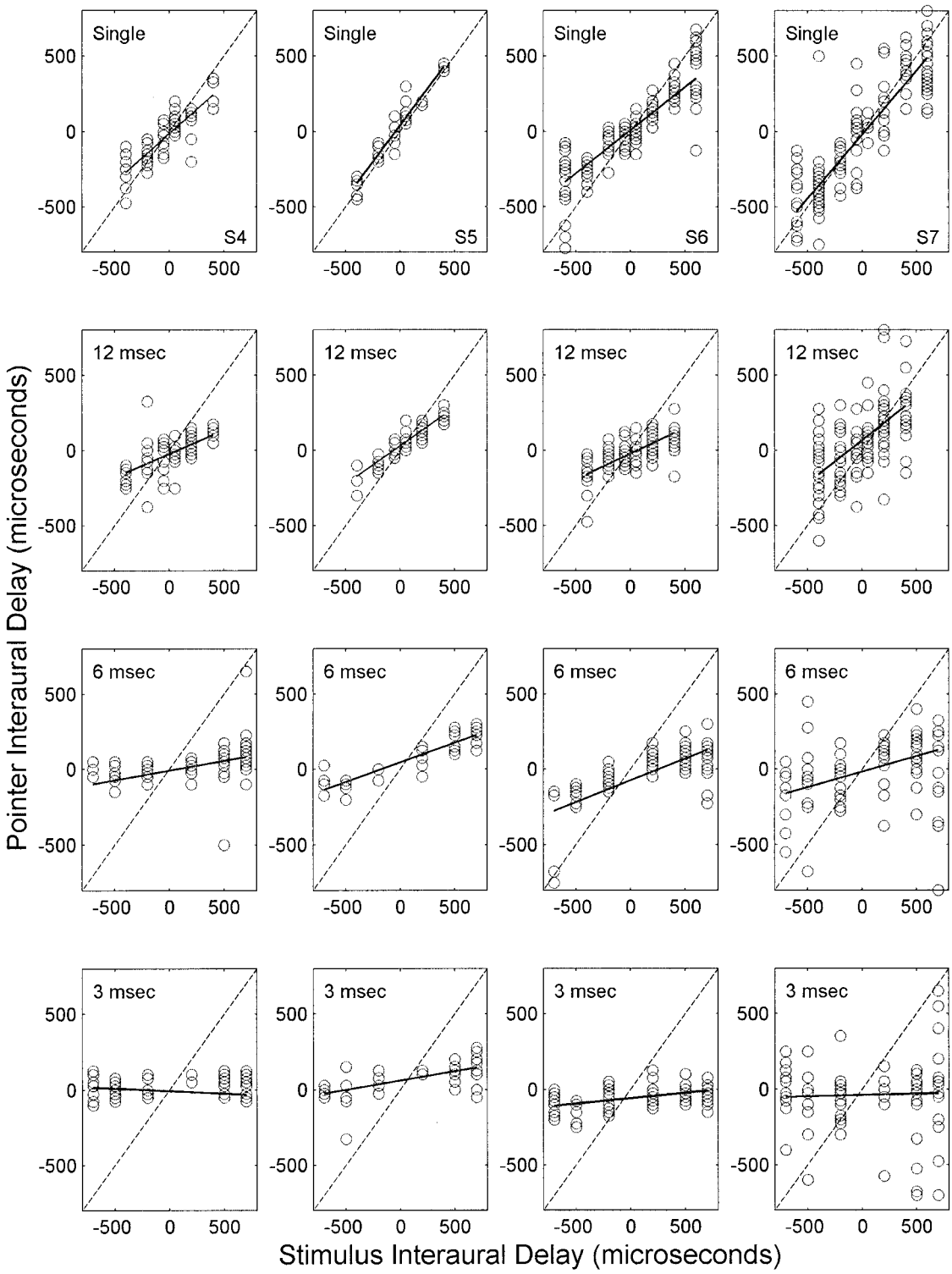

Figure 4. Position judgments from the pointing task. The abscissa is the stimulus interaural time difference τ_s, either of the single pulse or the second pulse of the two-pulse waveform. The ordinate shows the interaural time difference of the pointer, that is, the position estimate $\tau_{\hat{p}}$. Each column of panels shows data from one observer (S4 to S7 from left to right). Each row shows data for one condition: From top to bottom, these are the single-pulse control condition and the 12-ms, 6-ms, and 3-ms interpulse interval conditions. The dashed line shows unity slope, and the solid lines are regression fits.

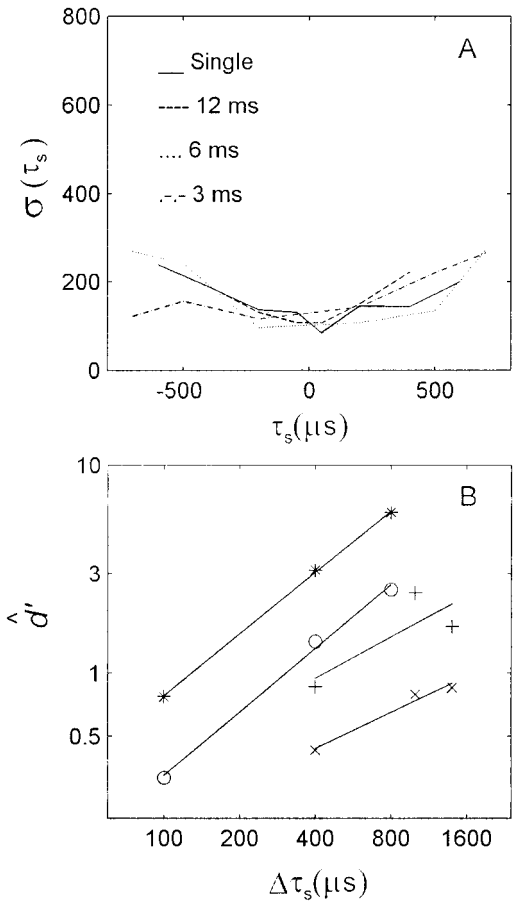

Figure 5. A: The standard deviation of pointing estimates from Figure 4 at all combinations of interpulse interval (IPI) and τ_s, averaged across observers. Solid line = single-pulse condition; dashed line = 12-ms IPI; dotted line = 6-ms IPI; dashed–dotted line = 3-ms IPI. B: $\hat{d}'$ for the single-pulse control (asterisks) and the 12-ms IPI condition, predicted from pointing judgments (see text). ○ = 12-ms IPI; + = 6-ms IPI; × = 3-ms IPI.

each τ_s in Figure 4, the difference distribution of pointing estimates, $\tau_{\hat{p}}$, for corresponding right- and left-leading τ_s:

$$\hat{d}' = \frac{(\bar{\tau}_{\hat{p}}|+\tau_s) - (\bar{\tau}_{\hat{p}}|-\tau_s)}{\sqrt{\sigma^2_{(\bar{\tau}_{\hat{p}}|+\tau_s)} + \sigma^2_{(\bar{\tau}_{\hat{p}}|-\tau_s)}}}, \tag{4}$$

where $(\bar{\tau}_{\hat{p}}|+\tau_s)$ is the mean position estimate given a positive signal interaural time difference (i.e., target pulse leading to the right ear), $(\bar{\tau}_{\hat{p}}|-\tau_s)$ is the mean estimate given a negative time difference, and $\sigma^2_{(\bar{\tau}_{\hat{p}}|+\tau_s)}$ and $\sigma^2_{(\bar{\tau}_{\hat{p}}|-\tau_s)}$ are the corresponding variances. Figure 5B shows mean $\hat{d}'$s with regression fits in double-log coordinates.[7] The abscissa is the difference in stimulus interaural time difference between left- and right-ear leading stimuli. The functions appear reasonably consistent with those from the discrimination experiment shown in Figure 2.

Position-Estimate Model

The position estimate in this task may thus be modeled as a stationary normal deviate with expectation,

$$E(\hat{P}) = \tau_s\gamma(T), \tag{5}$$

where T in milliseconds is the IPI bound by 1 and infinity, and γ is the logistic function

$$\gamma(T) = \frac{2\alpha}{1 + e^{-\lambda(T-0.9)}} - \alpha, \tag{6}$$

which relates the scalar γ to the IPI and has a value near zero for $T = 1$ ms and near unity at $T = \infty$ (representing the single-pulse condition).[8] The constants λ and α are the slope and upper asymptote of the logistic function, respectively, and are described below. We selected the logistic model for this weighting parameter because it conforms to the data that relate performance to IPI as shown in Figure 6. This figure shows that the empirically measured detection performance at a fixed $\Delta\tau$ (400 μs) for three observers increases precipitously for IPIs from 2 ms to 10 ms, and gradually above this limit with an upper asymptote at a value of d' that depends on the interaural time difference. The dashed line is mean d', and the solid line is a logistic function fitted to the mean data using a multivariate Nelder–Mead Simplex algorithm. This function was modified to produce weights from near zero to near unity as shown in Equation 6. The estimation of $\gamma(T)$ from d' measures as a scalar for determining $E(\hat{P})$ is justified because a change in d' may be interpreted as a proportional change in the mean of the distribution that underlies its estimation (Figure 3). The constants λ and α, as noted above, are the slope and upper asymptote, respectively, of the logistic function. λ was set to 0.17 and α to 0.9 for all model predictions in this article. These values were selected to visually produce reasonable functions shown in Figure 7 (as discussed below). The asymptote α represents a small proportional offset in position judgments for single-pulse waveforms (i.e., slopes of slightly less than 1 in the top row of Figure 4) and may be ignored or set to 1 in studies where $\tau_{\hat{p}}$ matches τ_s for this control condition (see Footnote 5).

The function in Equation 6 has a mean of 0.9 where it crosses the T axis. Because the lower limit on T is 1, the lower bound on γ will be near but not equal to zero. This value for the mean was selected in accord with findings from previous reports that the precedence effect, even at maximum strength at $T = 1$, will not result in complete loss of spatial information (Zurek, 1987).

The model assumes that on a given trial, $\hat{P}$ is sampled from

$$\hat{P} \sim \text{normal}(\tau_s\gamma(T),\ \sigma^2_{\hat{P}}(\tau_s)), \tag{7}$$

[7] $\hat{d}'$ was obtained for single-pulse and 12-ms IPI conditions for all subjects. However, we could not derive reliable estimates of $\hat{d}'$ for IPIs of 6 ms and 3 ms because the data of one observer (S4) either included outliers that inflated the variance estimates or produced unreliable functions. For example, in the lower left panel of Figure 4, the solid line has a negative slope. We speculate that a larger number of trials for this observer would eliminate these anomalies and allow a better estimation of the differences in the mean of position estimates and thus a better estimate of $\hat{d}'$. However, we no longer have access to this subject and have therefore excluded his data from the 3-ms and 6-ms functions.

[8] For the subject for whom we had the most number of pointing estimates (S7), we performed 13 separate Kolmogorov–Smirnov tests for normality of the distribution of pointing estimates across four conditions and all τ_s. In all cases, the hypothesis of normality was retained and we therefore consider this assumption of the model to be consistent with the data (see also Saberi, 1995). Shinn-Cunningham et al. (1993) made a similar assumption.

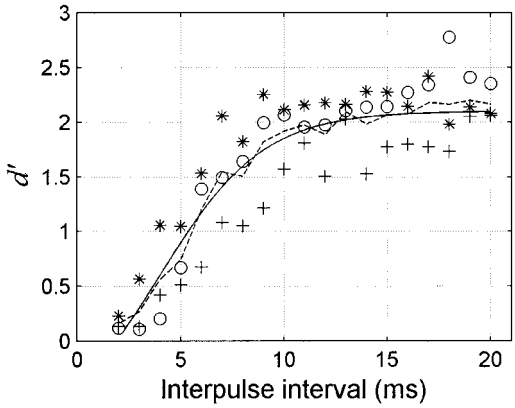

Figure 6. Detection index d' as a function of interpulse interval measured for 3 subjects (different symbols) for the second pulse of a two-pulse waveform with $\Delta\tau$ constant at 400 μs. The dashed curve shows the mean of the data and the solid curve is a logistic function fitted to this mean.

where the variance $\sigma^2_{\hat{P}}(\tau_s)$ is independent of T and is estimated from position judgments of Figure 4.[9] Figure 7 shows model predictions for the four stimulus conditions used in the experiments (single pulse, and IPIs of 3 ms, 6 ms, and 12 ms). The dashed line in each panel is the diagonal with slope of 1, the solid line is the expected value of $\hat{P}$, and the symbols are 20-sample position estimates at each of eight τ_s (Equation 7).

Several features of the model predictions are worth noting. First, the $E(\hat{P})$ values represented by the solid lines are proportional to τ_s and decline as a function of IPI. The magnitude of decline is similar to that observed empirically (Figure 4). The variance of position estimates is independent of IPI, but mildly dependent on τ_s, with larger variances at the extreme values of τ_s, consistent with data of Figures 4 and 5A. Finally, at an IPI of 1 ms (not shown), the slope of the solid line would be near but not equal to zero, as predicted from previous reports (Saberi & Perrott, 1990; Zurek, 1980). This suggests that one may make position predictions from Equation 7 for virtually any IPI and τ_s.

For the same parameter values in Equation 7, one may also make predictions of the detectability of $\Delta\tau_s$ for discriminating any pair of values of τ_s, that is, one value for each interval of the 2IFC detection task. In the current experiments, τ_s and $-\tau_s$ comprise the signal interaural time differences to be discriminated across the two intervals and thus for the equal-variance case:

$$\hat{d}' = \frac{\tau_s\gamma(T) - [-\tau_s\gamma(T)]}{\sqrt{2\sigma^2_{\hat{P}}(\tau_s)}} = \frac{\sqrt{2}\tau_s\gamma(T)}{\sigma_{\hat{P}}(\tau_s)}. \quad (8)$$

Figure 8 shows $\hat{d}'$ as a function of T and $\Delta\tau_s$ (i.e., $2|\tau_s|$) for the four conditions tested, as well as two other interesting cases. Predictions for the four conditions are shown with the solid curves and labels as T = 3 ms, 6 ms, 12 ms, and infinity, corresponding to the single-pulse condition. Predictions for two other values of T = 1 ms and 20 ms are plotted as dashed curves for comparison.

Several features of the predictions are noteworthy. First, log $\hat{d}'$ is a nearly linear function of log $\Delta\tau_s$ at all T (r^2 = .99). The slight curvature is due to the fact that $\sigma^2_{\hat{P}}(\tau_s)$ is not constant but increases with τ_s (see Footnote 9). Such an increase is small and results in only minor deviations from linearity and, given measurement noise, is not apparent in Figure 2. All functions in Figure 8 are parallel as predicted from Equation 7 (cf. Equation 2). The slopes of all functions are 0.88, very near those reported for the functions of Figure 2. It should be noted that as T increases, the intercepts increase in a compressive form toward the upper asymptote at the curve represented by infinity. This feature is also observed in Figure 2, as the parallel distance between functions representing T of 3 ms and 6 ms is greater than that corresponding to T of 12 ms and 6 ms. The vertical solid lines represent threshold at $\hat{d}'$ = 1. The precise threshold values are not critical because they depend on $\sigma^2_{\hat{P}}(0)$, which itself depends on individual subjects. However, for reference, thresholds for T of infinity, 12 ms, 6 ms, and 3 ms for predictions of Figure 8 were 148 μs, 206 μs, 387 μs, and 1,045 μs, respectively. These values conform to the two functions in Figure 5B, because they are based on the same variance measures. These predicted thresholds are also in the general range of those shown in Figure 2.

We should note here that using a Gaussian model for position estimation, which is consistent with assumptions of statistical decision theory, has parallels in a variety of research fields including spatial vision (Allan, 1968; De Valois & De Valois, 1988), polysensory signal detection and localization (Gescheider et al., 1975; Haessly et al., 1995; Mulligan & Shaw, 1980), and even applied fields such as machine vision and robotics (Stroupe, Martin, & Balch, 2001). Our discussions here, we believe, should thus be considered within the context of a more generalized systems-level description that extends beyond mechanisms specific to a single modality. That notwithstanding, it is useful at this stage to speculate on candidate neural structures that underlie a shift in the mean position of the echo image without affecting the variance of these position judgments. Comparable neurophysiological processes may be involved in localization of temporally proximate events in other sensory domains.

Explanations of echo inhibition have predominantly focused on binaural cross-correlation, which simulates coincidence detectors and axonal delay lines of the medial superior olivary (MSO) complex (Hartung & Trahiotis, 2001; Lindemann, 1986a, 1986b; Zurek & Saberi, 2003). Traditional position-estimate models have, as well, been based on the circuitry of the MSO (Colburn, 1973; Sayers & Cherry, 1957; Stern et al., 1988). Although these explanations are helpful to understanding aspects of echo suppression, empirical psychophysical evidence implicates a broader and higher level mechanism beyond a binaural process (Blauert, 1971; Litovsky, Rakerd, Yin, & Hartmann, 1997; Rakerd & Hartmann, 1997). Physiological studies (Fitzpatrick, Kuwada, Batra, & Trahiotis, 1995; Fitzpatrick, Kuwada, Kim, Parham, & Batra, 1999; Yin & Litovsky, 1995) have shown that echo suppression may be associated with a loss of neural samples in the inferior colliculus

[9] The variances of position estimates and discrimination thresholds are moderately dependent on τ_s (Hafter & De Maio, 1975). This variance may be defined in the model as $\sigma_{\hat{P}}(\tau_s) = \sigma_{\hat{P}}(0) + \omega|\tau_s|$, where ω is a constant of proportionality estimated from the data of Hafter and De Maio (1975) to have a value of 0.05. $\sigma_{\hat{P}}(0)$ is the standard deviation of position judgments and is estimated from $\sigma_{\hat{P}}(50)$ in Figure 4 averaged across 4 subjects. Having derived $\sigma_{\hat{P}}(0)$, we can determine $\sigma_{\hat{P}}(\tau_s)$ in the model for any value of τ_s.

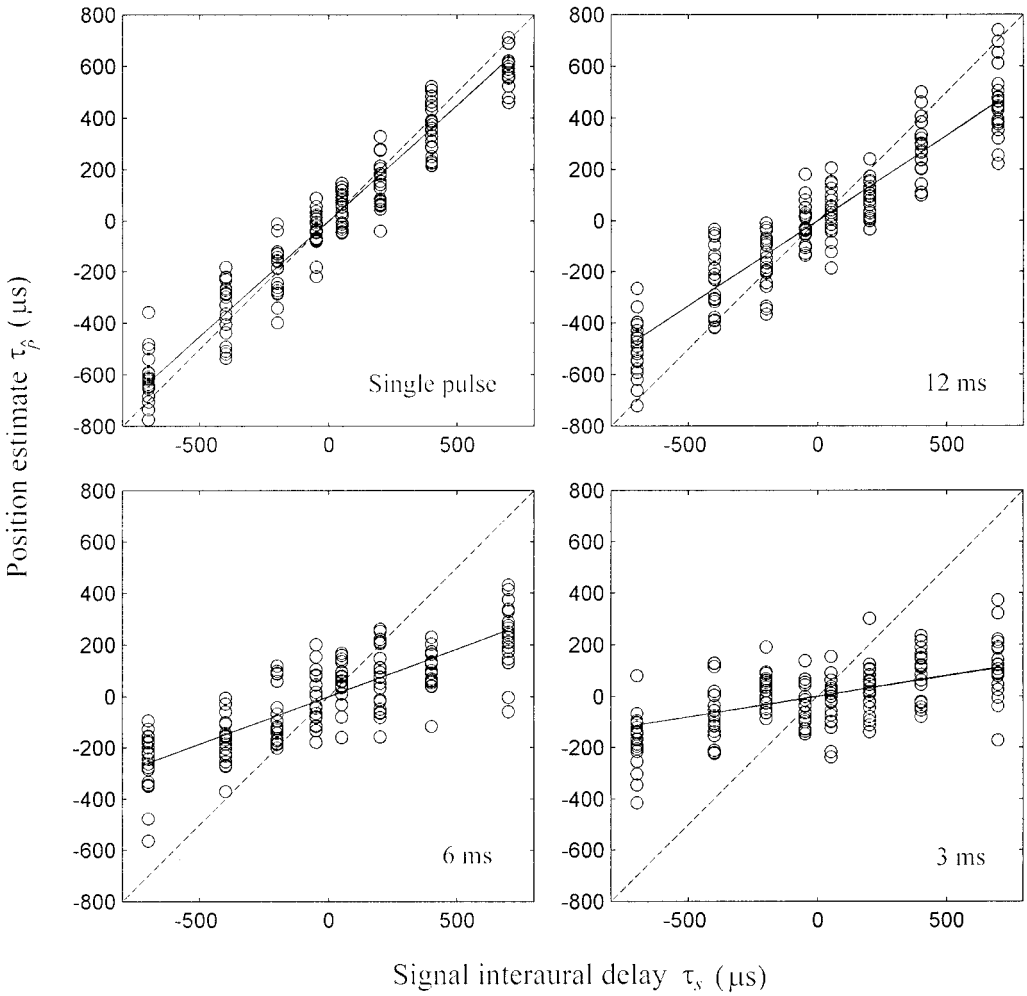

Figure 7. Model predictions of pointing estimates for the four conditions used in the experiments. At each combination of interaural time difference and interpulse interval, a 20-sample estimate was obtained, with each symbol corresponding to one estimate. The solid line shows the predictions for the expected values of pointing judgments. The dashed line is the diagonal with a slope of 1.

(IC), which receives direct as well as indirect input from the MSO through the nucleus of the lateral lemniscus (Pickles, 1988). Neurons in the IC encode for both horizontal and vertical sound-source positions and show diminished firing rates in response to echo stimuli. Reduced firing rates or loss of neural samples alone, however, imply changes in the variance of the distribution of perceived echo position, and not a change in mean position, and are therefore incomplete explanations.

A more convincing explanation would require a shift in the receptive field of neurons (e.g., preferred location) without necessarily a reduction of firing rate. This would be equivalent to a change in, for example, the best interaural time difference of a given neuron. Studies of IC neurons, to our knowledge, have not shown such a displacement of the entire ITD-tuning curve. An alternate explanation is that echo inhibition may result from a neurally based weighted lowpass filter (Koch, 1999), with greater weight given to the onset pulse. Models of this type can explain perceptual fusion of onset and echo pulses, as well as changes in the perceived position of the echo. Furthermore, if one assumes that the neural noise that limits the detection of ITD is primarily central (i.e., downstream of the lowpass filter), then given that the two pulses occur in rapid succession, they may well share a common noise whose variance would not change as a result of processing either a single pulse or two pulses. The result is that a change in perceived position of the echo is predicted without a change in the variance of the position estimates. This explanation fits much of the data, except that it cannot account for the perception of two distinct images at long IPIs (e.g., 10 ms), especially when the echo image is shifted toward the position of the onset pulse, unless one assumes that the lowpass filter acts on position information but not on other image features (e.g., extensity, perceived sharpness).

Considering that the principal physiological data on echo inhibition points to the reduction of firing rate of IC neurons as a candidate mechanism, then the function that maps this loss-of-samples to a shift in perceived location should be specified as a next step. No evidence of such a mapping function has yet been provided. Furthermore, given the complexity of effects associated

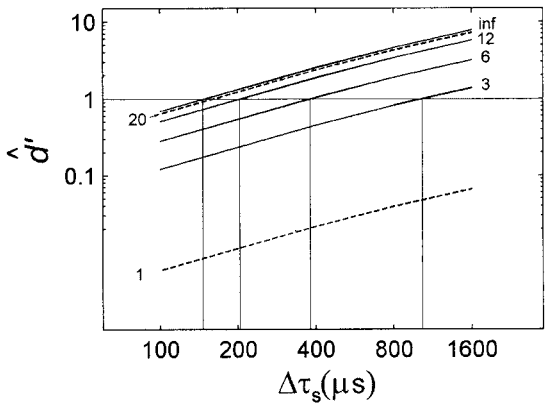

Figure 8. Model predictions for detection of a change in interaural time difference $\Delta\tau_s$ for various combinations of $\Delta\tau_s$ and interpulse interval (IPI). The solid curves correspond to the four experimental conditions, that is, single-pulse condition and IPIs of 12 ms, 6 ms, and 3 ms. The dashed curves are predictions for IPIs of 1 ms and 20 ms, plotted for comparison. Note that the parallel distance between functions is an inverse compressive function of IPI. Vertical lines correspond to the threshold at $d' = 1$. inf = infinity.

with echo suppression, different aspects of this process may be mediated by different neural structures. The origins of the reduced firing rates in the IC may, for example, be prior to the brainstem, whereas other aspects of echo inhibition, such as recovery with learning (Saberi & Antonio, 2003; Saberi & Perrott, 1990) or breakdown effects (Clifton, 1987), are likely associated with cortical structures.

Efficiency as a Measure of Onset Dominance

Several previous reports have quantified the magnitude of onset dominance in lateralization (Hafter & Dye, 1983; Hafter et al., 1983, 1988; Shinn-Cunningham et al., 1993). Although these measures have been quite useful in the analysis of onset dominance, they do not specify how dominance varies as a function of the main parameter of interest, T. In addition to describing a functional form for the relation between T and onset dominance, we prefer here to adopt a detection-theoretic measure of onset dominance, defined as the efficiency (Berg, 1990; Saberi, 1996; Tanner & Birdsall, 1958) with which the binaural system receives spatial information from the second pulse of a two-pulse waveform. The measure of efficiency η is defined as the squared ratio of d' for the second pulse at a given τ_s to ideal d' at that τ_s:

$$\eta = \frac{\hat{d}_s'^2}{d_{\text{ideal}}'^2} = \frac{(\tau_s\gamma(T) - [-\tau_s\gamma(T)])^2/2\sigma_{\hat{p}}^2(|\tau_s|)}{(\alpha\tau_s - \alpha[-\tau_s])^2/2\sigma_{\hat{p}}^2(|\tau_s|)} \qquad (9)$$

$$= \left(\frac{\gamma(T)}{\alpha}\right)^2 = \left(\frac{1 - e^{-\lambda(T-0.9)}}{1 + e^{-\lambda(T-0.9)}}\right)^2,$$

where d_{ideal}' is simply that associated with T of ∞, that is, the single-pulse control. As the temporal separation between the pulses increases, the information from the second pulse is received with greater efficiency, and as this separation approaches infinity, the transmission efficiency approaches unity, that is, an impulse in isolation. An important feature of this measure is that it is independent of the specific values of interaural time difference (which cancel in Equation 9), depending exclusively on T. This useful characteristic is consistent with the interpretation of the strength of the precedence effect as depending on the time delay between the lead and lag events, and not their degree of spatial separation. Equation 9 also shows that α cancels in estimating efficiency, allowing η to reach a maximum value of unity and, therefore, be easily extended to discrimination experiments. Furthermore, the use of this efficiency measure allows comparison of the strength of onset dominance across a variety of stimuli, experimental conditions, observers, and even to different psychoacoustic phenomena. Figure 9A shows η calculated from Equation 9 for the parameters described for Equation 7 as a function of T. The efficiency is close but not equal to zero at $T = 1$ ms, and increases monotonically with an upper asymptote of unity. At $T = 20$ ms, which is often assumed to be outside the temporal region of the precedence effect, efficiency is about 0.85, and at $T = 10$ ms, it is less than 0.50. The asterisks in Figure 9A are the efficiencies averaged across three observers from Figure 2. To obtain these values, we selected the point ($\Delta\tau$) on the single-pulse psychometric function that corresponds to $d' = 1$ for each observer and determined the d's for the other functions at that $\Delta\tau$ through interpolation. The circles are η values calculated for the data of Figure 6.[10] The form of this function is consistent with the weighting estimates reported by Stellmack et al. (1999), who used a correlation method (Lutfi, 1995; Richards & Zhu, 1994) to estimate weights given to echo and onset pulses in a variety of experimental conditions. They observed less than optimal weighting of echo pulses for long IPIs (greater than 12 ms) that are traditionally considered outside the range of echo suppression. Stellmack et al. examined IPIs as long as 256 ms, and for the condition that most resembled our design (target echo/fixed source), percent correct detection of echo τ_s interestingly did not reach that for a single pulse, even at the longest IPI (80% vs. 85% correct detection). This difference is larger than we would have expected from the efficiency function shown in Figure 9A and may be related to the trial-by-trial perturbation of the echo τ_s in that study. Figure 9B shows η calculated for several types of stimuli from previously published reports.[11] These estimates are averaged across observers within each study. Estimates of efficiency from all these reports show that even for IPIs of about 10 ms, efficiency is notably below 1 and is, at least for one study, less than 0.2. The data of Figure 9B also demonstrate a large variability in the efficiency of processing information from the echo pulse, which may be partially related to the different types of stimuli used in the various studies.

Relation Between Three Measures of Onset Dominance

We describe here two additional cases for which onset dominance in localization has been quantified. These are the models

[10] For these data we estimated d_{ideal}' from the single-pulse psychometric functions of Figure 2 at $\Delta\tau = 400$ μs because the same subjects were used to obtain the data of Figures 2 and 6.

[11] In many of the studies for which η is shown in Figure 9B, the measure of performance was $\Delta\tau$ threshold at a fixed detection index (e.g., $d' = 1$). For these conditions $\eta = (\Delta\tau_{\text{ideal}}\sigma(\tau_s))^2/(\Delta\tau_s\sigma(\tau_{\text{ideal}}))^2$, where $\sigma(\tau)$ is used to correct for the effects of increasing variances of internal noise as τ increases in value (see Footnote 9). $\Delta\tau_{\text{ideal}}$ is the interaural time difference threshold for a single pulse, and $\Delta\tau_s$ is the threshold for the second pulse.

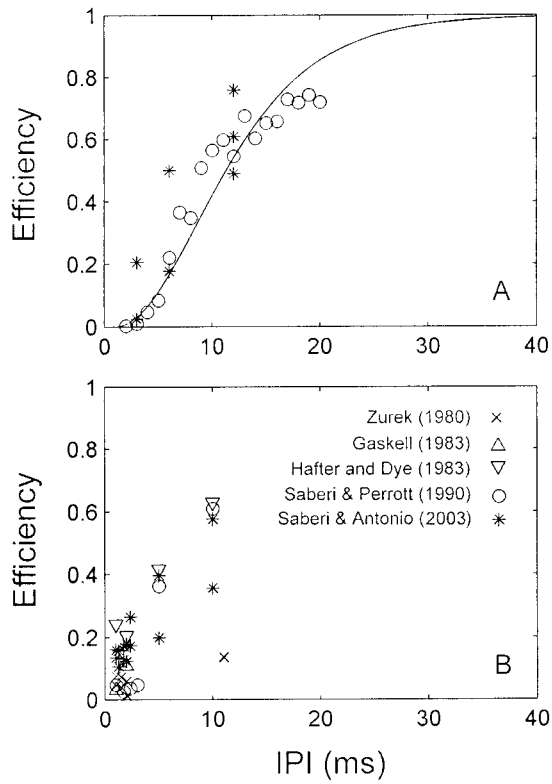

Figure 9. A: η as a function of the interpulse interval (IPI) for the data of Figure 6 (circles) and Figure 2 (asterisks). The curve is Equation 9. B: η calculated for data from several previous studies. See the text for details.

described by Hafter and his colleagues under the topic of binaural adaptation, and that described by Shinn-Cunningham et al. (1993). Hafter et al. (1983, 1988) have described a set of experiments in which they show that $\Delta\tau$ threshold for a train of n pulses having the same interaural time difference improves according to

$$\Delta\tau_n = \Delta\tau_1/n^{0.5k}, \tag{10}$$

or in its more convenient logarithmic form,

$$\log\ \Delta\tau_n = \log\ \Delta\tau_1 - 0.5k \log\ n, \tag{11}$$

where $0 \leq k \leq 1$ is a function of IPI and represents maximum loss of information when it is equal to zero and ideal detection when k is unity. On double-log coordinates, k affects the slope of Equation 11. For $k = 1$, the slope of the line relating n to threshold $\Delta\tau_n$ is 0.5, that is, $\sqrt{n}$ improvement. When k is zero, there is complete loss of postonset information, and threshold for n pulses is the same as that of the onset pulse, $\Delta\tau_1$. We address here how k relates to the measure of efficiency η. Hafter and Buell (1990) have defined the "relative effectiveness" of a single pulse as a function of its position j in the train as

$$\text{relative effectiveness} = n_j^k - n_{j-1}^k, \tag{12}$$

where n_j is the number of pulses up to position j and is assumed to be proportional to the number of neural samples that encode for the signal interaural time difference. It can be shown that this measure of relative effectiveness is, in fact, precisely the same as Tanner and Birdsall's (1958) efficiency measure η as used in our model. Because Equation 12 represents the change in the proportion of neuronal samples that code for a given interaural time difference, and because the standard deviation of this process is inversely proportional to the square-root of the sample size, the detection index for pulse j is related to that of the leading pulse by

$$d'_j = d'_1\sqrt{n_j^k - n_{j-1}^k} \tag{13}$$

or

$$n_j^k - n_{j-1}^k = \left(\frac{d'_j}{d'_1}\right)^2 = \eta, \tag{14}$$

where d'_1 is that associated with a pulse in isolation, that is, d'_{ideal}. If there are two pulses in the train ($n_j = 2$; $n_{j-1} = 1$), as is the case for the current set of experiments, then solving for k in Equation 14 yields

$$k = \log_2(1 + \eta) = \log_2\left(1 + \left(\frac{d'_s}{d'_{\text{ideal}}}\right)^2\right). \tag{15}$$

Figure 10A shows that k is a curvilinear function of η, both have

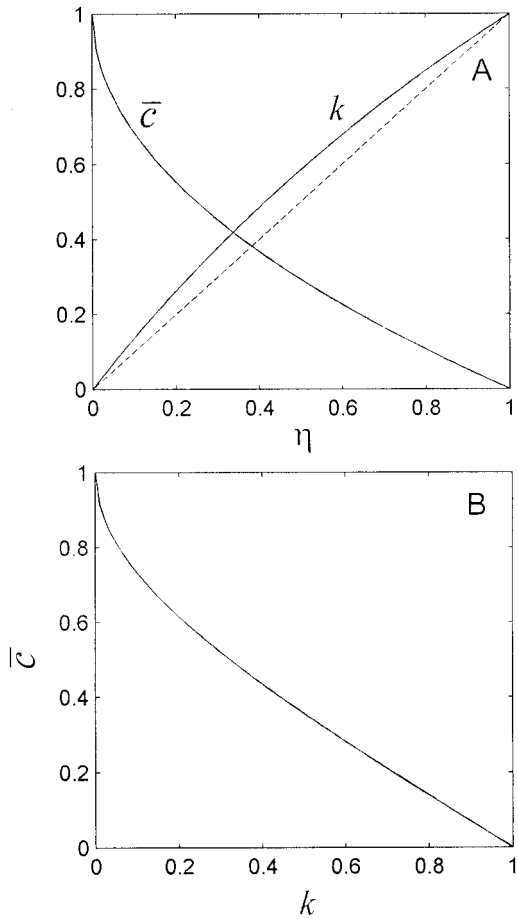

Figure 10. Comparison between three measures of onset dominance. A: How k, a measure of binaural adaptation (Hafter & Dye, 1983), and the metric $\bar{c}$ (Shinn-Cunningham et al., 1993) are related to efficiency η (Tanner & Birdsall, 1958). The relation between $\bar{c}$ and η is meaningful only for stimuli that yield a single image because $\bar{c}$ does not apply to dual-image percepts. B: How $\bar{c}$ relates to k.

identical ranges (zero to one) and are positively correlated. From the right-hand side of Equation 15, for example, one can calculate that if the detection index for the second pulse is half that of the single-pulse control, then the efficiency with which the binaural system receives information from the second pulse is 0.25, whereas k has a slightly larger value at 0.32.

The second model previously used to quantify the magnitude of the precedence effect is that described by Shinn-Cunningham et al. (1993). Although this model is not designed to make predictions of the magnitude of precedence as a function of T, it does describe a metric $\bar{c}$ that relates the empirically measured change in the perceived position of a dual-pulse composite (τ_{comp}) to the interaural time differences of the first (τ_1) and second (τ_s) pulses:

$$\bar{c} = \frac{\bar{\tau}_{comp} - \tau_s}{\tau_1 - \tau_s}. \tag{16}$$

For the paradigm used in the current study the interaural time difference of the first pulse is zero and thus

$$\bar{\tau}_{comp} = (1 - \bar{c})\tau_s. \tag{17}$$

It is important to point to a distinction between the measure of position judgments based on the Shinn-Cunningham et al. (1993) metric $\bar{c}$ and that of the current study. The metric $\bar{c}$ assumes a single composite image that is a weighted average of the interaural time differences of the lead and lag pulses. If the two pulses are equal in effectiveness, $\bar{c}$ has a value of 0.5. If the lag pulse dominates completely the perceived position of the dual-pulse composite, $\bar{c}$ has a value of 0. Thus, $\bar{c}$ is a centroid measure, whereas η refers to the efficiency of information transmission from only the lag pulse. The implications of this distinction are important especially when a dual-pulse waveform produces a dual-image percept, as is the case for IPI $\geq$ 10 ms, and often for shorter IPI between 5 ms and 10 ms. The $\bar{c}$ metric cannot apply to dual images, whereas η allows for a dual-image percept. If a dual image is perceived with equal effectiveness of the lead and lag pulses (i.e., no precedence) and subjects are required to point to the centroid of this composite according to assumptions of the $\bar{c}$ metric, the pointer interaural time difference would be half that of the lag, whereas, on the basis of the assumptions of the current model, the pointer would be adjusted exactly to the lag-pulse interaural time difference. When a single image is perceived and when the lead interaural time difference is zero, any displacement of the image away from zero is necessarily based on the interaural time difference of the lag pulse. For this case, the position judgments are effectively the same in the current study as that used by Shinn-Cunningham et al. (1993; i.e., $\bar{\tau}_{comp} = \bar{\tau}_{\hat{p}}$) or

$$\bar{\tau}_{\hat{p}} = (1 - \bar{c})\tau_s. \tag{18}$$

If we make the simplifying assumption (see Footnote 5) that the position estimate for a single pulse is equal to its interaural time difference, as would be the case when the pointer and signal pulses are the same type of stimuli ($\alpha = 1$ in Equation 6), then from Equation 9,

$$\eta = \left(\frac{\hat{d}_s}{d'_{ideal}}\right)^2 = \left(\frac{\sqrt{2}\Delta\bar{\tau}_{\hat{p}}/\sigma(\tau_s)}{\sqrt{2}\Delta\tau_s/\sigma(\tau_s)}\right)^2 = \left(\frac{\Delta\bar{\tau}_{\hat{p}}}{\Delta\tau_s}\right)^2, \tag{19}$$

and because $\Delta\tau = 2\tau$ for all conditions, Equation 19 yields

$$\bar{\tau}_{\hat{p}} = \tau_s\sqrt{\eta}. \tag{20}$$

Combining Equations 18 and 20 yields the relationship between Tanner and Birdsall's (1958) measure of efficiency η and Shinn-Cunningham et al.'s (1993) metric of onset dominance $\bar{c}$:

$$\bar{c} = 1 - \sqrt{\eta}. \tag{21}$$

It should be noted that because of the way in which $\bar{c}$ is defined, this relation applies only to cases where the dual-pulse waveform produces a single-image percept. Figure 10A shows the relation between η and $\bar{c}$ in comparison with that between η and k. Whereas k is directly related to η, $\bar{c}$ is an inverse function of η. The relation between $\bar{c}$ and k may also be specified by combining Equations 21 and 14:

$$\bar{c} = 1 - \sqrt{2^k - 1}. \tag{22}$$

Figure 10B shows that this function is, not surprisingly, similar to that between $\bar{c}$ and η, given the close correspondence between η and k.

The metric $\bar{c}$ has certain features that are consistent with and some that are dissimilar to our model. We have adopted a signal-detection approach in which the strength of onset dominance is quantified as the efficiency with which the binaural system receives information from a pulse preceded by another. Ideal efficiency is that for which the temporal separation between the lead and lag pulse is infinity. The model applies this efficiency measure only to the lag pulse. The metric $\bar{c}$ represents a weighted average of the interaural time differences of the lead and lag pulses. In the current study we set the lead interaural time difference to zero. If this time difference has a nonzero value, the derivations are simply offset by a constant equal to the lead time difference. Whereas the $\bar{c}$ metric is a centroid measure, η represents an efficiency applied solely to the second impulse. As noted above, the implication of this difference is most evident for values of IPI that lead to a dual percept. Whereas for short IPIs of less than 5 ms, observers usually report a single fused percept, at larger IPIs, a double image is often reported, corresponding to lead and lag impulses. Because the efficiency measure applies only to the second image, we had instructed our observers to adjust pointers to this image when two images were perceived. The centroid measure $\bar{c}$, however, estimates a weighted average of these two. Shinn-Cunningham et al. (1993) acknowledged this feature and noted that at an IPI of 10 ms, at least 1 of their subjects displayed idiosyncratic patterns of $\bar{c}$ attributable to dual-image percepts. This particular observer is reported to have selected one or the other image on different trials (see also the nonmonotonic patterns of $\bar{c}$ in their Figure 6). A second strategy, which other observers may use, is to adjust a pointer to a centroid position even when two images are perceived. Such cases, where dual images are perceived but a single weighted response is required, affect the interpretation of the metric $\bar{c}$. Others have also pointed to shortcomings in the use of $\bar{c}$ as a measure of onset dominance. Tollin and Henning (1998a), for example, pointed out that the $\bar{c}$ metric lacks a common frame of reference to judge the absolute size of echo inhibition (see their Footnote 2). They proposed an alternative measure that they referred to as threshold elevation factor (TEF), which compares the ratio of thresholds for echo with single-click conditions and is an inverse monotonic function of efficiency. The TEF has many desirable properties that allow comparison across different condi-

tions and estimation of an absolute effect of echo inhibition. Efficiency has the additional appeal of a well-established universal measure based on the ratio of detection indexes that normalize across not only experimental conditions, but across psychophysical paradigms.

The model described in the present article also does not preclude the idea that the perceived position of the lead pulse may vary, because it clearly does. However, our analysis concerns the efficiency of information transmission by the second impulse. In principle, nearly identical analysis may be applied to the effects of lag pulse on the percept of the lead pulse, with different model parameter values. It is obvious that in cases where a single image is perceived, that is, for short IPIs, the perceived position of lead and lag are identical.

The issue of perceptual fusion is interesting and merits some consideration. When the IPI is long enough to produce two images (i.e., greater than 3–4 ms), the echo image will have a distinct perceived position. This position will depend on the IPI as well as τ_s. The shorter the IPI, the closer will the image associated with a fixed τ_s be to the midline because the onset pulse does not carry an interaural time difference in our experiments. We have illustrated that the distribution of perceived echo positions has a mean that depends on IPI and a constant variance, independent of IPI. The model fits these descriptions, allowing a straightforward calculation of efficiency for the two-image case. If, however, a single fused image is perceived, then one must consider how the position mean and variance of this image are affected by τ_s within the context of our model. We consider the fused-image condition as a special case for which an efficiency value may readily be derived. To determine this efficiency, it is appropriate to describe the effects of the echo τ_s on the perceived position of the fused image. This image will have a single perceived position, and the extent to which its mean is displaced from the midpoint of the interaural axis reflects uniquely the influence of the echo τ_s. That this position is described as a normal deviate with a constant variance is an empirical observation that appears to hold for both a separate or fused echo. The use of an efficiency measure, representing the degree of information transmission from the echo waveform, is suitable here because the mean image displacement from zero midline is based only on *signal magnitude* (i.e., the echo τ_s). That there may be fusion is inconsequential to calculation of η or predictions of the model. η simply reflects signal transmission and makes no statement on other aspects of the process, which is part of its appeal as a universal measure. As an aside, we should note that in the model we have not specified the source of the internal noise that limits the accuracy of position estimation and whether it may be associated with the onset or echo pulse. In fact, one may speculate that this limiting noise is primarily central (downstream) to the initial encoding of the onset and echo images. Such a central-noise process would certainly be consistent with the constant variance of position estimates that we have empirically observed and independent of whether the images are fused or not.

Finally, two other dissimilarities between the three measures of onset dominance described above are worth considering. First, the function that relates onset dominance to IPI has not been specified in the Shinn-Cunningham et al. (1993) or Hafter et al. (1988) models, although both $\bar{c}$ and k are clearly primarily affected by IPI. In the current model we describe a logistic relation between the IPI, detection, and efficiency. The second difference between the three measures is related to the assumption of the underlying variances of position judgments and/or detection. The $\bar{c}$ metric assumes a constant variance of internal noise at all interaural time differences and IPIs in making predictions of discrimination performance, whereas both k and the current model assume a dependency of the variance of position judgments on interaural time difference τ_s, consistent with earlier work (Hafter & De Maio, 1975). This dependency affects predictions for detection or discrimination but not efficiency η or k.

Concluding Remarks

We described here a model of onset dominance in lateralization in which the perceived position of an echo pulse of a two-pulse waveform is a normal deviate whose mean is scaled by a logistic function of the IPI and whose variance depends on the echo interaural time difference. Simulations showed that this model accurately predicts both pointing judgments as well as detection performance for a number of combinations of interaural time difference and IPI. We have also suggested that the magnitude of onset dominance be quantified as a detection-theory efficiency measure. The advantage of using this measure compared with previously described measures of onset dominance is that it may be compared with efficiencies of other auditory processes. For echo inhibition, the rate at which efficiency declines with IPI is specified by a single parameter that depends on a number of factors such as the type of stimulus (e.g., transient, noise bursts, speech) or individual differences. Finally, we should note that the model has been designed and tested for a lateralization paradigm and an interaural delay cue. It would be of interest to investigate its application to free field-presented stimuli and real-world sounds since echo inhibition also holds for vertically distributed sources.

References

Allan, L. G. (1968). Visual position discrimination: A model relating temporal and spatial factors. *Perception & Psychophysics, 4,* 267–278.

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America, 43,* 648–657.

Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience.* Commack, NY: Nova Science.

Berg, B. G. (1990). Observer efficiency and weights in multiple observation tasks. *Journal of the Acoustical Society of America, 86,* 1743–1746.

Bernstein, L. R., & Trahiotis, C. (1985). Lateralization of low-frequency, complex waveforms: The use of envelope-based temporal disparities. *Journal of the Acoustical Society of America, 77,* 1868–1880.

Blauert, J. (1971). Localization and the law of the first wavefront in the median plane. *Journal of the Acoustical Society of America, 50,* 466–470.

Blauert, J. (1989). Binaural technology: Fundamentals and applications. *Journal of the Acoustical Society of America, 86,* S66.

Blauert, J., & Col, J. (1991). Irregularities in the precedence effect. In Y. Cazals, L. Demany, & K. Horner (Eds.), *Auditory physiology and perception: Proceedings of the Ninth International Symposium on Hearing* (pp. 531–538). Oxford, England: Pergamon Press.

Boyer, W. N., Cross, H. A., Guyot, G. W., & Washington, D. M. (1970). A TSD determination of a DL using two-point tactual stimuli applied to the back. *Psychonomic Science, 21,* 195–196.

Cain, W. S. (1977, February). Differential sensitivity for smell: "noise" at the nose. *Science, 195,* 796–798.

Clifton, R. K. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America, 82,* 1834–1835.

Colburn, H. S. (1973). Theory of binaural interaction based on auditory nerve data: I. General strategy and preliminary results on interaural

discrimination. *Journal of the Acoustical Society of America, 54,* 1458–1470.

Cox, R. H., & Hawkins, H. L. (1976). Application of theory of signal detectability to kinesthetic discrimination tasks. *Journal of Motor Behavior, 8,* 225–232.

Cranford, J., & Oberholtzer, M. (1976). Role of neocortex in binaural hearing in the cat: II. The "precedence effect" in sound localization. *Brain Research, 111,* 225–239.

Cross, H. A., Boyer, W. N., & Guyot, G. W. (1970). Determination of a DL using two-point tactual stimuli: A signal-detection approach. *Psychonomic Science, 21,* 198–199.

Dai, H. P. (1994). Signal-frequency uncertainty in spectral-shape discrimination—psychometric functions. *Journal of the Acoustical Society of America, 96,* 1388–1396.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision.* New York: Oxford University Press.

Domnitz, R. H. (1973). A headphone monitoring system for binaural experiments below 1 kHz. *Journal of the Acoustical Society of America, 58,* 510–511.

Egan, J. P. (1965). Masking-level differences as a function of interaural disparities in intensity of signal and of noise. *Journal of the Acoustical Society of America, 38,* 1043–1049.

Fitzpatrick, D. C., Kuwada, S., Batra, R., & Trahiotis, C. (1995). Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *Journal of Neurophysiology, 74,* 2469–2486.

Fitzpatrick, D. C., Kuwada, S., Kim, D. O., Parham, K., & Batra, R. (1999). Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex. *Journal of the Acoustical Society of America, 106,* 3460–3472.

Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review, 107,* 768–785.

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.

Gardner, M. B. (1968). Historical background of the Haas and/or precedence effect. *Journal of the Acoustical Society of America, 43,* 1243–1248.

Gaskell, H. (1983). The precedence effect. *Hearing Research, 12,* 277–303.

Geldard, F. A., & Sherrick, C. E. (1986). Space, time, and touch. *Scientific American, 255,* 91–95.

Gescheider, G. A., Sager, L. C., & Ruffolo, L. J. (1975). Simultaneous auditory and tactile information processing. *Perception & Psychophysics, 18,* 209–216.

Green, D. M. (1990). Stimulus selection in adaptive psychophysical procedures. *Journal of the Acoustical Society of America, 87,* 2662–2674.

Green, D. M. (1995). Maximum-likelihood procedures and the inattentive observer. *Journal of the Acoustical Society of America, 97,* 3749–3760.

Green, D. M., & Luce, R. D. (1975). Parallel psychometric functions from a set of independent detectors. *Psychological Review, 82,* 483–486.

Green, D. M., McKey, M. J., & Licklider, J. C. R. (1959). Detection of a pulsed sinusoid in noise as a function of frequency. *Journal of the Acoustical Society of America, 31,* 1446–1452.

Green, D. M., & Swets, J. (1988). *Signal detection theory and psychophysics.* Los Altos, CA: Peninsula. (Original work published 1966)

Haas, H. (1972). The influence of a single echo on the audibility of speech. *Journal of the Audio Engineering Society, 20,* 145–159. (Original work published 1949)

Haessly, A., Sirosh, J., & Miikkulainen, R. (1995). A model of visually guided plasticity of the auditory spatial map in the barn owl. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 154–158). Hillsdale, NJ: Erlbaum.

Hafter, E. R., & Buell, T. N. (1990). Restarting the adapted binaural system. *Journal of the Acoustical Society of America, 88,* 806–812.

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.

Hafter, E. R., & De Maio, J. (1975). Difference threshold for interaural delay. *Journal of the Acoustical Society of America, 57,* 181–187.

Hafter, E. R., & Dye, R. H. (1983). Detection of interaural differences of time in trains of high-frequency pulses as a function of interpulse interval and number. *Journal of the Acoustical Society of America, 73,* 1708–1713.

Hafter, E. R., Dye, R. H., & Wenzel, E. (1983). Detection of interaural differences of intensity in trains of high-frequency pulses as a function of interpulse interval and number. *Journal of the Acoustical Society of America, 73,* 644–651.

Hartung, K., & Trahiotis, 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.

Jesteadt, W., Reimer, J. F., Schairer, K., & Nizami, L. (2001). Psychometric functions for detection of tones following a forward masker. *Abstracts of the 2001 Midwinter Meeting of the Association of Research in Otolaryngology* (Abstract No. 884).

Koch, C. (1999). *Biophysics of computation: Information processing in single neurons.* New York: Oxford University Press.

Laming, D. (1986). *Sensory analysis.* London: Academic Press.

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.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America, 106,* 1633–1654.

Litovsky, R. Y., & Macmillan, N. A. (1994). Sound localization precision under conditions of the precedence effect—effects of azimuth and standard stimuli. *Journal of the Acoustical Society of America, 96,* 752–758.

Litovsky, R. Y., Rakerd, B., Yin, T. C. T., & Hartmann, W. M. (1997). Psychophysical and physiological evidence for a precedence effect in the median saggital plane. *Journal of Neurophysiology, 77,* 2223–2226.

Lutfi, R. A. (1995). Correlation-coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks. *Journal of the Acoustical Society of America, 97,* 1333–1334.

McFadden, D. (1973). Precedence effect and auditory cells with long characteristic delays. *Journal of the Acoustical Society of America, 54,* 528–530.

Mulligan, R. M., & Shaw, M. L. (1980). Multimodal signal detection: Independent decisions vs. integration. *Perception & Psychophysics, 28,* 471–478.

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.

Perrott, D. R., Marlborough, K., Merrill, P., & Strybel, T. Z. (1989). Minimum audible angel thresholds obtained under conditions in which the precedence effect is assumed to operate. *Journal of the Acoustical Society of America, 85,* 282–288.

Pickles, J. O. (1988). *An introduction to the physiology of hearing* (2nd ed.). San Diego, CA: Academic Press.

Rakerd, B., & Hartmann, W. M. (1997). The Haas effect with and without binaural differences. *Journal of the Acoustical Society of America, 101,* 3083.

Richards, V. M., & Zhu, S. P. (1994). Relative estimates of combination weights, decision criteria, and internal noise based on correlation-coefficients. *Journal of the Acoustical Society of America, 95,* 423–434.

Saberi, K. (1995). Some considerations on the use of adaptive methods for estimating interaural-delay thresholds. *Journal of the Acoustical Society of America, 98,* 1803–1806.

Saberi, K. (1996). Observer weighting of interaural delays in filtered impulses. *Perception & Psychophysics, 58,* 1037–1046.

Saberi, K., & Antonio, J. (2003). Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval. *Journal of the Acoustical Society of America, 114,* 420–429.

Saberi, K., & Green, D. M. (1996). Adaptive psychophysical procedures and imbalance in the psychometric function. *Journal of the Acoustical Society of America, 100,* 528–536.

Saberi, K., & Green, D. M. (1997). Evaluation of maximum-likelihood estimators in nonintensive auditory psychophysics. *Perception & Psychophysics, 59,* 867–876.

Saberi, K., & Perrott, D. R. (1990). Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. *Journal of the Acoustical Society of America, 87,* 1732–1737.

Saberi, K., & Perrott, D. R. (1995). Lateralization of pulse-trains with opposing onset and ongoing interaural delays. *Acustica, 81,* 272–275.

Saberi, K., Sadralodabai, T., & Perrott, D. R. (1991). Judgments of lateral distance using transients presented with interaural differences of time. *Bulletin of the Psychonomic Society, 29,* 59–61.

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.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., pp. 739–811). New York: Wiley.

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.

Sperling, G., Reeves, A., Blaser, E., Lu, Z.-L., & Weichselgartner, E. (2001). Two computational models of attention. In J. Braun, C. Koch, & J. L. Davis (Eds.), *Visual attention and cortical circuits* (pp. 177–214). Cambridge, MA: MIT Press.

Stellmack, M. A., Dye, R. H., & Guzman, S. J. (1999). Observer weighting of interaural delays in source and echo clicks. *Journal of the Acoustical Society of America, 105,* 377–387.

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., & Trahiotis, C. (1988). Lateralization of complex binaural stimuli: A weighted-image model. *Journal of the Acoustical Society of America, 84,* 156–165.

Stroupe, A. W., Martin, M. C., & Balch, T. (2001). Merging Gaussian distributions for object localization in multi-robot systems. *Experimental Robotics VII: Lecture Notes in Control and Information Sciences, 271,* 343–352.

Swets, J. A. (1984). Mathematical models of attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 183–242). Orlando, FL: Academic Press.

Tanner, W. P., & Birdsall, T. G. (1958). Definitions of d' and h as psychophysical measures. *Journal of the Acoustical Society of America, 30,* 922–928.

Tollin, D. J., & Henning, G. B. (1998a). Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect. *Journal of the Acoustical Society of America, 104,* 3030–3038.

Tollin, D. J., & Henning, G. B. (1998b). Some aspects of the lateralization 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.

Trahiotis, C., & Stern, R. M. (1989). Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase. *Journal of the Acoustical Society of America, 86,* 1285–1293.

van Trees, H. L. (1968). *Detection, estimation, and modulation theory, Part I.* New York: Wiley.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *American Journal of Psychology, 52,* 315–336.

Yin, T. C. T., & Litovsky, 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.

Yost, W. A. (1997). The cocktail party problem: Forty years later. In R. Gilkey & T. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 329–348). New York: Erlbaum.

Yost, W. A., & Soderquist, D. R. (1984). The precedence effect—revisited. *Journal of the Acoustical Society of America, 76,* 1377–1383.

Zurek, P. M. (1980). The precedence effect and its possible role in the avoidance of interaural ambiguities. *Journal of the Acoustical Society of America, 67,* 952–964.

Zurek, P. M. (1987). The precedence effect. In W. A. Yost & G. Gourevitch (Eds.), *Directional hearing* (pp. 85–105). New York: Springer-Verlag.

Zurek, P. M., & Saberi, K. (2003). Lateralization of two-transient stimuli. *Perception & Psychophysics, 65,* 95–106.

Received October 22, 2001
Revision received November 28, 2002
Accepted December 13, 2002 ■