

Physiological and Psychophysical Modeling of the Precedence Effect

JING XIA¹, ANDREW BRUGHERA², H. STEVEN COLBURN², AND BARBARA SHINN-CUNNINGHAM^{1,2}

¹*Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA*

²*Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA*

Received: 21 October 2009; Accepted: 25 February 2010; Online publication: 1 April 2010

ABSTRACT

Many past studies of sound localization explored the precedence effect (PE), in which a pair of brief, temporally close sounds from different directions is perceived as coming from a location near that of the first-arriving sound. Here, a computational model of low-frequency inferior colliculus (IC) neurons accounts for both physiological and psychophysical responses to PE click stimuli. In the model, IC neurons have physiologically plausible inputs, receiving excitation from the ipsilateral medial superior olive (MSO) and long-lasting inhibition from both ipsilateral and contralateral MSOs, relayed through the dorsal nucleus of the lateral lemniscus. In this model, physiological suppression of the lagging response depends on the inter-stimulus delay (ISD) between the lead and lag as well as their relative locations. Psychophysical predictions are generated from a population of model neurons. At all ISDs, predicted lead localization is good. At short ISDs, the estimated location of the lag is near that of the lead, consistent with subjects perceiving both lead and lag from the lead location. As ISD increases, the estimated lag location moves closer to the true lag location, consistent with listeners' perception of two sounds from separate locations. Together, these simulations suggest that location-dependent suppression in IC neurons can explain the behavioral phenomenon known as the precedence effect.

Keywords: computational model, localization dominance, echo threshold, neural correlates, inferior colliculus

Abbreviations: AN – Auditory nerve; CF – Characteristic frequency; DNLL – Dorsal nucleus of the lateral lemniscus; IC – Inferior colliculus; IHC – Inner hair cell; IPD – Interaural phase delay; ISD – Inter-stimulus delay; ITD – Interaural time delay; LSO – Lateral superior olive; MLE – Maximum likelihood estimate; MSO – Medial superior olive; PE – Precedence effect; PST – Peri-stimulus time; SMAX – Suppression-at-maximum; SMIN – Suppression-at-minimum

INTRODUCTION

Listeners have a remarkable ability to localize sounds accurately in reverberant settings, a feat attributed to the fact that they give greater perceptual weight to location cues at sound onsets, suppressing cues from later-arriving reflections (Zurek 1980; Freyman et al. 1997; Devore et al. 2009). This “precedence effect” (PE; Wallach et al. 1949) has been studied using a range of paradigms whereby a pair of dichotic clicks is presented with a brief inter-stimulus delay (ISD). Psychophysical studies in humans (e.g., Zurek 1980; Freyman et al. 1991; Litovsky and Shinn-Cunningham 2001) and nonhuman species (e.g., Kelly 1974; Cranford 1982; Wyttenbach and Hoy 1993; Keller and Takahashi 1996) reveal different phases of the PE (Blauert 1997). *Summing localization*, in which listeners perceive one single fused auditory image located somewhere between the lead and lag sources (often biased towards the leading source), occurs for ISDs

Correspondence to: Barbara Shinn-Cunningham · Department of Cognitive and Neural Systems · Boston University · 677 Beacon Street, Boston, MA 02215, USA. Telephone: +1-617-3535764; fax: +1-617-3537755; email: shinn@cns.bu.edu

from 0 to 1 ms. *Localization dominance*, in which the fused image is localized near the lead, occurs for ISDs ranging from 1 to 10 ms. For ISDs greater than 10 ms, the *echo threshold* (the shortest ISD at which two separate auditory images are heard) is reached, but the perceived locations of the two sounds are both often close to the lead. For ISDs considerably longer than the echo threshold, listeners localize lead and lag independently, at the locations where lead and lag would be perceived in isolation (Litovsky and Shinn-Cunningham 2001).

Neural correlates of the PE are observed in extracellular responses in inferior colliculus (IC; Carney and Yin 1989; Yin 1994; Fitzpatrick et al. 1995; Litovsky and Yin 1998a, b; Tollin et al. 2004). For small ISDs, neural responses to the lag are reduced or eliminated, consistent with the leading source dominating sound localization and the lagging source location having little influence on perception. Neural responses in awake animals recover at ISDs comparable to psychophysically measured echo thresholds (Fitzpatrick et al. 1995; Tollin et al. 2004).

Some previous models have simulated behavioral aspects of the PE (Lindemann 1986; Tollin and Henning 1999; Hartung and Trahiotis 2001; Dizon and Colburn 2006). The current study accounts for both physiological and behavioral PE results using a population of biologically plausible, model IC neurons. Responses of single IC neurons are simulated for pairs of binaural click stimuli with ISDs spanning the range from localization dominance to echo threshold, where longer-term inhibition is the dominant factor. Responses from the model population are combined to predict the perceived location of paired clicks. Results are compared with corresponding physiological and psychophysical data. Finally, we comment on some open questions about our modeling approach, including the role of inhibitory pathways to IC, the contributions of cochlear interactions of low-frequency neurons in PE conditions, and the possible explanations for minor discrepancies between simulated and observed behavioral results.

METHODS

Stimuli

Stimuli consist of pairs of binaural clicks (the lead and the lag) separated by different ISDs. The ISD was defined as the time difference between the onsets of the lead and lag clicks delivered to the right ear. Interaural time delays (ITDs) were imposed separately on the lead and lag binaural clicks. Positive ITDs were generated by advancing the stimulus in the ear contralateral to the model cell (right).

The click stimuli were generated in MATLAB at a sampling rate of 20 kHz using square pulses of 50- μ s duration passed through a 5th-order Butterworth low-pass filter with a cutoff frequency of 2 kHz. The intensity of a single, monaural click was set to 70-dB peak-equivalent SPL. Binaural PE stimuli were generated by adding the desired ITDs to the lead and lag clicks, superimposing the resulting binaural lead and lag stimuli, and then presenting the resulting left- and right-ear stimuli to the auditory nerve model. For each model neuron and tested stimulus, 50 repetitions of the paired PE stimuli were presented to the model to generate peri-stimulus time (PST) histograms. These average results were used to predict both physiological and behavioral results.

Model structure

The model (Fig. 1) consists of a hierarchy of processing stages, mimicking the stages of the auditory periphery, brainstem, and midbrain. Our IC model neurons, which are based on a previous IC model (Cai et al. 1998), are innervated by medial superior olive (MSO) model neurons (Brughera et al. 1996), with inhibitory inputs that reach the IC via the dorsal nucleus of the lateral lemniscus (DNLL). The relative strength of inhibition from ipsi- vs. contra-lateral DNLL determines the type of the IC cell (suppression-at-maximum (SMAX) or suppression-at-minimum (SMIN), see below). The MSO model neurons receive bilateral inputs from the model bushy cells in the cochlear nucleus (Rothman et al. 1993), which receive convergent inputs from the model auditory-nerve (AN) fibers (Carney 1993). In our modeling, we assumed that the left and right ICs were mirror symmetric. By convention, half of the individual model neurons comprising the IC population model were in the left IC (i.e., responding primarily to sources with positive ITDs, leading in the right ear), and the other half were in the right IC (i.e., responding primarily to negative ITDs, leading in the left ear). Both left and right IC responses were combined to generate the total population response, which contributed to predictions of perceived locations.

Auditory nerve model. The Carney (1993) model is used to generate neural spikes of low-frequency AN fibers in response to the stimuli. The frequency tuning around the characteristic frequency (CF) of the AN fiber is determined by a nonlinear band-pass filter, as described by Carney (1993). The output of the filter is then passed through models of the inner hair cell (IHC, a compressive nonlinearity) and the IHC-AN synapse (which incorporates adaptation and refractoriness) to simulate the corresponding processing stages in the cochlea. A non-homogeneous Poisson process model is

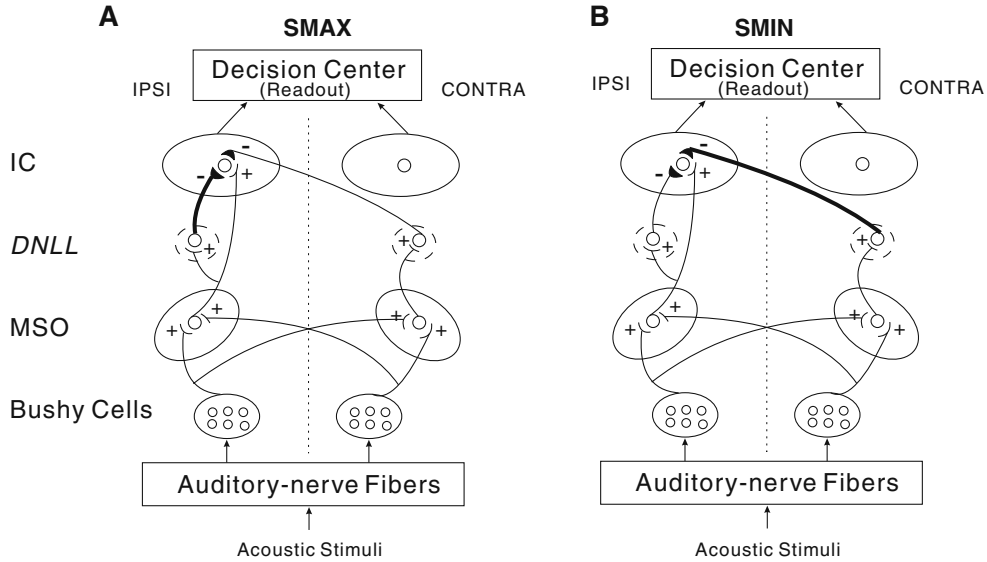


FIG. 1. Structure of the inferior colliculus (IC) model, which incorporates models of medial superior olive (MSO) neurons, bushy cells in cochlear nucleus, and auditory-nerve fibers. The DNLL is included in the model only as a relay mechanism for generating delayed inhibitory input to the IC from the MSO. The details of DNLL

cell behaviors are not included in the model. Excitatory synapses are marked by *plus signs* and inhibitory synapses by *minus signs*. **A** The structure of SMAX model neuron, which has strong ipsilateral inhibition. **B** The structure of SMIN model neuron, which has strong contralateral inhibition.

used to generate random spike times of the AN fiber. All of the parameter values used here are identical to the Carney (1993) model used in Brughera et al. (1996) and Cai et al. (1998).

Bushy cell model. The temporal information in low-frequency AN-fiber responses is enhanced in the responses of model spherical bushy cells in the anteroventral cochlear nucleus. These cells, which provide input to the next model stage (MSO), have primary-like PST histograms with high synchronization indices in response to low-frequency tone-burst stimuli at CF (Joris et al. 1994). The activity of spherical bushy cells is produced by the Rothman et al. (1993) model, which is based on the assumption that the soma of the bushy cell is uniform and adendritic. The model membrane contains three voltage-dependent ion channels (a low-threshold, slow potassium channel, B ; a fast potassium channel, K ; and a fast sodium channel, Na) as well as a voltage-independent leakage channel (L). The membrane potential (V) is determined by the currents of these channels as well as those of the excitatory and inhibitory synaptic inputs:

$$C \frac{dV}{dt} + G_B(V - E_K) + G_K(V - E_K) + G_{Na}(V - E_{Na}) + G_L(V - E_L) + G_I(V - E_I) + G_E(V - E_E) = I_{ext} \quad (1)$$

where C is the membrane capacitance, the subscripted- E parameters are the reversal potentials of the corresponding channels, and I_{ext} is the applied external current, which is set to zero throughout this

study. The conductances of the low-threshold slow potassium channel (G_B), the fast potassium channel (G_K), and the fast sodium channel (G_{Na}) are described by Hodgkin and Huxley (1952) type equations. The time course of the excitatory synaptic conductance (G_E) in response to a single input discharge is described by the following alpha function with a time constant of τ_{ex} :

$$G_E(t - t_0) = G_{E \max} \frac{t - t_0}{\tau_{ex}} \times \exp \left[1 - \frac{t - t_0}{\tau_{ex}} \right] u(t - t_0). \quad (2)$$

The conductance starts to increase when an input action potential arrives at time t_0 , reaching its maximum value $G_{E \max}$ at time $t_0 + \tau_{ex}$. The model cell receives excitatory inputs from 25 model AN fibers with a CF of 500 Hz. The convergence of many excitatory inputs onto one model spherical bushy cell, each of which is individually too weak to depolarize the bushy cell, results in responses with very high synchrony that support and enhance the close dependence of the MSO discharge rate on the stimulus ITD. There are no inhibitory inputs in the implemented bushy cell model. Further details concerning supporting equations of the Hi-Sync bushy cell model can be found in Rothman et al. (1993). The parameter values used here are the same as those used in Brughera et al. (1996) and Cai et al. (1998); these values are the same as those used by Rothman et al. (1993), except for β_n and α_n (Brughera, personal

communication). Brughera et al. (1996) modified these parameters to reduce regularity in inter-spike intervals (reduced via inhibition in Rothman et al. 1993). The modifications provide smoother transitions in β_n and α_n as a function of membrane voltage, a characteristic also present in a more recent model by Rothman and Manis (2003).

MSO model. The MSO is thought to be the initial site of low-frequency binaural interaction. MSO neurons are “tuned” (respond preferentially) both to a particular sound frequency and a particular ITD in input binaural stimuli. At least to a first-order approximation, this characteristic arises because these cells act as narrowband interaural “coincidence detectors” (Colburn et al. 1990; Joris et al. 1998), generating output spikes only if they receive nearly simultaneous neural spikes from matched frequency, narrowband ipsilateral, and contralateral excitatory spherical bushy cell inputs (note that inhibition has also been shown to influence the ITD tuning of MSO neurons; e.g., see Brand et al. 2002). The delays of spikes from the ipsi- and contralateral ears to a particular MSO neuron can differ; as a result, different model MSO neurons are characterized by different “best” ITDs (the ITD that leads to the maximal firing rate for a particular neuron), which is that ITD that compensates for difference in the neural transmission delays to the neuron from the ipsi- and contralateral ears. In the current study, we assume that the inputs reaching a model MSO neuron from the contralateral side are always delayed relative to ipsilateral inputs; therefore, the model MSO neurons prefer (respond most vigorously to) sound sources from the contralateral sound field. With our conventions, model neurons in the left MSO have positive best ITDs.

The current MSO model neurons are the same as those developed by Brughera et al. (1996), which received only Hi-Sync excitatory inputs from model bushy cells on both sides, leaving out any inhibitory inputs. The Hodgkin–Huxley point neuron model for spherical bushy cells (Rothman et al. 1993) is also used to model the MSO cell. This type of MSO model cell has been shown to be particularly sensitive to the relative timing of its inputs due to the contribution of a slow, low-threshold potassium channel (Brughera et al. 1996). The parameters for the ipsilateral and contralateral excitatory synaptic conductances, including the number of input neurons, the maximum value and the time constant of the synaptic conductance, and the delay of the input arrival, are identical to those used in Brughera et al. (1996).

IC model. Connections from model MSO cells to a model IC cell are based on anatomical and physiological evidence. Specifically, the MSO provides ipsilateral

projections to the IC (Henkel and Spangler 1983), while both the ipsilateral and contralateral DNLL provide GABAergic, inhibitory projections to the IC (Adams and Mugnaini 1984). Excitatory inputs from MSO to ipsilateral IC lead to ITD sensitivity in low-frequency IC cells (Kuwada and Yin 1983; Carney and Yin 1989; Loftus et al. 2004). Delayed inhibitory inputs to the IC from DNLL are thought to contribute to the neural correlates of the PE (Carney and Yin 1989; Yin 1994; Fitzpatrick et al. 1995; Litovsky and Yin 1998a, b; Litovsky and Delgutte 2002).

The differential equations describing the membrane potential of the model IC neurons are the same as those in the IC model developed by Cai et al. (1998), which are common to both principal cells in the MSO (Brughera et al. 1996) and the spherical bushy cells in the anteroventral cochlear nucleus (Rothman et al. 1993). In the Cai et al. (1998) IC model, the time course of the inhibitory synaptic conductance (G_I) is described by a linear summation of an alpha function and an exponential function. Detailed parameter specifications for excitatory and inhibitory synaptic conductances in the current IC model are provided in Table 1. Similar to the Cai et al. (1998) model, the time constant associated with inhibitory inputs are relatively long-lasting (several milliseconds) compared with the brief conductance changes evoked by excitatory inputs (tenth of milliseconds). Inhibitory inputs to the IC are delayed, presumably because of the extra synapse when passing through the DNLL. As a result, inhibition suppresses responses after sound onsets but has little effect on onset responses except for very short ISDs.

In contrast to the Cai et al. (1998) model (which receives inhibition only from contralateral MSO), directionally tuned parallel inhibition from both ipsilateral and contralateral MSOs (via the corresponding DNLL) are used in our IC model to produce suppression of the lagging response that depends strongly on the relative locations of lead and lag clicks in the acoustic inputs, consistent with responses observed in past physiological studies of the PE (Carney and Yin 1989; Yin 1994; Fitzpatrick et al. 1995). Some IC neurons show greatest suppression of the lagging response when the leading sound comes from the neuron’s best ITD. These neurons, known as SMAX neurons (Litovsky and Yin 1998b), receive more inhibition from the ipsilateral DNLL than from the contralateral DNLL in the model (see Fig. 1A and SMAX model parameters in Table 1). In contrast, SMIN neurons (Litovsky and Yin 1998b) show greatest suppression of the lag when the lead comes from positions that elicit little response. These neurons receive stronger inhibition from the contralateral DNLL in the model (see Fig. 1B and SMIN model parameters in Table 1). The parameters for the

TABLE 1

Parameters for a model SMAX and SMIN IC neuron			
	Excitation		Inhibition
Parameters (SMAX)			
Number of MSO inputs	1 (I)	1 (I)	1 (C)
Peak conductance (ns)	25	8	5
Time constant (ms)	0.1	3	2
Best ITD of MSO inputs (us)	300	300	-300
Delay of arrival (ms)	0	2	2
Parameters (SMIN)			
Number of MSO inputs	1 (I)	1 (I)	1 (C)
Peak conductance (ns)	25	5	8
Time constant (ms)	0.1	2	3
Best ITD of MSO inputs (us)	300	300	-300
Delay of arrival (ms)	0	2	2

C inputs come from the contralateral side of that neuron, I inputs come from the ipsilateral side of that neuron

ipsilateral and contralateral inhibitory channels of the model SMIN neuron are symmetrically opposite those of the model SMAX neuron. These parameters were constrained by the values suggested by Cai et al. (1998) and were chosen to fit the physiological responses of empirical SMAX and SMIN neurons observed in Litovsky and Yin (1998a, b) study.

Individual neuron analysis

Responses of model neurons were calculated as the running average of the number of spikes over 50 stimulus repetitions. The techniques used here to allocate responses to the leading and lagging stimulus were taken from past physiological studies (Fitzpatrick et al. 1995; Litovsky and Yin 1998a; Tollin et al. 2004). Specifically, to quantify responses, we analyzed the number of the spikes falling within a time window covering a fixed post-stimulus time (Fig. 3A). The temporal position of the analysis window for an input click was determined by model responses to an isolated binaural click. Specifically, the window began when the discharge rate in the output spike train first exceeded the mean spontaneous rate by at least two standard deviations (the mean spontaneous rate was computed over the 10 ms prior to the stimulus presentation) and ended when the response fell below the mean spontaneous rate. Response latency was defined as the start time of the analysis window.

For pairs of binaural clicks, two windows (the *leading window* and the *lagging window*) were used to calculate responses, based on the window latency and window duration determined from the isolated click input. In other words, the relative start and end times of the leading window were taken to be the same as those of the analysis window in response to a single click. The lagging window start time equaled the

response latency plus the ISD; the lagging window had the same duration as the leading window.

For large ISDs (e.g., Fig. 3B), the leading and lagging windows did not overlap; we were able to separately compute the response to the lead and lag from the spike counts in the leading and lagging window. For small ISDs (e.g., Fig. 3C), the two windows overlapped; in these cases, the response to the lag was estimated by subtracting the number of spikes in response to an isolated binaural click (identical to and with the same absolute timing as the lead) from the number of spikes counted from the onset of the leading window to the offset of the lagging window.

Population model analysis

The absolute discharge rate of a single IC neuron cannot account for the perceived location of acoustic inputs. The ITDs of the leading and lagging stimuli were estimated based on the responses of a population of model IC neurons with the same frequency tuning. The perceived location of the PE stimuli was hypothesized to be a weighted sum of the estimated leading and lagging ITDs. As described below, the relative weights given to the leading and lagging ITD estimates were assumed to depend on the reliability of the estimated ITDs, which were also computed directly from the population response of the model neurons.

In the current work, we concentrated on neurons with low CFs of 500 Hz and assumed that the population consisted of neurons whose best ITDs were uniformly distributed over a symmetrical range from -1 to $+1$ ms in 0.05 ms increments, for a total of 41 model neurons. A summary of the response \mathbf{P} of a population of neurons is given by a vector average of complex values, representing the individual neurons' responses (see Shinn-Cunningham and Kawakyu 2003). The magnitude of the complex value associated with a given IC neuron is given by the number of spikes (L) falling within the time window being considered (see individual neuron analysis, above); the phase of this complex value is determined by the best ITD (τ_m) of that neuron, so that the response of a particular neuron is given by the complex vector $L_{k,\tau_m} e^{j2\pi f \tau_m}$, where L_{k,τ_m} is the spike count falling in the *leading* or *lagging window* of the neuron tuned to τ_m . The complex average of these values (computed over best ITD) is the parameter \mathbf{P}_k specified here for each of the time windows (k =lead or lag):

$$\mathbf{P}_k = \frac{1}{2T} \sum_{\tau_m=-T}^{\tau_m=T} L_{k,\tau_m} e^{j2\pi f \tau_m}, \quad (3)$$

where $T=1$ ms and f is the CF of the peripheral band-pass filter ($f=500$ Hz). For pure sinusoidal inputs, the phase of \mathbf{P} gives the maximum likelihood estimate

(MLE) of the interaural phase delay (IPD) of a binaural input with an IPD that is constant overtime (Colburn and Isabelle 1992; Shinn-Cunningham and Kawakyu 2003); given the narrowband nature of the IC responses analyzed here, the phase of \mathbf{P} approximates the IPD MLE for broadband clicks estimated from the frequency band centered on f :

$$\varphi_k = \angle \mathbf{P}_k \quad (4)$$

While the angle of \mathbf{P} estimates IPD, the magnitude of \mathbf{P} varies with the reliability of the observed population response estimate (Shinn-Cunningham and Kawakyu 2003) and is directly related to the interaural correlation of the left and right ear inputs in response to a single sound source. We thus estimate the reliability of the IPD estimate in response to a lead or lag input as:

$$r_k = |\mathbf{P}_k|. \quad (5)$$

For the PE stimuli used here, we hypothesized that the perceived IPDs corresponding to the lead (θ_1) and the lag (θ_2) are weighted sums of the lead and lag IPDs estimated by Eq. 4:

$$\theta_i = c_i \varphi_{lead} + (1 - c_i) \varphi_{lag}, \quad (6)$$

where the weights (c_i) differ for estimates of lead and lag location ($i=1$ and 2 , respectively). In general, the weights depend on the reliability of the population response to the lead (or to the lag), relative to the reliability of the population response to the lead (or lag) alone. If the instruction is to match the location of the leading stimulus, $i=1$,

$$c_1 = \frac{r_{lead}}{r_{Slead}}, \quad (7)$$

while if the instruction is to match the location of the lagging stimulus, $i=2$,

$$c_2 = 1 - \frac{r_{lag}}{r_{Slag}}, \quad (8)$$

where r_{Slead} and r_{Slag} are the reliability measures estimated using Eq. 5 to a single binaural click presented in isolation at the leading or lagging locations (for r_{Slead} and r_{Slag} respectively). Finally, for the low-frequency that we examined ($f=500$ Hz), we assumed that the perceived ITD, α_i is given by

$$\alpha_i = \theta_i / f. \quad (9)$$

For high-frequency neurons, where the maximum IPD observed can exceed $2\pi f$, further assumptions must be made to resolve inherent ambiguity in the IPDs. The weights (c_i) are restricted to fall between 0 and 1. A value of $c_1=1$ (which occurs when the reliability of the estimate of the leading ITD, r_{lead} equals the reliability of the estimated ITD for the lead

click in isolation, r_{Slead}) indicates that the lead dominates lateralization entirely ($\alpha_1 = \varphi_{lead}/f$). Conversely, $c_2=0$ (which occurs when the reliability of the estimate of the lagging ITD, r_{lag} equals the reliability of the estimated ITD for the lag click in isolation, r_{Slag}) indicates that the lag dominates lateralization completely ($\alpha_2 = \varphi_{lag}/f$). We expect the response to the leading stimulus to be only modestly affected by the presence of the lag ($c_1 \approx 1$ and $\alpha_1 \approx \varphi_{lead}/f$); therefore, in the current study, the perceived leading location is expected to roughly equal the location of a single click at the lead location. If the instruction is to match the lagging ITD, the precedence effect is strong when the population response to the lagging stimulus is substantially suppressed ($c_2 \approx 1$ and $\alpha_2 \approx \varphi_{lag}/f$); in these cases, the model predicts that the lagging image is near the leading source. When the lagging stimulus begins to evoke a population response that closely resembles the neural response to the lagging source presented in isolation, the precedence effect is weak ($c_2 \approx 0$ and $\alpha_2 \approx \varphi_{lag}/f$); in these cases, the model predicts that the lagging image would be localized near the location of the lagging source.

RESULTS

In this section, we first illustrate MSO model responses in order to demonstrate that long-lasting DNLL inhibition is necessary to explain lead-lag interactions observable for ISDs longer than 5 ms, i.e., in the range studied in past physiological work. We then present simulations of the responses of a single IC neuron, as well as of a population of IC neurons. Finally, psychophysical predictions based on the physiological responses are compared to past behavioral results.

Model MSO cell's response to paired clicks

The model MSO cell's response to a single click lasts for up to 10 ms due to the ringing of the MSO inputs, coming from the 500-Hz AN. For ISDs shorter than 10 ms, the MSO will still be responding to the lead when the initial response to the lag begins. As a result, the lead and lag MSO responses are nearly impossible to separate from each other at short ISDs; the definitions of "lead" and "lag" responses adopted from past IC physiology studies (see "Individual neuron analysis," above) are not appropriate for analyzing the MSO results. Therefore, in order to examine the effect of the leading response on the lag at different ISDs, we calculated the MSO cell's response in a window whose start time equaled the response latency plus the ISD (i.e., at the moment the MSO response should be affected by the lag click)

and whose duration was the same as that used for analyzing responses to a single click.

Figure 2 shows the model MSO cell's response to the PE stimuli as well as to a single click (solid line), as a function of the leading ITD. The lagging click was held constant at the cell's best ITD ($+300 \mu\text{s}$). For ISDs shorter than 5 ms (Fig. 2A), the MSO cell's response was reduced when lead and lag had different ITDs, even though there is no inhibition in the MSO circuitry. The amount of the reduction depended on the ITD of the lead and the ISD and was generally stronger for shorter ISDs than for longer ISDs. Moreover, for the 500-Hz cell, the patterns of the ITD-rate curve for the 1- and 3-ms ISDs were similar (crosses and asterisks, respectively); the patterns for the 2- and 4-ms ISDs were also similar (triangles and squares, respectively).

This kind of suppression can be explained by peripheral interactions between the lead and lag responses at the AN, which shift the effective ITD at the time of the lag onset (Hartung and Trahiotis 2001; Trahiotis and Hartung 2002). As a result, the neuron whose best ITD matches that of the lag may not be active, since the effective ITD at the lag onset depends on interactions between the residual lead response and the onset of the lag response. This peripheral interaction depends critically on the ISD, the ITD, and the CF of the AN (Hartung and Trahiotis 2001). Specifically, such interactions should be similar for ISDs that differ by exactly one cycle of the best frequency of the neuron (here 500 Hz), but should be weaker for longer ISDs (as the lead response dies out). This explains why results for the 3-ms ISD were affected by the lead ITD in a way similar to, but weaker than, results for the 1-ms ISD (compare also the responses to 4- and 2-ms ISDs). The focus of the current paper is on lead-lag interactions that cannot be explained by peripheral interactions of this sort; thus, later analysis at the level of the IC focuses on longer ISDs.

As seen in Figure 2B, which plots the MSO response for longer ISDs, peripheral interactions are not influential beyond about 5 ms. Regardless of the

lead ITD, the MSO response to the PE stimuli was roughly equal to the response to a single click presented at the cell's best ITD (i.e., equal to the expected response to the lag alone, without any noticeable effect of the lead). Therefore, for ISDs longer than 5 ms, any suppression of the lag response at the level of the IC must be the result of inhibition not included in these MSO responses.

Simulations of physiological data

Model SMAX and SMIN neurons were constructed with best ITDs = $300 \mu\text{s}$ to simulate the response properties observed in physiological studies of PE. Two parameters were varied when pairs of binaural clicks were presented: (1) the ITD of the leading click and (2) the ISD between the leading and lagging clicks. The ITD of the lagging click was fixed at the best ITD of the model IC neuron. We varied the ITD of the leading click from -900 to $+900 \mu\text{s}$ in $150\text{-}\mu\text{s}$ increments. We concentrated on ISDs from 1 to 30 ms, which span the range of ISDs from localization dominance to echo threshold. We also presented isolated binaural clicks (with no lag) as controls.

Temporal responses of the model IC neuron to a single binaural click were not significantly different from previous observations by Carney and Yin (1989), who studied extracellular responses to broadband clicks of low-frequency neurons in the central nucleus of the IC in cat. Figure 3A shows the dot raster (top) and PST histogram (bottom) of a model cell's response to a single binaural click located at the cell's best ITD. The model cell responded to a binaural click with only one or two spikes at a latency of approximately 8 ms. The low-frequency model cell was characterized by its phase-locked response with the periodicity determined by its CF (500 Hz). For pairs of binaural clicks with a large ISD (e.g., ISD = 20 ms, as shown in Fig. 3B), clearly separated responses were seen corresponding to the lead and the lag. For this long ISD, the transient nature of the response to a click allowed individual discharges to be attributed to either the leading or the lagging

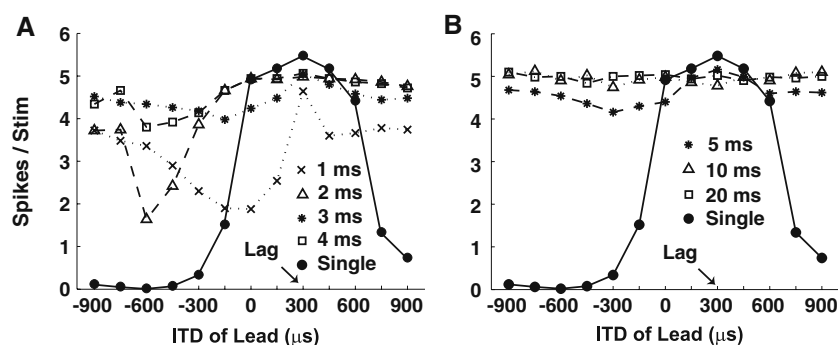


FIG. 2. Responses of a model MSO cell to paired clicks, as well as to a single click as a function of lead location. Arrow indicates the location of the lagging click, which is at the cell's best ITD. **A** Responses for ISDs of 1, 2, 3, and 4 ms. **B** Responses for ISDs of 5, 10, and 20 ms.

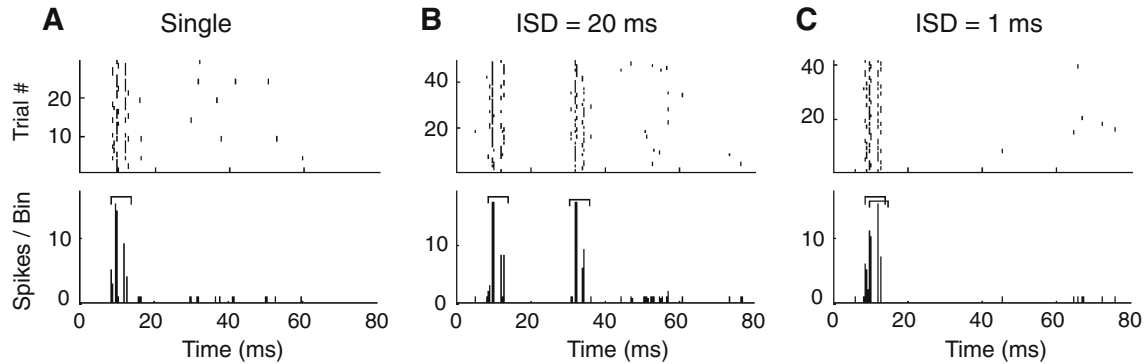


FIG. 3. Responses of a model IC cell. **A** Responses to a single binaural click with an ITD equal to the cell's best ITD, shown as a dot raster (*top*) and a summary histogram (*bottom*). The histogram bin width is 0.1 ms. An *analysis window* (\square) is defined to quantify the

response. **B, C** Responses to a pair of binaural clicks separated by ISDs of 20 ms (**B**) and 1 ms (**C**). Both leading and lagging clicks are located at the cell's best ITD. The *lagging window* is shifted by the ISD from the *leading window*.

stimulus. However, as ISD decreased (e.g., an ISD of 1 ms, as shown in Fig. 3C), the responses to the leading and lagging stimuli overlapped, making it impossible to assign clicks unambiguously to the lead or the lag input. Moreover, for short ISDs, portions of the response to the leading click are likely to be affected by the presence of the lag. Although it is difficult to allocate responses to the lead and lag separately for short ISDs, we adopted the same procedure used in previous physiological studies of the PE (see “[Individual neuron analysis](#),” above).

Varying the ISD between the lead and lag. Figure 4 shows responses of model SMAX (panels A and C) and SMIN (panels B and D) neurons to a pair of lead–lag clicks, both of which were located at the neuron's best ITD. Temporal discharge patterns of the responses are shown in Figure 4A and B (dot rasters as a function of post-stimulus time) for ISDs varying from 1 to 30 ms (y axis). Responses to a single binaural click are plotted at zero ISD. Figure 4C, D displays the spike counts in response to the lead and lag, respectively, as a function of ISD. These values were obtained by analyzing spikes in the leading and lagging windows, as described in the individual neuron analysis.

The responses to the lead were generally the same as the responses to a single binaural click presented in isolation, except for very short ISDs (about 1 ms) where the responses to the leading and lagging clicks overlapped. For ISDs shorter than 10 ms, the responses of both model SMAX and SMIN neurons were substantially reduced for the lag. As ISD increased, the lagging response recovered gradually, approaching that of the lead at ISDs of 20–30 ms. In physiological studies, the half-maximal ISD (as shown by dashed lines in Fig. 4C and D), which represents the ISD at which the lag response reaches 50% of the response to a single source at the lagging location

(see asterisk in far left of panels C and D), is hypothesized to be related to the psychophysical echo threshold (Tollin et al. 2004). The suppression of the response to the lagging sound is similar to behavioral results for ISDs eliciting localization dominance, where the single perceived location of a pair of clicks is dominated by the location of the leading click. This neural recovery is qualitatively similar to what occurs at ISDs past the behavioral echo threshold, where both the leading and lagging sounds are heard as separate images located near their respective sources.

The half-maximal ISD was about 20 ms for the model SMAX neuron and was about 17 ms for the model SMIN neuron. This result shows that when the lag came from the same direction as the lead, the lagging response of a model SMAX neuron (which shows greatest suppression of the lagging response when the leading sound comes from the neuron's best ITD) is suppressed for slightly longer than is the lagging response of a model SMIN neuron (which shows least suppression of the lag when the lead comes from the neuron's best ITD).

Varying the lead ITD as well as the ISD. Figure 5 compares physiological results recorded from single-unit IC cells of anesthetized cats (Litovsky and Yin 1998b; left column) to results from the model (right column). Although the CFs of the empirical IC cells (1–3 kHz) were higher than the CF of the model cell (500 Hz), Litovsky and Yin (1998b) suggested that ITDs still play an important role in binaural processing for the units from which they recorded. The top row shows responses for the leading click (Fig. 5A and B). The middle row compares lag responses for an SMAX neuron, which yields maximum suppression when the leading click is at the neuron's best location or ITD (Fig. 5C and D). Finally, the bottom row gives empirical and model responses for an SMIN neuron, where maximal

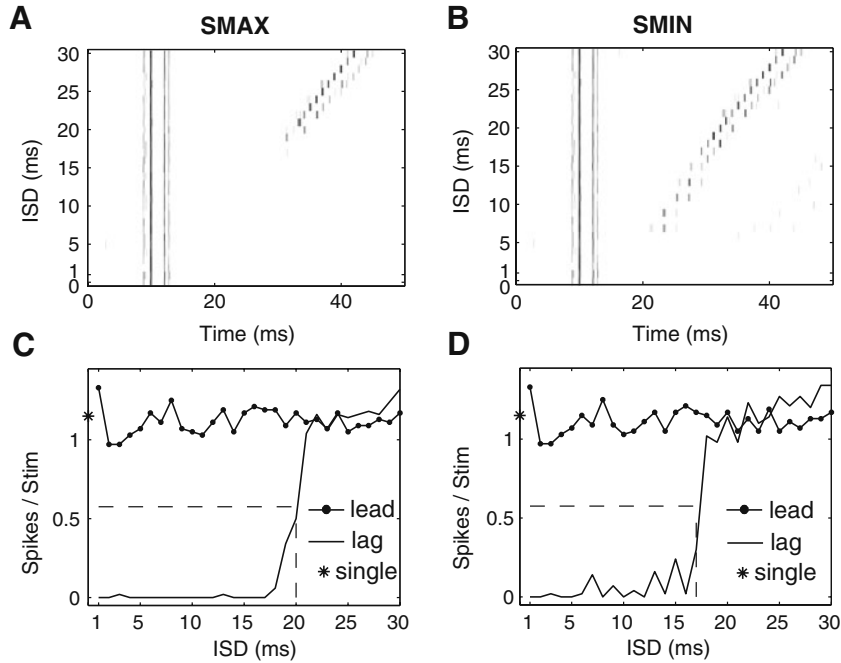


FIG. 4. Responses of the model SMAX (left column) and SMIN (right column) cell to a pair of binaural clicks, presented with an ISD varying from 1 to 30 ms in 1-ms increments. The ITD of both leading and lagging stimuli equals the cell's best ITD. **A, B** Dot raster responses at each ISD as a function of the time, with zero ISD indicating the responses to a single binaural click. **C, D** The spike count vs. the ISD for the lead, lag, and the single click. The dashed lines indicate the model cell's half-maximal ISD.

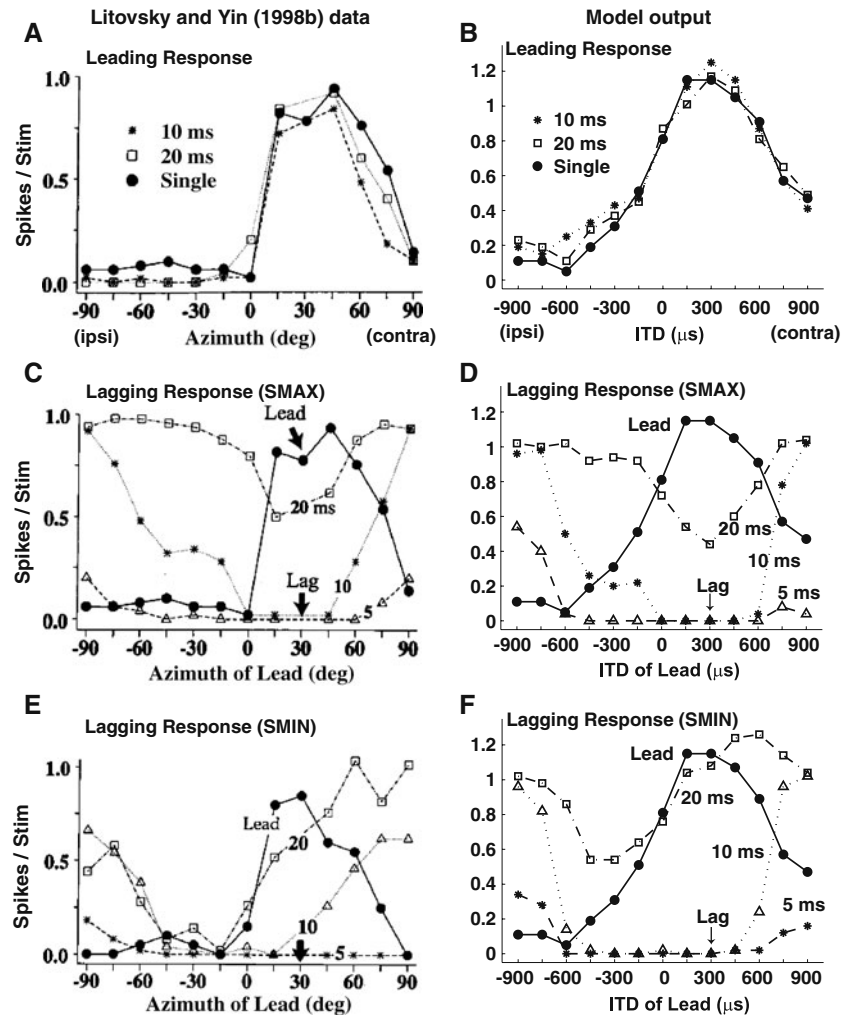


FIG. 5. Comparison of neuron response patterns to model response patterns. **A, C, E** Responses of neurons (from figures 1c, d, 2f of Litovsky and Yin (1998b), reprinted with permission). **B, D, F** Responses of model IC cells. *Top:* Responses to the leading stimulus with the lag delayed by 10 and 20 ms, as well as to a single click as a function of the location of the lead. *Middle:* Responses of SMAX neurons to the lagging stimulus as a function of lead location. *Bottom:* Responses of SMIN neurons. Responses to single clicks are shown for comparison. Arrow indicates the location of the lagging click, which evokes the maximal possible response from the neuron if presented in isolation. Responses to the lagging click are shown for ISDs of 5, 10, and 20 ms.

suppression occurs when the leading click is contralateral to the best location/ITD of the neuron (Fig. 5E and F, respectively).

For leading responses (Fig. 5A and B), the presence of the lagging click had a negligible effect for both the actual neuron and the model neuron. Specifically, responses to a single click (shown by the filled dots) are very similar to the responses to the lead click when the lag was presented 10 or 20 ms after the lead (asterisks and squares, respectively, in Fig. 5A and B). Both the actual neuron and the model neuron had broadly tuned responses, with peak responses at the best location (about 30° for the actual neuron in Fig. 5A) or ITD ($300 \mu\text{s}$ for the model neuron in Fig. 5B). The responses dropped off gradually as location or ITD deviated from these best values. The only noticeable difference between the shape of the actual and model neuron's responses is that the actual neuron's responses dropped off more rapidly for ipsilateral azimuths than for contralateral azimuths. The asymmetry is stronger for real neurons but is also present for model neurons. In considering this difference, it is worth noting that the physiological rate-ITD functions of other real neurons show less asymmetry than the neuron plotted in Figure 5A (from Litovsky and Yin 1998b), particularly if the stimulus level is higher than was used by Litovsky and Yin (e.g., see figure 4 of Carney and Yin 1989; figure 9 of Yin 1994).

The sensitivity of the lagging click response to the location of the leading click is summarized in panels C–F by plotting the responses to the lagging click at ISDs of 5, 10, and 20 ms as a function of the location of the leading click. In both physiological and model results, the lagging click was held constant at the best location or ITD (of the actual or model neuron, respectively). In the limit, at long enough ISDs, the effect of the lead will be negligible, and the lag response will be constant (equal to the cell's response to a single click at the neuron's best location or ITD) as a function of leading location. Any decrease in the lagging response below the single-click response at the best location or ITD reflects the suppressive effect of the leading stimulus (which the model assumes arises from delayed inhibition through DNLL). However, given that the lead response generally matches the single-click response (Fig. 5A and B), we can also compare the lag response to the response to the lead in order to estimate the effects of the lead on the lag.

As seen in Figure 5C–F, for both SMAX and SMIN neurons, the amount of suppression of the lagging response varied with the leading source location (or ITD; for a given connected line in a given panel, there are variations in the response with the abscissa value) as well as with the ISD (responses differ for the

different connected lines within a panel). The amount of lagging suppression decreased as the temporal separation between the lead and lag (i.e., the ISD) increased. For the shortest ISD (5 ms), the lagging response for both actual and model neurons was almost completely suppressed at nearly all locations of the leading click for both SMAX (plotted as triangles in Fig. 5C and D) and SMIN (plotted as asterisks in Fig. 5E and F) neurons. However, responses for SMAX and SMIN neurons differed at longer ISDs of 10 and 20 ms.

For both the actual and model SMAX neurons, the amount of lagging suppression at longer ISDs was maximal when the leading click was at the neuron's best location (Fig. 5C, near 30° azimuth) or best ITD (Fig. 5D, $300\text{-}\mu\text{s}$ ITD). At 20-ms ISD (plotted as squares), the suppression was weak when the leading click was located far from the best location (or ITD) of the neuron. At 10-ms ISD (plotted as asterisks), the suppression was greater overall and spread over a broader range of leading locations (or ITDs). However, the dependence on the leading location (or ITD) was still significant. The lagging response recovered almost fully from the suppression evoked by the lead at 10-ms ISD when the difference between the location of the lead and lag was large (i.e., the leading location/ITD was far to the ipsilateral side, far from the best location/ITD).

The lagging responses of the model SMIN neuron (Fig. 5F) were similar to those of the actual SMIN neuron (Fig. 5E): both gradually recovered with increased ISDs and both showed greatest suppression when the leading click was presented on the ipsilateral side (away from the best location/ITD of the neuron). One discrepancy between the empirical and model results is that the suppression was stronger for the actual neuron at 20-ms ISD than for the model neuron.

Discussion. At small ISDs, the lag response was greatly suppressed for all values of the lead ITD; the lagging responses of both SMAX and SMIN neurons were essentially eliminated for leads located in any position. In the model, this lack of dependence on the lead ITD at short ISDs arises because the lead causes almost all MSO neurons to fire at onset (i.e., the MSO onset response is poorly tuned in location). As a result, the onset response to a lead from any location will cause suppression in IC. Such poor ITD sensitivity of IC onset response to clicks is seen in some physiological data (e.g., Carney and Yin 1989, Fig. 5). However, there are no data on MSO responses to clicks in the cat of which we are aware; this prediction from our model is something that could be examined in future physiological studies. After the onset response, MSO responses become more spatially tuned. Therefore, the

later-arriving suppression is more spatially specific and causes suppression that depends on the relative locations of lead and lag. It is this later suppression that differentiates SMAX and SMIN neurons. Specifically, the model SMAX cell receives stronger inhibition from an ipsilateral MSO cell tuned to the same ITD as its excitatory projection; the model SMIN cell receives stronger inhibition from a contralateral MSO cell tuned to the ITD to which its excitatory projection responds minimally.

The strength of ipsilateral and contralateral inhibition in the model SMIN neuron are symmetrically opposite those of the model SMAX neuron (Table 1), which suggests that the amount of inhibition in an SMIN neuron when the lead is at +300 μs might be expected to equal the amount of inhibition in an SMAX neuron when the lead is at -300 μs . However, the model simulation for 10-ms ISD shows a further reduction in the model SMIN neuron's response when the lead and lag were both on the contralateral side (triangles in Fig. 5F, around +300 μs), compared with that of the model SMAX neuron when the lead was ipsilateral and lag was contralateral (asterisks in Fig. 5D, around -300 μs). This "further suppression" may simply be due to the extended refractory or adaptation-like effects in responses for contralaterally placed leading sources, which produce more activity overall, and which are therefore more likely to show such adaptation. As a result, later responses may be lower for contralateral leading sources than ipsilateral leads, which cause less adaptation. This kind of adaptation mechanism is included not only explicitly in the auditory nerve model but also implicitly in the membrane equations of the bushy cell, MSO cell, and IC cell model. We are not aware of any physiological data that directly addresses whether such adaptation mechanisms contribute to the suppression of lagging spatial information in the IC neuron responses; this is a question that could be tested in future studies.

Simulations of psychophysical data

To generate psychophysical predictions, a population of SMAX IC neurons was constructed. The population consisted only of SMAX units because SMAX units are thought to be more prevalent in the auditory pathway (Litovsky and Yin 1998b; Litovsky and Delgutte 2002). We used this population to simulate the results of behavioral experiments in which subjects were asked to indicate the perceived location(s) of the lead/lag target by adjusting the ITD of a pointer stimulus (Litovsky and Shinn-Cunningham 2001). For each lead-lag stimulus configuration, two sets of matches were made: one in which subjects matched the "right-most" image, and one in which they matched the "left-most" image. In the model simulations, the ITDs

of the lead took on values of -400, 0 or +400 μs ; the lagging ITD was held constant at +400 μs . To allow a direct comparison with the perceptual measures, simulated responses for a lagging stimulus on the left (ITD=-400 μs) were generated by assuming left/right symmetry of the model.

Responses of a population of neurons. Figure 6A displays the output of a population of SMAX neurons in response to pairs of binaural clicks as dot rasters showing the number of spikes as a function of time. The ordinate shows the best ITD of the neurons making up the population. Higher activity is indicated by darker gray scale. The top row shows the response to a single binaural click whereas the second, third, and the bottom rows show the response to a pair of lead-lag clicks with ISDs of 5, 10, and 20 ms, respectively. The ITD of the leading click equaled -400, 0, and +400 μs , in the left, center, and right columns, respectively. The ITD of the lagging click was +400 μs . Each panel in Fig. 6B, C shows the activity of every neuron (each point along the abscissa corresponds to a neuron with a different ITD). The leading response (Fig. 6B) is obtained by summing the spikes falling within the leading window (dotted box in Fig. 6A). At long ISDs, the lagging response (Fig. 6C) equals the number of spikes falling within the lagging window (dashed box in Fig. 6A). At short ISDs, the lagging spike count is corrected to account for overlap with the lead response (see [individual neuron analysis](#) above). The vector sum P_k obtained by Eq. 3 is plotted in each panel of Figure 6B (P_{lead}) and C (P_{lag}) as vertical bars: the ITD at which the bar is plotted is the estimated ITD of the leading and lagging clicks (calculated from the phase of P_k , see Eq. 4) and the height of the bar is the reliability of the estimated ITD (corresponding to the magnitude of P_k , see Eq. 5).

The response of the population of SMAX neurons was consistent with the results of a single SMAX model neuron. The population response to a single click located at the leading position (first row in Fig. 6B) was similar to the response to the leading click when the lag was 5, 10, or 20 ms (second to fourth rows in Fig. 6B). For all ISDs, the estimated ITD based on the leading response matched the ITD of the leading stimulus (-400, 0, and +400 μs in the left, center, and right columns, respectively).

Figure 6C shows the lagging response. For the 5-ms ISD, almost all of the response to the lagging click was eliminated, no matter whether the lead ITD was -400, 0 or +400 μs . In these cases, the estimated ITD differed from the true ITD of the lagging stimulus (the location of the vertical bar differed from the "correct" ITD, +400 μs) and the reliability of the estimates (the height of the vertical bar) was low. The suppression of the lagging

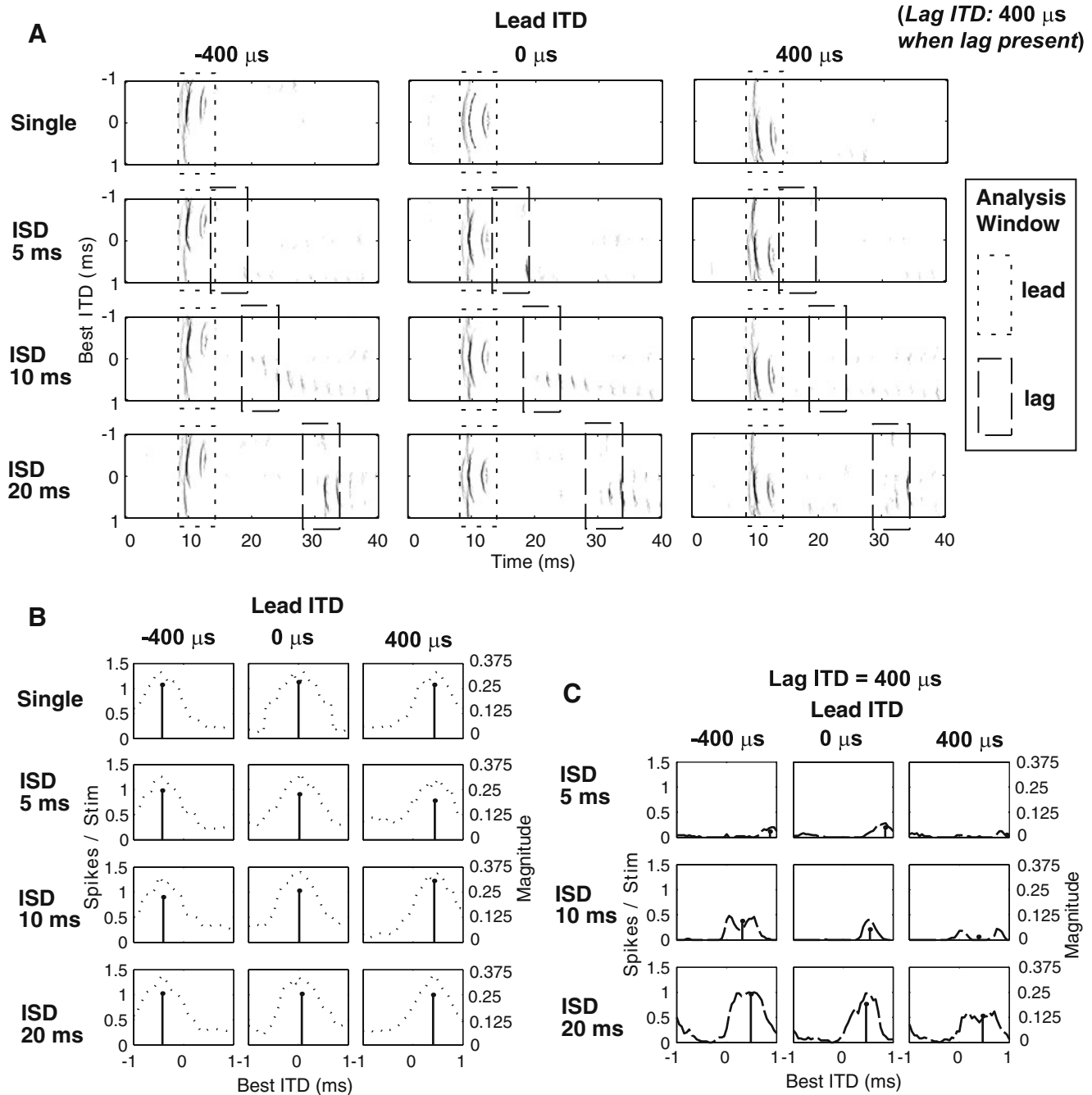


FIG. 6. Responses to a pair of binaural clicks from a population of model IC neurons whose best ITDs vary from -1 to $+1$ ms. In all panels, the left column shows results for lead on the left ($-400\text{-}\mu\text{s}$ ITD) and lag on the right ($400\text{-}\mu\text{s}$ ITD). The center column shows results for lead at center ($0\text{-}\mu\text{s}$ ITD) and lag on the right ($400\text{-}\mu\text{s}$ ITD). The right column shows results for both lead and lag on the right ($400\text{-}\mu\text{s}$ ITD). **A** The number of spikes as a function of time for each neuron in the population. The responses to a single click are shown in the top row, and the responses shown in the second, third, and bottom row are for a pair of clicks with ISDs of 5, 10 and 20 ms, respectively. Higher activity is indicated by a darker gray scale.

Within each panel, the leading window is indicated by the dotted box and the lagging window is indicated by the dashed box. **B** The number of spikes attributed to the lead as a function of each neuron's best ITD. The responses to a single click are shown in the top row, and the responses shown in the second, third, and bottom row are for a pair of clicks with ISDs of 5, 10 and 20 ms, respectively. **C** The number of spikes attributed to the lag as a function of each neuron's best ITD. The responses shown in the top, middle, and bottom row are for a pair of clicks with ISDs of 5, 10 and 20 ms, respectively. The magnitude of the vector average of responses is also plotted in B and C (vertical bars, scale on right).

response decreased with increasing ISDs. For the 10-ms ISD, some of the neurons had already recovered, but the percentage of recovered neurons depended on the lead ITD. More neurons recovered when the lead was at

$-400\text{-}\mu\text{s}$ ITD (far left) than when the lead was at $+400\text{-}\mu\text{s}$ ITD (far right).

For the 20-ms ISD, many of the neurons had fully recovered, although the lagging response was still

partially suppressed for some neurons. When the lead was at $-400\text{-}\mu\text{s}$ ITD, the magnitude of the vector sum was relatively large, reflecting the fact that the model predicted that the estimated ITD was relatively reliable. Moreover, in this case, the estimated lagging ITD was around $400\text{ }\mu\text{s}$, near the true ITD of the lagging stimulus. The vector sum had a smaller magnitude when the lead and lag were at the same location (Fig. 6C, ISD=20 ms, right panel) compared to when they were spatially far apart (left panel), even though the estimated lagging ITD was relatively accurate in both cases. For the model SMAX neurons, suppression was stronger when the lead and lag were at the same location, resulting in a reduction of the responses of neurons whose best ITD matched the leading and lagging ITD. As a result, the population activity was spread more evenly across a larger number of neurons with different best ITDs than when lead and lag were spatially separated. This kind of flat distribution of activity across neurons produced a less focused population response, resulting in a smaller-magnitude vector sum.

Estimates of perceived location. Litovsky and Shinn-Cunningham (2001) calculated the metric c to quantify the relative influence of the lead and lag in localization by comparing the perceptual location to the locations of the lead and lag. They assumed that the perceived lateral position of a PE stimulus is a weighted average of the leading and lagging ITDs. The value of c was estimated as the perceptual weight listeners gave to the lead location relative to that of the lag in producing their responses. This perceptual c value can be directly compared to the model weights in Eq. 6. The precedence weights c are plotted in Figure 7A and B for behavioral and model results, respectively. Model results are shown for conditions with the lead on the left and the lag on the right (left of Fig. 7B); the lead at center and the lag on the right (center of Fig. 7B); and both lead and lag on the right (right of Fig. 7B), a condition for which behavioral results could not be measured. In the legend, lead-lag positions are denoted by the two ordered letters (right, R, $+400\text{-}\mu\text{s}$ ITD; center, C, $0\text{-}\mu\text{s}$ ITD; and left, L, $-400\text{-}\mu\text{s}$ ITD). The bold letter indicates which of the two stimuli the listener was instructed to match. Therefore, c_1 (calculated when the instruction was to match the lead) is plotted for conditions in which the first letter is in bold and c_2 (calculated when the instruction was to match the lag) is plotted for conditions in which the second letter is in bold.

In general, behavioral and model results are in reasonable agreement. The weight for the lead-matching instruction c_1 (filled symbols) was always near one in both behavioral and model results. These results show that when instructed to match the

leading image, listeners heard the lead near the location at which the lead would be heard in isolation, with little influence of the lag. The influence of the lead on the localization of the lag is quantified by the weight for the lag-matching instruction, c_2 (open symbols), with values near zero indicating that listeners heard the lag near its own source location (weak precedence) and values near one indicating that the lag was perceived near the leading location (strong precedence). In both behavioral and model simulations, precedence was strong for short ISDs (<5 ms), with c_2 close to 1. As ISD increased, precedence weakened, and the value of c_2 decreased. This decrease was larger for the behavioral results than for the model results. In Figure 7A, c_2 dropped below 0.5 for ISDs larger than 10 ms whereas in the corresponding panels in Figure 7B, c_2 was still around 0.5 for long ISDs. For both behavioral and model results, at long ISDs, precedence was stronger when the lead and lag location were closer together (middle panels) than when the lead and lag were farther apart (left panels). In particular, c_2 was smaller for 10- and 15-ms ISDs when the lead and lag were from different hemifields (left panels) than when the lead was at center and lag was lateral (middle panels).

Figure 8A shows the ITD the listeners used to match the locations of the lead and lag in behavioral experiments (Litovsky and Shinn-Cunningham 2001). The corresponding model simulation results are shown in Figure 8B. The predicted responses when the lag was on the left (dashed lines in Fig. 8B) were based on the model's predictions when the lag was on the right (solid lines in Fig. 8B) by assuming left/right symmetry. In both behavioral and model results, the matched ITD was near the lead location for short ISDs regardless of whether listeners were instructed to match the lead or the lag, suggesting that the localization cues of the lead dominated. When listeners matched the lead image (filled symbols), both behavior and predictions were similar: the matched ITD was near the lead ITD for all ISDs from 1 to 15 ms, no matter whether the lead and lag were located on the left, right, or at the center. When instructions were to match the lag (open symbols), the matched ITD for both behavior and model predictions approached the lag ITD for ISDs longer than 10 ms. In the model, the dominance of the lead on the perceived location of the lag was determined by the precedence weight c_2 (Eq. 6), which depends on the recovery of the responses of the model IC population attributed to the lagging stimuli (Eq. 8).

Consistent with Figure 7, precedence was slightly stronger in the model predictions than in behavioral results at long ISDs. For the 10-ms ISD, the matched ITD was closer to the "true" lagging ITD in Figure 8A than in Figure 8B. Also consistent with the results shown in Figure 7, for both behavioral and model

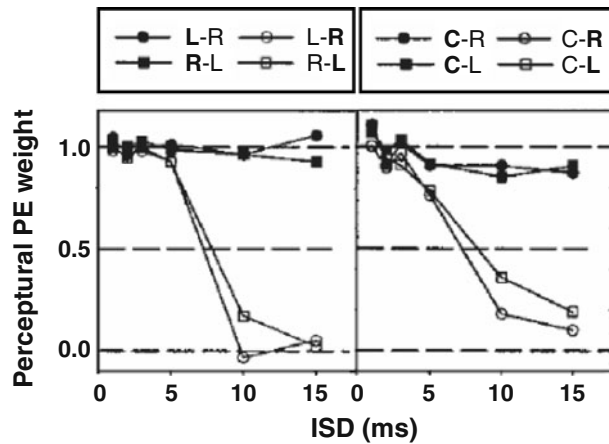
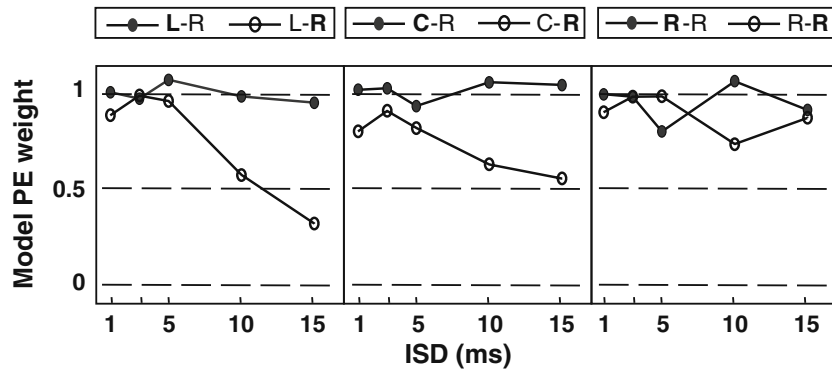
A Litovsky and Shinn-Cunningham (2001) data**B Model output**

FIG. 7. Precedence weight c as a function of inter-stimulus delay. **A** Estimates based on subject S1's response (from figure 8 of Litovsky and Shinn-Cunningham (2001), reprinted with permission). **B** Estimates based on the model cells' response. *Left* Lead and lag on opposite sides. *Center* Lead at center and lag lateral. *Right* lead and lag to the same side (note that in this condition, precedence weights could not be obtained from psychophysical measures). For each condition, the stimulus that listeners are instructed to match is indicated in bold. *Filled* and *open symbols* reflect conditions in which instructions are to match the lead and lag, respectively.

results, precedence was stronger when lead and lag were spatially near to one another than when they were from opposite hemifields. In Figure 8A and B, for long ISDs, the matched ITD was closer to the "true" lagging ITD in L-R and R-L conditions (left panels) than in C-R and C-L conditions (middle panels). These results indicate that when the lead and lag were spatially near to one another, the likelihood of perceiving two distinct images at their correct locations was lower than when they were far apart.

Discussion. In precedence-effect conditions, the dominance of the lead over the lag depends on the ISD as well as the relative location of the lead and lag. For short ISDs, the lagging responses of almost all the neurons in the model population were greatly suppressed and the model predicted that listeners heard one image near the location of the leading source. The lagging response recovered for long ISDs and the model predicted that listeners heard a second image near the location of the lagging source (i.e., where the lag would be heard in isolation). Moreover, due to the fact that model SMAX cells generate the strongest suppression of the lag when the lead and lag locations are close together, the model predicted that

precedence was stronger when the lead and lag were relatively near one another in space than when they were from opposite hemifields. In contrast, if the neural population consisted only of SMIN units that generate the strongest suppression when the lead and lag are from opposite hemifield, the predicted localization dominance would be stronger when two stimuli were further apart in space, which is inconsistent with psychophysical results (Litovsky and Shinn-Cunningham 2001; Dent et al. 2009).

For the 5-ms ISD, the lack of complete suppression of all the model neurons is compatible with behavioral results showing that the lagging stimulus can be detected at short ISDs even when it is not localized at the true lag location (Blauert 1983; Freyman et al. 1998). For the 10-ms ISD, some of the model neurons began to respond to the lag; the model predicted that listeners would perceive both a source near the lead and a second source somewhere between the lead and lag locations. In this case, a flat distribution of activity across neurons responding to the lag resulted in an unreliable estimate of the IPD of the lagging stimulus (Eq. 5). As a result, the model gives little weight to the unreliable estimate of the lag location (Eqs. 6 and 8), causing strong lead dominance. These results suggest

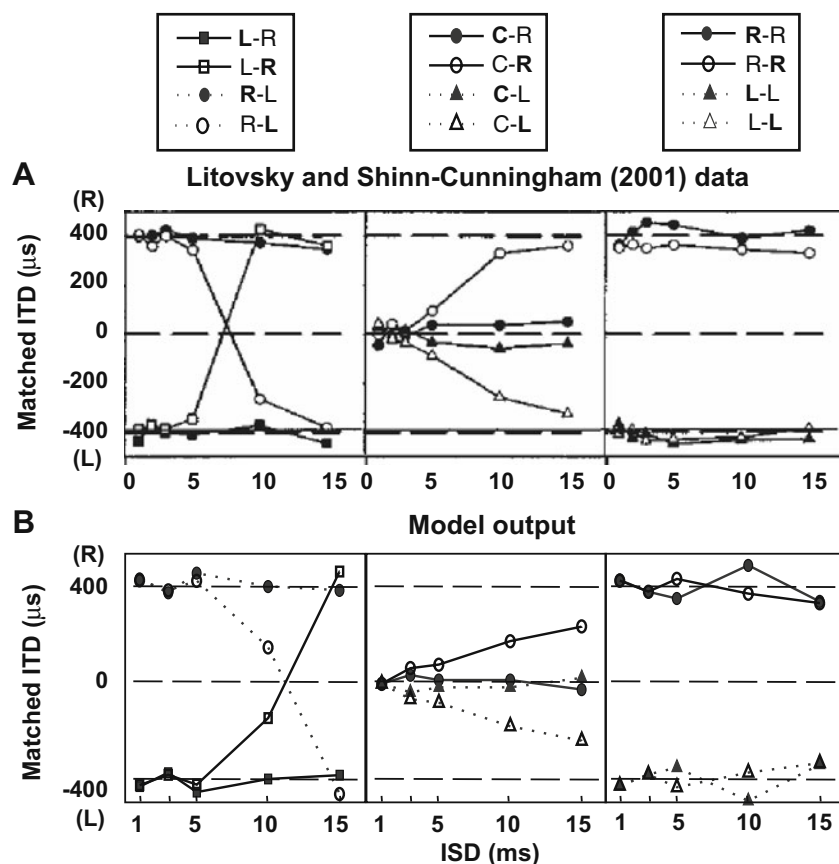


FIG. 8. The matched ITD under various conditions, for ISDs from 1 to 15 ms. In the legend, lead-lag positions are denoted by order. The **bold letter** indicates whether instructions are consistent with matching the lead or lag. *Left column* lead and lag on opposite sides. *Center column* lead at center and lag on either right or left. *Right column* lead and lag to the same side. **A** Matching results for subject S1 (from figure 5 of Litovsky and Shinn-Cunningham (2001), reprinted with permission). **B** Estimates of the model.

that the lagging image is heard at its own source location only when the lagging response reliably encodes the lag location, i.e., at long ISDs. For the 20-ms ISD, although the lagging responses of some model cells were still partially suppressed, the model predicted that both the lead and lag were heard at the locations from which they would be heard in isolation. These results suggest that full recovery of the lagging response of all neurons requires an ISD longer than the ISD at which listeners first perceive two sources (echo threshold).

Although the current model assumes a uniform distribution of best ITDs in the neural population, physiological data show that the distribution of best ITDs is highly dependent on CF and in general does not correspond to the range of naturally occurring ITDs (McAlpine et al. 2001; Hancock and Delgutte 2004). Instead, best IPD is more independent of CF and the steepest slopes of neural rate-ITD functions tend to occur near the midline. A distribution that was more “physiological” could be modeled by imposing a non-uniform weighting of the neural responses that we simulated. Such weighted responses would have peaks at symmetrically positioned positive and negative ITDs, corresponding to the two populations of neurons located in the left and right hemispheres, respectively. The perceived ITD could then be calcu-

lated as a difference between ipsilateral and contralateral responses (e.g., as in Hancock 2007), rather than based on the vector average over the uniform distribution. For the lead response, the difference between the ipsilateral and contralateral responses must map, perceptually, to -400 , 0 , and $+400$ μs in the three columns, respectively (Fig. 6B). For the lag response (Fig. 6C), such a difference would (1) contain no information about the lag at the shortest ISD, (2) become larger as the ISD increases, and (3) be smaller for lead and lag both at $+400$ μs than for lead at -400 μs and lag at $+400$ μs (bottom right panel compared to bottom left panel of Fig. 6C). All of these predictions, therefore, are qualitatively similar to predictions of the current model.

GENERAL DISCUSSION

Physiological evidences of inhibition

Similar to previous models of PE (e.g., Lindemann 1986; Zurek 1987; Dizon and Colburn 2006), the current model suggests that localization dominance arises because the response to the lagging source is suppressed while the response to the leading source is preserved. This kind of physiological suppression has been observed at several levels of the ascending

auditory system, including the auditory nerve (Parham et al. 1996), the cochlear nucleus (Parham et al. 1998), the superior olivary complex (Fitzpatrick et al. 1995), the inferior colliculus (Fitzpatrick et al. 1995; Litovsky and Yin 1998a, b; Tollin et al. 2004), and the auditory cortex (Fitzpatrick et al. 1999). In the current model, the critical suppression, which depends on the lead location, arises from inhibition from MSO via the DNLL to IC. The suppression in the auditory nerve and cochlear nucleus may be able to explain some of the suppression observed in the IC (Hartung and Trahiotis 2001; Trahiotis and Hartung 2002). However, we argue that this kind of suppression is too brief and too weak to explain all the suppression seen in the IC (see below). Also, most known inhibitory inputs to the MSO are monaural and are insensitive to ITDs (Fitzpatrick et al. 1995), so they could not generate interactions between lead and lag locations like those observed physiologically. For SMIN neurons, a leading stimulus that by itself elicits few or no spikes actually suppresses the lag more effectively than does a leading stimulus at the best ITD. The existence of SMIN neurons rules out the possibility that a long refractory period or recurrent inhibition among IC neurons is the only cause of long-lasting suppression of the lag at the IC level.

The current model assumes that the long suppression observed in the IC is due to synaptic inhibition coming from the DNLL on both sides (Adams and Mugnaini 1984; Shneiderman et al. 1988) and that the DNLL receives ipsilateral excitatory projections from the MSO (Oliver et al. 1987). Thus, for an IC neuron, the best ITD and the worst ITD preferentially activate the ipsilateral and contralateral DNLLs, respectively. Since the IC neuron receives inhibitory inputs from both DNLLs, leading clicks from both the best and worst ITDs evoke some suppression in IC. The balance of these two ITD-tuned inhibitions varies from neuron to neuron, giving rise to SMAX (stronger inhibition from ipsilateral DNLL) and SMIN (stronger inhibition from contralateral DNLL) model neurons, consistent with the observed physiology. Anatomically, the ascending, inhibitory projections to the IC primarily come from the DNLL on both sides and the low-frequency region of ipsilateral lateral superior olive (LSO) (Saint Marie et al. 1989; Loftus et al. 2004).

Though not included in the present model, the inhibitory inputs from the ipsilateral LSO are likely to produce responses in IC that are consistent with the inhibition present in the model that is driven by the contralateral MSO (via the corresponding DNLL). LSO neurons are driven by ipsilateral excitation and contralateral inhibition, resulting in trough-type ITD sensitivity that is phase-inverted with respect to the peak-type ITD sensitivity of MSO neurons at the same CF and characteristic delay (Fitzpatrick et al. 2002).

This phase inversion in ITD sensitivity means that the inhibition from the ipsilateral LSO would suppress at the worst ITD of the ipsilateral MSO; thus, such responses could contribute to SMIN responses at the IC. The current model could be extended to include realistic inputs from LSO. By adjusting the relative strengths of these LSO inputs and the strengths of the current MSO-driven excitation and inhibition, the extended model should be able to generate predictions very much like those presented here.

The influence of peripheral processing

Hartung and Trahiotis (2001) suggest that peripheral interference at the level of the auditory nerve between directional information in the lead and the lag can explain the PE in some conditions. Such peripheral interference is greatest in low-frequency neurons due to the band-pass filtering and adaptation mechanisms in the cochlea. When the ISD is comparable to the duration of the click response at the AN level, the response to the lead causes ringing in the basilar membrane that causes significant lead-lag interactions. Any residual lead response can add constructively or destructively with the responses to the lag, depending on the relative monaural phases of lead and lag, which depends on the ISD, the ITD, and the CF of the auditory nerve in question. When lead and lag have different ITDs, the relative monaural phases of the lead and lag can differ in the left and right ears, resulting in shifts in the effective lag ITD due to the different monaural interactions. The effects of such peripheral interaction were shown in the model MSO's response for ISDs shorter than 5 ms (Fig. 2A), altering the outputs of the model AN fiber that drive model MSO responses.

If peripheral interference were the only factor contributing to the PE, precedence would only occur for ISDs shorter than 5 ms. Moreover, without additional suppression, a lagging source following shortly after a lead should evoke some responses reflecting the internal, effective ITDs caused by peripheral interactions, which suggest a localizable event that is not heard at either the leading or the lagging location. Neither of these predictions is consistent with past results. Physiological data (and the current model simulations) shows a more general suppression of the lagging response than can be explained by short-lasting interactions in the cochlea. Specifically, in the IC, suppression lasts for as long as 20 ms. When the ISD is short, the lagging response is diminished no matter where the lead is located. Consistent with this, listeners do not hear the location of the lagging source when the ISD is short.

The similarity between cats and humans

In the current study, we simulated physiological data from cats and psychophysical data from humans, even though there are likely differences between both the representations of ITD and the distributions of best-ITDs in the two species (e.g., see Harper and McAlpine 2004). Although we did not specifically model any perceptual data from cats, results from behavioral experiments in cats are generally similar to those found in humans (Cranford 1982; Populin and Yin 1998). For example, Tollin and Yin (2003) measured the PE for horizontally positioned sources in cats using direct localization procedures. During the time course of localization dominance for humans, cats localized stimuli near the leading source location. In the range of echo threshold for humans, cats were able to perceive the lead and the lag at distinct locations, and at the longest ISDs, the perceived lead and lag locations were like those that the lead and lag would produce in isolation.

The similarity of results suggests that any underlying neural mechanisms of PE measured physiologically in cats may be similar to those in humans. The current model successfully simulates the recovery time of IC neurons measured in anesthetized cats. However, our model, in which the parameter values were chosen to fit the cat's physiological data, predicts a time course of localization dominance longer than that measured in the human psychophysical experiments (see Figs. 7 and 8). This discrepancy may be due to either an effect of anesthesia or to species differences, rather than a quantitative failure of the model. Previous studies have shown that the mean neural recovery time for PE stimuli is about 35 ms in anesthetized cats (Litovsky and Yin 1998a, b; Yin 1994), but only about 7 ms in unanesthetized rabbits (Fitzpatrick et al. 1995) and awake behaving cats (Tollin et al. 2004). Thus, we suspect that the discrepancy in recovery time between our model results and behavioral results is due to anesthesia rather than species differences or a failure of the model.

General notions regarding the PE

The precedence effect is one of the few well-studied auditory phenomena in both physiology and psychophysics. Physiologically, different IC neurons show considerable differences in their responses to PE stimuli. Some neurons showed a period of reduced suppression for short ISDs (Yin 1994), and some neurons showed suppression that was independent of the leading location (Fitzpatrick et al. 1995; Litovsky and Yin 1998b). These results are consistent with the fact that IC is an obligatory station for all

ascending projections from the lower auditory brain stem, including multiple inhibitory pathways. Each of these projections could contribute to behavior associated with the PE. However, our predictions are currently based only on a population of SMAX neurons, which are thought to be more numerous than other neurons. Psychophysically, echo thresholds vary widely with stimulus characteristics (Blauert 1997), and the PE may “build up” with repeated stimulus presentations in human observers (Freyman et al. 1991). Our model cannot account for these effects, which may be due to the feedback from the auditory cortex. Future extensions to the current model could include such factors by adding ascending pathways from different types of neurons as well as descending pathways from higher levels of the auditory system.

CONCLUSIONS

A model IC was developed that simulates physiological responses and predicts psychophysical behavior in response to precedence-effect click stimuli. The single IC neuron model was based on the Cai et al. (1998) model, which incorporates existing models for auditory-nerve fibers (Carney 1993), bushy cells in the cochlear nucleus (Rothman et al. 1993), and principal cells of the MSO (Brughera et al. 1996). The IC model cell received excitatory inputs from an ipsilateral MSO model cell, as well as inhibitory inputs from both ipsilateral and contralateral MSO model cells via the DNLL. Most of the suppression of the lagging response in the model IC was due to the long-lasting inhibition from MSO evoked by the leading stimulus. This suppression was modulated by ITD because the inhibition came from cells that were themselves sensitive to stimulus ITD. Consistent with previous data (Yin 1994; Fitzpatrick et al. 1995; Tollin et al. 2004), the model neuron cells showed suppression of the lagging response at short ISDs, with greatest suppression at ISDs from 1 to 5 ms. By adjusting the relative strength of inhibition from both sides, some model neurons displayed strongest suppression of the lagging response for a lead at the neuron's best ITD, whereas others had the strongest suppression for a lead placed in the hemifield opposite the best ITD, just as has been observed in IC (Litovsky and Yin 1998a, b). A population model of IC readout of the responses of a population of the first type of model neurons explained localization dominance reported in psychophysical studies of PE, whereby at short ISDs, the perceived location of a pair of clicks is dominated by the leading source; the strength of dominance decreases and the lagging sound is more likely to be heard near its own true location as

the spatiotemporal separation of the lead and lag increases (Litovsky and Shinn-Cunningham 2001).

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (DC009477 to BGSC and DC00100 to HSC).

REFERENCES

- ADAMS JC, MUGNAINI E (1984) Dorsal nucleus of lateral lemniscus: a nucleus of GABAergic projection neurons. *Brain Res Bull* 13:585–590
- BLAUERT J (1983) *Spatial hearing*. MIT, Cambridge
- BLAUERT J (1997) *Spatial hearing: the psychophysics of human sound localization*. MIT, Cambridge
- BRAND A, BEHREND O, MARQUARDT T, McALPINE D, GROTHE B (2002) Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417:543–547
- BRUGHERA AR, STUTMAN ER, CARNEY LH, COLBURN HS (1996) A model with excitation and inhibition for cells in the medial superior olive. *Aud Neurosci* 2:219–233
- CAI H, CARNEY LH, COLBURN HS (1998) A model for binaural response properties of inferior colliculus neurons. I. A model with ITD-sensitive excitatory and inhibitory inputs. *J Acoust Soc Am* 103:475–493
- CARNEY LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. *J Acoust Soc Am* 93:401–417
- CARNEY LH, YIN TCT (1989) Responses of low-frequency cells in the inferior colliculus to interaural time differences of clicks: excitatory and inhibitory components. *J Neurophysiol* 62:144–161
- COLBURN HS, ISABELLE SK (1992) Models of binaural processing based on neural patterns in the medial superior olive. In: Cazals Y, Demaney L, Horner K (eds) *Auditory physiology and perception*. Pergamon, Oxford, pp 539–545
- COLBURN HS, HAN Y, CULOTTA CP (1990) Coincidence model of MSO responses. *Hear Res* 49:335–346
- CRANFORD JL (1982) Localization of paired sound sources in cats: effects of variable arrival times. *J Acoust Soc Am* 72:1309–1311
- DENT ML, TOLLIN DJ, YIN TC (2009) Influence of sound source location on the behavior and physiology of the precedence effect in cats. *J Neurophysiol* 102:724–734
- DEVORE S, IHLEFELD A, HANCOCK K, SHINN-CUNNINGHAM BG, DELGUTTE B (2009) Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron* 62:123–134
- DIZON RM, COLBURN HS (2006) The influence of spectral, temporal, and interaural stimulus variations on the precedence effect. *J Acoust Soc Am* 119:2947–2964
- FITZPATRICK DC, KUWADA S, BATRA R, TRAHOTIS C (1995) Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol* 74:2469–2486
- FITZPATRICK DC, KUWADA S, KIM DO, PARHAM K, BATRA R (1999) Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. *J Acoust Soc Am* 106:3460–3472
- FITZPATRICK DC, KUWADA S, BATRA R (2002) Transformations in processing interaural time differences between the superior olivary complex and inferior colliculus: beyond the Jeffress model. *Hear Res* 168:79–89
- FREYMAN RL, CLIFTON RK, LITOVSKY RY (1991) Dynamic processes in the precedence effect. *J Acoust Soc Am* 90:874–884
- FREYMAN RL, ZUREK PM, BALAKRISHNAN Y, CHIANG YC (1997) Onset dominance in lateralization. *J Acoust Soc Am* 101:1649–1659
- FREYMAN RL, MCCALL DD, CLIFTON RK (1998) Intensity discrimination for precedence effect stimuli. *J Acoust Soc Am* 103:2031–2041
- HANCOCK KE (2007) A physiologically-based rate code for interaural time differences predicts bandwidth-dependent lateralization. In: Kollmeier B, Klump G, Hohmann V, Langemann U, Mauermann M, Uppenkamp S, Verhey J (eds) *Hearing—from basic research to applications*. Springer, New York, pp 389–398
- HANCOCK KE, DELGUTTE B (2004) A physiologically based model of interaural time difference discrimination. *J Neurosci* 24:7110–7117
- HARPER NS, McALPINE D (2004) Optimal neural population coding of an auditory spatial cue. *Nature* 430:682–686
- HARTUNG K, TRAHOTIS C (2001) Peripheral auditory processing and investigations of the “precedence effect” which utilizes successive transient stimuli. *J Acoust Soc Am* 110:1505–1513
- HENKEL CK, SPANGLER KM (1983) Organization of the efferent projections of the MSO nucleus in the cat as revealed by HRP and autoradiographic tracing methods. *J Comp Neurol* 221:416–428
- HODGKIN AL, HUXLEY AF (1952) A quantitative description of membrane current and its application to conduction and excitation in the nerve. *J Physiol (London)* 117:500–544
- JORIS PX, CARNEY LH, SMITH PH, YIN TCT (1994) Enhancement of synchronization in anteroventral cochlear nucleus. I. Responses to tones at characteristic frequency. *J Neurophysiol* 71:1022–1036
- JORIS PX, SMITH PH, YIN TCT (1998) Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* 21:1235–1238
- KELLER CH, TAKAHASHI TT (1996) Responses to simulated echoes by neurons in the barn owl’s auditory space map. *J Comp Physiol* 178:499–512
- KELLY JB (1974) Localization of paired sound sources in the rat: small time differences. *J Acoust Soc Am* 55:1277–1284
- KUWADA S, YIN TCT (1983) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol* 50:981–999
- LINDEMANN W (1986) Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front. *J Acoust Soc Am* 80:1623–1630
- LITOVSKY RY, DELGUTTE B (2002) Neural correlates of the precedence effect in the inferior colliculus: effect of localization cues. *J Neurophysiol* 87:976–994
- LITOVSKY RY, SHINN-CUNNINGHAM BG (2001) Investigation of the relationship among three common measures of precedence: fusion, localization dominance, and discrimination suppression. *J Acoust Soc Am* 109:346–358
- LITOVSKY RY, YIN TCT (1998a) Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. *J Neurophysiol* 80:1285–1301
- LITOVSKY RY, YIN TCT (1998b) Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms. *J Neurophysiol* 80:1302–1316
- LOFTUS WC, BISHOP DC, SAINT MARIE RL, OLIVER DL (2004) Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive. *J Comp Neurol* 472:330–344
- McALPINE D, JIANG D, PALMER AR (2001) A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4:396–401
- OLIVER D, SHNEIDERMAN A, HENKEL CK (1987) Morphological substrates for binaural interactions in the midbrain: the dorsal nucleus of the lateral lemniscus (Abstract). *Assoc Res Otolaryngol* 10:218

- PARHAM K, ZHAO HB, KIM DO (1996) Responses of auditory nerve fibers of the unanesthetized decerebrate cat to click pairs as simulated echoes. *J Neurophysiol* 76:17–29
- PARHAM K, ZHAO HB, YE Y, KIM DO (1998) Responses of anteroventral cochlear nucleus neurons of the unanesthetized decerebrate cat to click pairs as simulated echoes. *Hear Res* 125:131–146
- POPULIN LC, YIN TCT (1998) Behavioral studies of sound localization in the cat. *J Neurosci* 18:2147–2160
- ROTHMAN JS, MANIS PB (2003) The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J Neurophysiol* 89:3097–3113
- ROTHMAN SR, YOUNG ED, MANIS PB (1993) Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J Neurophysiol* 70:2562–2583
- SAINT MARIE RL, OSTAPOFF EM, MOREST DK (1989) Glycine-immunoreactive projection of the cat lateral superior olive: possible role in midbrain ear dominance. *J Comp Neurol* 279:382–396
- SHINN-CUNNINGHAM BG, KAWAKYU K (2003) Neural representation of source direction in reverberant space. In: *Proceedings of the 2003 IEEE workshop on applications of signal processing to audio and acoustics*, New Paltz, NY, pp 79–82
- SHNEIDERMAN A, OLIVER DL, HENKEL CK (1988) Connections of the dorsal nucleus of the lateral lemniscus: an inhibitory parallel pathway in the ascending auditory system. *J Comp Neurol* 276:188–208
- TOLLIN DJ, HENNING GB (1999) Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum. *J Acoust Soc Am* 105:838–849
- TOLLIN DJ, YIN TCT (2003) Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. *J Neurophysiol* 90:2149–2162
- TOLLIN DJ, POPULIN LC, YIN TCT (2004) Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J Neurophysiol* 92:3286–3297
- TRAHIOTIS C, HARTUNG K (2002) Peripheral auditory processing, the precedence effect and responses of single units in the inferior colliculus. *Hear Res* 168:55–59
- WALLACH H, NEWMAN EB, ROSENZWEIG MR (1949) The precedence effect in sound localization. *Am J Psychol* 52:315–336
- WYTTENBACH RA, HOY RR (1993) Demonstration of the precedence effect in an insect. *J Acoust Soc Am* 94:777–784
- YIN TCT (1994) Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J Neurosci* 14:5170–5186
- ZUREK PM (1980) The precedence effect and its possible role in the avoidance of interaural ambiguities. *J Acoust Soc Am* 67:953–964
- ZUREK PM (1987) The precedence effect. In: *Yost WA, Gourevitch G (eds) Directional hearing*. Springer, New York, pp 85–105