

Modeling Physiological and Psychophysical Responses to Precedence Effect Stimuli

Jing Xia¹, Andrew Brughera², H. Steven Colburn², and Barbara Shinn-Cunningham^{1,2}

¹ Department of Cognitive and Neural Systems, Boston University, Boston

{jxia@bu.edu, shinn@cns.bu.edu}

² Department of Biomedical Engineering, Boston University, Boston

{brughera@bu.edu, colburn@bu.edu}

Corresponding author:

Barbara Shinn-Cunningham, Ph.D.

Department of Cognitive and Neural Systems,

Boston University, 677 Beacon Street,

Boston, Massachusetts, 02215

Tele: (617) 353-5764

Fax: (617) 353-7755

Email: shinn@cns.bu.edu

ABSTRACT

Many perceptual and physiological studies of sound localization have explored the precedence effect (PE), whereby a pair of brief sounds coming from different directions and arriving at the ears close together in time is perceived as coming from a location near that of the first-arriving sound. Here, we used a computational model of low-frequency inferior colliculus (IC) neurons to account for both physiological and psychophysical responses to PE stimuli. In the model, physiological suppression of the ITD-tuned lagging response depends on the inter-stimulus delay (ISD) between the leading and lagging sound as well as the ITD of the lead. Psychophysical predictions generated from a population of model IC neurons estimate the perceived location of the lagging click as near that of the lead click at short ISDs, consistent with subjects perceiving both lead and lag as coming from the lead location. As ISD increases, the estimated location of the lag becomes closer to the true lag location, consistent with listeners perceiving two sounds coming from separate locations. Together, these physiological and perceptual simulations suggest that ITD-dependent suppression in IC neurons can explain the behavioral phenomenon known as the precedence effect.

INTRODUCTION

Listeners are remarkably accurate at localizing sound despite the reflections that are present in most settings. It is thought that the auditory system accomplishes this feat by attributing greater perceptual weight to the location cues at sound onsets and suppressing cues from later-arriving reflections. Wallach et al. (1949) introduced the term “precedence effect” (PE) to describe the phenomenon whereby a pair of dichotic clicks

presented with a brief inter-stimulus delay (ISD) are typically heard as a single “fused” sound image whose perceived direction is near the location of the leading stimuli.

Psychophysically, the PE can be described by three phases. *Summing localization*, in which listeners perceive one single fused auditory image located about halfway between the lead and lag sources, occurs for ISDs from 0 to about 1 ms. *Localization dominance*, in which the paired sounds are localized near the lead, occurs for ISDs ranging from about 1 to 10 ms. Finally, for ISDs more than about 10 ms, two separate auditory images can be heard, one near the location of the lead and one near the lag location.

Neural correlates of the PE have also been observed in physiological studies of extracellular responses, especially in inferior colliculus (IC) (Fitzpatrick et al. 1995; Tollin et al. 2004). For small ISDs, neural responses to the lag are reduced or eliminated, while at longer ISDs, responses to the lagging source recover, mirroring psychophysical results (Litovsky and Yin 1998a, b).

The objective of this study is to build a computational model that simulates the three phases of the PE as ISD changes, as well as the way these phenomena depend on the relative directions of the leading and lagging stimuli. In contrast to previous, generic models that simulate the behavioral aspects of the PE (Lindemann 1986; Hartung and Trahiotis 2001), the modeling approach in this study emphasizes the physiological basis of PE found in the binaural responses of IC neurons.

METHODS

Stimuli

PE stimuli were generated by presenting a pair of binaural clicks (the lead and the lag) separated by an ISD (defined as the time difference between the onsets of the lead and lag

delivered to the right ear). ITDs were imposed separately on the lead and lag. We assumed that the left and right ICs were mirror symmetric and generated models of individual left IC neurons. By convention, positive ITDs were generated by advancing the stimulus in the ear contralateral to the model cell (right) and delaying the stimulus in the ear ipsilateral to the model cell (left).

Model Structure

The model (Fig. 1) consists of a hierarchy of processing stages, mimicking the stages of the auditory periphery. Our IC model neurons, which are based on the previous IC model (Cai et al. 1998), are innervated by medial superior olive (MSO) model neurons (Brughera et al. 1996), with inhibitory inputs passed through the DNLL. The MSO model neurons receive bilateral inputs from models of bushy cells in the cochlear nucleus (Rothman et al. 1993), which receive convergent inputs from existing models of auditory-nerve fibers (Carney 1993).

Connections from model MSO cells to a model IC cell are based on anatomical and physiological evidence. Excitatory inputs from MSO to ipsilateral IC lead to ITD sensitivity in low-frequency cells of the IC (Kuwada and Yin 1983; Carney and Yin 1989; Loftus et al. 2004). Different model MSO neurons are characterized by different “best” ITDs, the ITDs that lead to the maximal firing rates (e.g., the ITD that best compensates for, among other things, any difference in the neural transmission delays to the neuron from the ipsi- and contralateral ears). With our conventions, model cells with positive best ITD are in the left IC.

Our IC model is based on the hypothesis that directionally tuned parallel inhibition from ipsilateral and contralateral MSOs (via the corresponding dorsal nuclei of the lateral

lemniscus, DNLL) produces the suppression of the lagging response (Fitzpatrick et al. 1995). The differential equations describing the membrane potential of the model IC neurons are the same as those used by Cai et al. (1998). Inhibitory inputs to the IC are delayed, presumably because of the extra synapse at the DNLL. As a result, inhibition suppresses responses after sound onsets, but has little effect on onset responses. Some model IC neurons receive more inhibition from the ipsilateral DNLL than from the contralateral DNLL (Fig. 1A). These neurons, known as suppression-at-maximum (SMAX) neurons (Litovsky and Yin 1998b), should show greatest suppression of the lagging response when the leading sound comes from the neuron's best ITD. In contrast, suppression-at-minimum (SMIN) neurons (Litovsky and Yin 1998b) receive stronger inhibition from the contralateral DNLL (Fig. 1B), and should show greatest suppression of the lag when the lead comes from positions that elicit minimal response.

Data Analysis

The techniques used here to compute the discharge rate in response to the leading and lagging stimuli were the same as those used in physiological studies (Fitzpatrick et al. 1995; Litovsky and Yin 1998a; Tollin et al. 2004). Specifically, to quantify the response to a single binaural click, the number of the spikes in an *analysis window* was counted. The window began when the number of discharges first exceeded two standard deviations of the mean spontaneous rate computed 10 ms prior to stimulus presentation and ended when the response fell below this level. Response latency was taken as the time of the onset of the analysis window. For a pair of binaural clicks, two windows (the *leading window* and the *lagging window*) were used to calculate responses. The start of the leading window was taken to be the latency of the response. The end of the leading

window was the end the analysis window in response to the single binaural click. The lagging window started at the time of the response latency plus the ISD and had a duration equal to the duration of the leading window. For small ISDs, the two windows overlapped with each other; the response to the lag was estimated by subtracting the number of spikes in response to a single leading stimulus from the number of spikes counted from the onset of the leading window to the offset of the lagging window.

The Readout

The absolute discharge rate of a single IC neuron cannot account for the perceived location of acoustic inputs. Readout of the responses of a population of IC neurons with the same frequency tuning was accomplished by forming an estimate of the perceived ITD from the population response. In the current work, we concentrated on the low characteristic frequency $f = 500$ Hz and assumed that the population consisted of neurons whose best ITDs were equally distributed over a symmetrical range from $-T$ to T ms, where T is some integer multiple of half the period of the characteristic frequency. For this neural distribution, the maximum likelihood estimate (MLE) of the input IPD (φ) can be computed as the complex angle of a weighted sum of complex values (Colburn and Isabelle 1992). The magnitude of each complex value representing one neuron's response is given by the firing rate (L) integrated across the time window corresponding to the leading and lagging stimulus (see data analysis). The phase depends on the best ITD (τ_m) of the neuron.

$$\hat{\varphi}_i(f) = \angle \sum_{\tau_m=-T}^{\tau_m=T} L_{i,\tau_m}(f) e^{j2\pi f \tau_m}, \quad (1)$$

where $i = 1, 2$, corresponds the response to the leading and lagging stimulus in the paired source, respectively. While the phase of the complex-valued sum estimates the source

IPD, the normalized magnitude of the sum (r) estimates the reliability of the observed population response estimate (Shinn-Cunningham and Kawakyu 2003).

$$\hat{r}_i(f) = \left| \sum_{\tau_m=-T}^{\tau_m=T} L_{i,\tau_m}(f) e^{j2\pi f \tau_m} \right| \quad (2)$$

For the PE stimuli used here, we assumed that the perceived ITD (α) to the lead ($i=1$) and to the lag ($i=2$) is a weighted sum of the lead and lag IPDs estimated by Eqn. 1.

$$\hat{\alpha}_i = \frac{c_i \hat{\phi}_1 + (1 - c_i) \hat{\phi}_2}{f}, \quad (3)$$

where the weights (c_i) depend on the population response to the lead or lag, relative to the population response to the lead or lag alone. If the instruction is to match the location of the leading stimulus, $i=1$.

$$c_1 = \frac{\hat{r}_1}{\hat{r}_{S1}} \quad (4)$$

If the instruction is to match the location of the lagging stimulus, $i=2$.

$$c_2 = 1 - \frac{\hat{r}_2}{\hat{r}_{S2}} \quad (5)$$

where r_{S1} and r_{S2} are the population response estimated using Eqn. 2 to a single binaural click presented in isolation at the leading and lagging location, respectively.

The weight c is restricted to fall between 0 and 1. Consistent with the weight c described by Shinn-Cunningham et al. (1993), $c_i=1$ indicates that the lead dominates lateralization entirely ($\alpha_i=\phi_1/f$), whereas $c_i=0$ indicates that the lag dominates lateralization completely ($\alpha_i=\phi_2/f$). In the current study, we assume that the response to the leading stimulus is not affected by the presence of the lag ($c_1 \approx 1$, $\alpha_1 \approx \phi_1/f$). 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 ($c_2 \approx 1$, $\alpha_2 \approx \varphi_1/f$). When the lagging stimulus begins to evoke a population response that closely resembles the neural response to the lagging source presented in isolation, precedence effect is weak ($c_2 \approx 0$, $\alpha_2 \approx \varphi_2/f$).

RESULTS

Simulations of Physiological Data

A model SMAX and SMIN neuron with the best ITD of 300 μs was built to simulate the response properties observed in physiological studies of PE. The ITD of the lagging click was fixed at the best ITD of the model IC neuron (300 μs). The ITD of the leading click varied from -900 to +900 μs , and ISDs ranged from 1 – 20 ms.

For both 10- and 20-ms ISDs, the responses to the leading sound were similar to those to a single stimulus, being strongest on the contralateral side for an ITD of 300 μs ; the model cell responded minimally at the negative ITDs on the ipsilateral side (see Fig. 2B). The similarity of the curves indicates that the response to the leading click is not affected by the presence of the lagging click. Fig. 2A displays corresponding physiological data (Litovsky and Yin, 1998b), which is in good agreement with the model.

Fig. 2C and D show the response of a SMAX cell observed by Litovsky and Yin (1998b) and the model simulation, respectively. The amount of suppression of the lagging response depends on the stimulus location of the leading source as well as on the ISD. At 20-ms ISD, the response dips at about 300 μs , when the leading click is near the neuron's best ITD. The model SMAX cell received more inhibition from an ipsilateral MSO cell tuned to the same ITD as its excitatory projection; therefore, the suppression of the lagging response was strongest when the leading source produced maximal response

of the neuron. At 10 ms, the suppression increased and the dip broadened. Finally, at 5 ms the response was suppressed completely at nearly all positions of the leading sound. This suppression of the lagging response is similar to the behavioral phase of localization dominance in that the perceived position of the paired sources appears to be governed by the leading source only.

Litovsky and Yin (1998b) also found a small number of units for which the lagging responses are most reduced when the lead produces the lowest responses. The model SMIN cell was able to simulate such response properties (compare Figs. 2E and 2F). The model SMIN cell received more inhibition from a contralateral MSO cell tuned to the ITD to which its excitatory projection would minimally respond. Therefore, suppression was strongest at the trough of the leading response.

Simulations of Psychophysical Data

For the generation of psychophysical predictions, the IC neuron population only consisted of SMAX units because SMAX units are thought to be more prevalent in the auditory pathway (Litovsky and Yin 1998b). 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 ITDs of a pointer stimulus (Litovsky and Shinn-Cunningham, 2001). For each stimulus, two 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 simulation, the ITDs of the lead were chosen between 0 and $-400 \mu\text{s}$ and the lagging ITD was held constant at $400 \mu\text{s}$. To allow a direct comparison with the conditions measured, all the data obtained when the lagging stimulus was located on the

left side (with the ITD of $-400 \mu\text{s}$) were generated by assuming the left/right symmetry of the model.

We found that lagging responses were suppressed, in some cases for ISDs as long as 20 ms. Clearly, a second sound image can be heard at ISDs for which some suppression remains. On the other hand, at 10-ms ISD, some individual neurons had already recovered completely. Although the responses to the lagging stimulus had recovered for some neurons at moderate ISDs, the population activity was spread out across neurons with different best ITDs. Therefore, the strength of the resulting perceived location was reduced, corresponding to a relatively low population response. It seems that perception of a second image occurs for ISDs that lead to two distinct neural responses across a population of neurons (one due to the lead and one due to the lag), both of them containing sufficient information to convey perception of a source at the corresponding location.

As in the behavioral measures, our model predictions show that the matched ITD was near the lead ITD at all ISDs when listeners matched the lead image (compare Figs. 3A and 3B, filled symbols). At short ISDs, the listener matched the lead ITD regardless of instructions, suggesting that the localization cues of the lead dominates (see open symbols in Fig 3). When instructions were to match the lag, the matched position approached the lag ITD for ISDs longer than 10 ms. One minor discrepancy between model and behavioral results is that in the model, localization dominance is slightly stronger at long ISDs. When lead and lag were spatially near to one another, the likelihood of perceiving two distinct images at their respective sources was lower than when lead and lag were spatial far apart in both model and behavior. However, this effect

is greater in the simulations than in the behavioral results, reflecting the fact that SMAX model cells showed stronger suppression of the lag when the lead and lag locations were close together.

CONCLUSIONS

A model was developed that simulated the responses of IC neurons in response to a pair of binaural clicks that could evoke the precedence effect. The model used existing models for auditory-nerve fibers, bushy cells in the cochlear nucleus, and cells in MSO. The IC model neuron 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. Suppression of the lagging response arises due to the long-lasting inhibition evoked by the leading stimulus. This suppression is modulated by ITD because the inhibition comes from cells that are themselves sensitive to stimulus ITD. Consistent with previous data, 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. A model IC built from the first type of model neurons can explain localization dominance reported in psychophysical studies of PE, whereby at short ISDs, the perceived location of a pair of clicks was dominated by the leading source. The strength of dominance decreased and the lagging sound was more likely to be heard near its own true location as the spatiotemporal separation of the lead and lag increased.

REFERENCES

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, and Horner K (ed) *Auditory Physiology and Perception*. Pergamon, Oxford, pp 539–545

Fitzpatrick DC, Kuwada S, Batra R, Trahiotis C (1995) Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol* 74: 2469–2486

Hartung K, Trahiotis C (2001) Peripheral auditory processing and investigations of the “precedence effect” which utilizes successive transient stimuli. *J Acoust Soc Am* 110: 1505–1513

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, 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(3):330-344

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

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, 79-82, October 2003

Shinn-Cunningham BG, Zurek PM, Durlach NI (1993) Adjustment and discrimination measurements of the precedence effect. *J Acoust Soc Am* 93: 2923–2932

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

Wallach H, Newman EB, Rosenzweig MR. (1949) The precedence effect in sound localization. *Am J Psychol* 52: 315–336

Figure Captions

FIG 1: Structure of the IC model, which incorporates models of medial superior olive (MSO) neurons, bushy cells in cochlear nucleus, and auditory-nerve fibers. The dorsal nucleus of the lateral lemniscus (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. Excitatory synapses are marked by “+” and inhibitory synapses by “-”. A: The structure of SMAX model neuron which has strong ipsilateral inhibition. B: The structure of SMIN model neuron which has strong contralateral inhibition.

FIG 2: Comparison of neuron response patterns to model response patterns. A, C, E: Responses of neurons. Reproduced from Litovsky and Yin (1998b) with permission from

J Neurophysiol. B, D, F: Responses of model IC cells. Top: Responses to the leading stimulus with the lag delayed by 10 ms 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 as well. ↓ 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 at ISDs of 5, 10, and 20 ms are shown.

FIG 3: The matched ITD α under various conditions, at ISDs of 1-15ms. 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: the lead is always at center ($0 \mu\text{s}$) and the lag is either on the right or left. Right column: lead and lag are on opposite sides. A: Matching results for subject S1. Reproduced from Litovsky and Shinn-Cunningham (2001) with permission from J Acoust Soc Am. B: Estimates of the model.

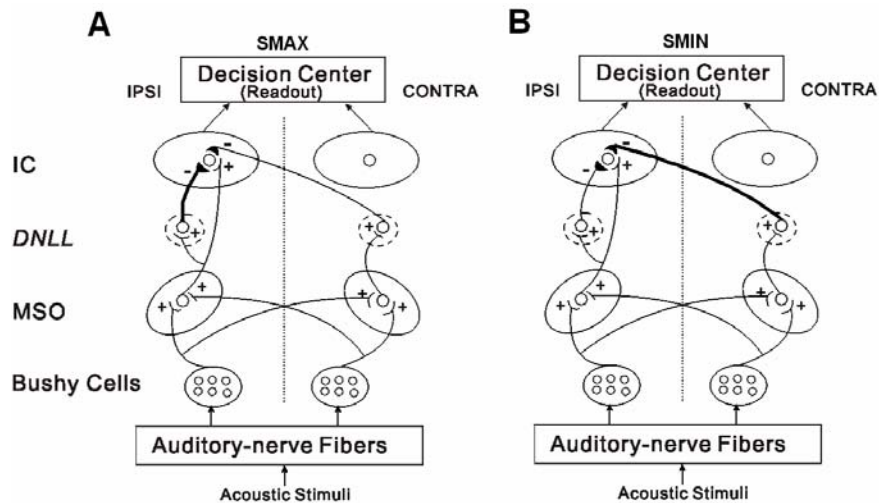


FIG 1

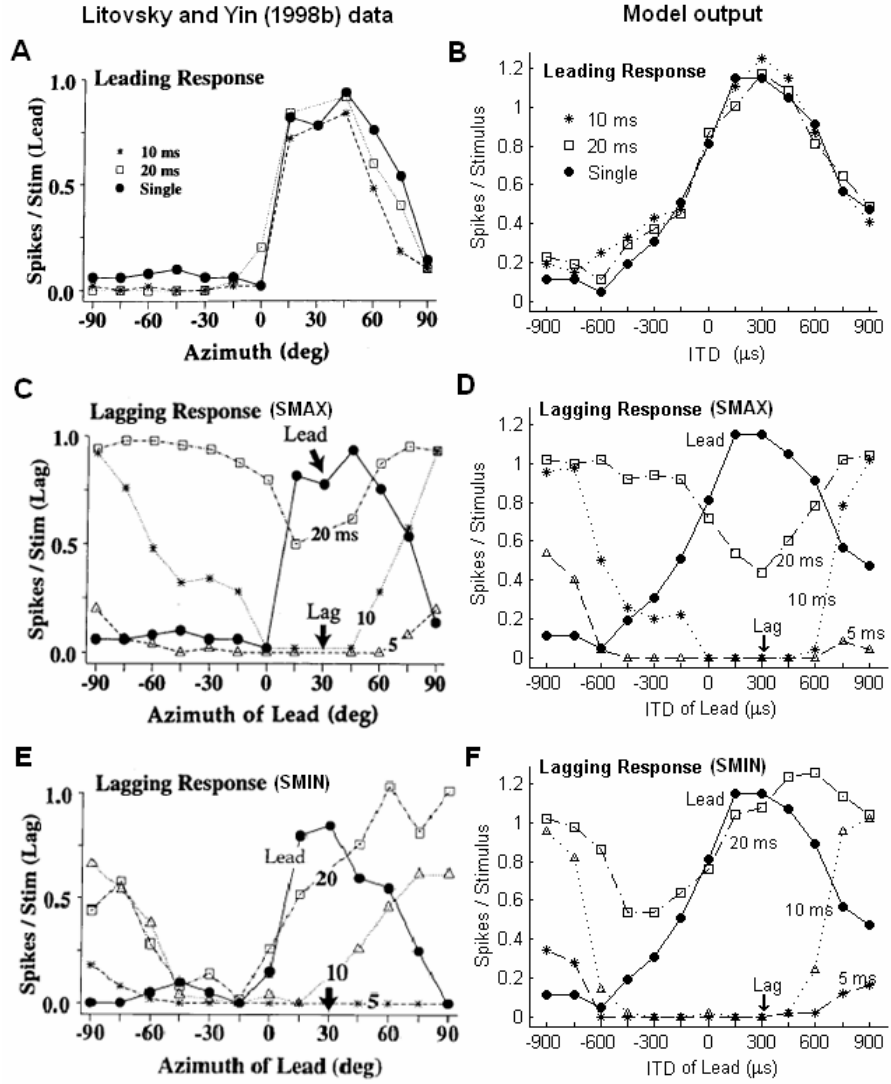


FIG 2

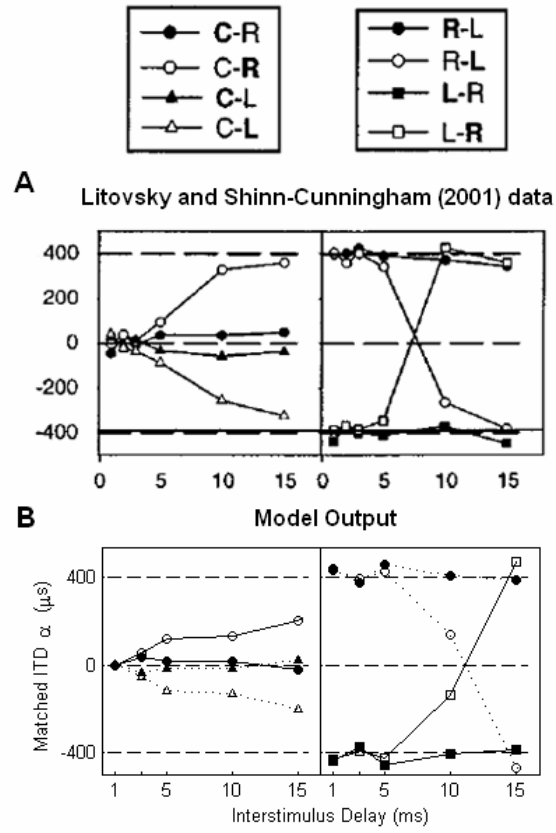


FIG 3