

Physiological measures of the precedence effect and spatial release from masking in the cat inferior colliculus.

R.Y. Litovsky^{1,3}, C. C. Lane^{1,2}, C. A. Atencio¹ and B. Delgutte^{1,2}

¹*Massachusetts Eye and Ear Infirmary, Eaton Peabody Laboratory
243 Charles St. Boston MA 02114, USA*

²*Harvard-MIT Speech and Hearing Sciences Program
Cambridge, MA, USA*

³*Boston University, Hearing Research Center
44 Cummington St., Boston MA 02215, USA
litovsky@bu.edu*

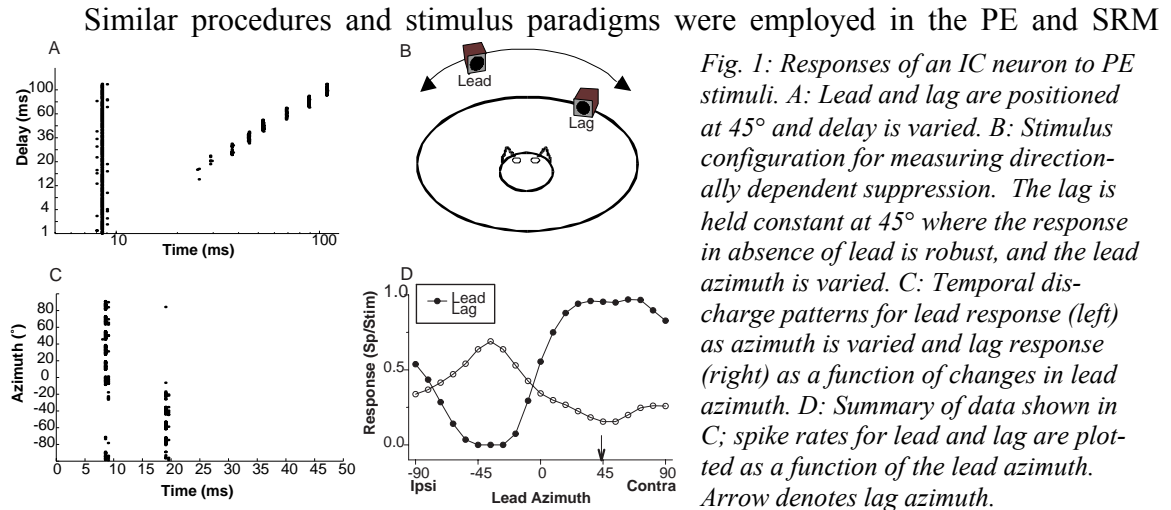
1. Introduction

Human listeners localize and recognize auditory objects in complex acoustic environments. Little is known about the neural mechanisms underlying this ability, which is often degraded in the hearing impaired. We investigated responses of single-units in the anesthetized cat inferior colliculus (IC) for two stimulus situations that have some characteristics of complex environments and that depend on the location of the sound sources: (1) two brief stimuli that simulate a direct sound and a single reflection, and (2) a pair of simultaneous sounds, one of which may mask the other. The IC is an obvious target for these investigations because most IC neurons are sensitive to source direction and individual localization cues such as interaural time (ITD) and level (ILD) differences. (For a review, see Irvine, 1992).

The first stimulus paradigm simulates the precedence effect (PE), which is the observation that two sounds occurring in rapid succession are perceived as a single auditory object localized near the leading sound (Litovsky et al., 1999). Single-unit studies in the IC (Yin, 1994; Fitzpatrick et al., 1995; Litovsky & Yin, 1998a, b) have identified a possible neural correlate of the PE in that the response to the lagging sound is suppressed for delays in which the PE occurs. A key question is whether these effects are due to a general suppressive mechanism akin to forward masking, or whether the neural suppression is specifically directional. To address this question, we characterize the relationship between the directional neural responses to the leading sound and the lagging sound and how this relationship depends on individual localization cues.

The second paradigm simulates a phenomenon called spatial release from masking (SRM), in which a signal is more easily detected when separated in space from a masker (e.g., Saberi et al., 1990). Single-unit studies of the IC have found that the neural detectability of a tone in noise can be improved if the signal ITD differs from the masker's (Jiang et al., 1997). We extend these results to neural masking release in simulated free field, which depends on other localization cues (ILD and spectral features) besides ITD (Bronkhurst and Plomp, 1988).

2. Methods



studies. Methods for recording from single units in the IC of anesthetized cats were described by Delgutte et al. (1999). Free-field stimuli were simulated by filtering signals through head-related transfer functions (HRTFs) measured in one cat by Musicant et al. (1990). All stimuli contained energy from 300 Hz to 30 kHz. PE stimuli consist of two delayed clicks, the “lead” and “lag”. For the SRM experiments, a 100-Hz click train or a 40-Hz chirp train was presented in continuous noise. Stimulus levels, when fixed, were chosen to produce a robust yet unsaturated response, usually 10–20 dB above a unit's threshold. In both experiments, we were primarily concerned with how the azimuth of the lead/masker affects responses to a fixed lag/signal.

3. Results

3.1 Precedence Effect

Responses to PE stimuli were initially measured for lead and lag stimuli both positioned at +45° contralateral to the recording site with delay varying from 1 to 100 ms. The temporal discharge patterns in Fig. 1A show strong responses to both stimuli at long delays and suppressed responses to the lagging stimulus at shorter delays. The half-maximal delay, or delay at which the lag response recovers to 50% of the maximum, is 30 ms. Half-maximal delays range from 3 to 55 ms in our data and are highly consistent with those reported in the anesthetized cat for actual free-field stimulation (Yin, 1994; Litovsky and Yin, 1998a).

To measure directional dependence of suppression, a delay slightly below the half-maximal delay was chosen, the lagging stimulus held at 45° contralateral, and the lead azimuth varied from -90° to +90° in the frontal hemifield (Fig. 1B). For the neuron in Fig. 1C–D, responses to the leading source vary with lead azimuth. If there were no suppression, the lag response would equal the response at +45° in the absence of a lead. The observed reduction in lag response indicates suppression resulting from the presence of the lead. Temporal discharge patterns for a 10-msec delay (Fig. 1C) show pronounced directionally dependent lag suppression; suppression is strongest for contralateral (positive) azimuths, where the lead evokes a response, and weaker for ipsilateral (negative) azimuths, where the lead evokes no response. A plot of spike rates for both stimuli against lead azimuth (Fig. 1D) makes this relationship between lead excitation and lag suppression evident.

For the neuron in Fig. 1, both the lead and lag responses are strongly modulated with lead azimuth. However, within the population, there are several types of relationships between lead and lag responses. Fig. 2 shows examples of 4 response types: (A) Strong modulation of both lead excitation and lag suppression, and a strong predictability of the lag response from the lead based on a linear regression analysis. This unit type represents about half of the population and suggests a simple relationship between excitatory and inhibitory mechanisms. (B) Strong modulation of both lead and lag responses, but no predictability of the lag suppression from the lead excitation. About one-third of neurons fall into this category. For this unit type, excitation and suppression do not have the same directional dependence. (C) Weak modulation in the lead response but strongly modulated lag suppression. (D) Strongly modulated lead excitation and weak or no

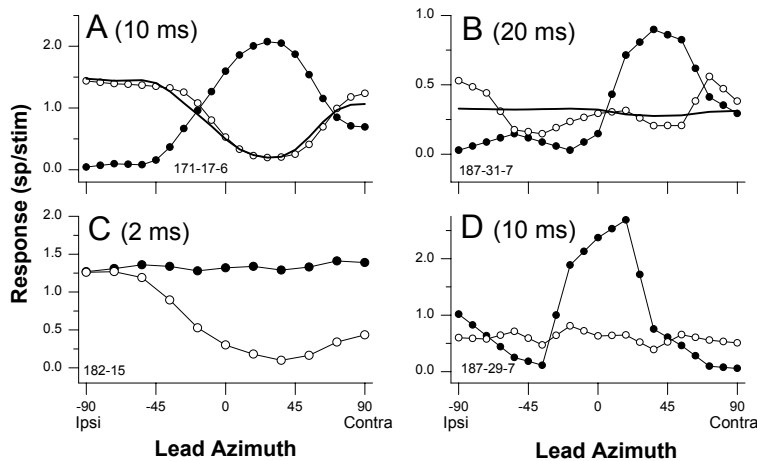


Fig. 2: Modulation response types. Lag stimulus was positioned at $+45^\circ$ and rates are plotted as a function of the lead azimuth. Lead-lag delay is indicated in each panel. Four types of neurons are shown, which exhibit different relationships between lead excitation and lag suppression (see text). Each panel plots lead response (\bullet) and lag response (\circ). In A and B the solid line represents a prediction of lag responses from lead responses using linear regression.

modulation in the lag suppression. About one tenth of neurons fall into each of C and D. These responses indicate a de-coupling between the directional mechanisms mediating excitation and suppression.

Our interest lies in identifying which localization cues determine the directional patterns of lead excitation and lag suppression. We used the approach of Delgutte et al. (1995), where stimuli were digitally manipulated so that some cues were varied with azimuth as in free field, while others were held constant. Fig. 3A-C shows results for a high-CF neuron (6.2 kHz). In response to the normal stimulus (“full cue”, A), the unit’s lead and lag responses are strongly modulated, and the lag response can be predicted from the lead response, the response type shown in Fig. 2A. When the lead ILD is held constant (B), the unit loses modulation of both the lead and lag responses, while when only lead ILD is varied (C), the responses resemble those seen with the full-cue stimuli. For this unit, then, ILD is a potent cue for both excitatory and suppressive responses.

Fig. 3D-F shows results for a lower-CF (1.9 kHz) unit where responses showed a dependence on multiple cues. The full-cue response in D again shows modulations in both the lead and lag responses. If only the lead ITD is varied (E), both lead and lag responses maintain their strong modulations with azimuth. Holding ITD constant (F) renders the lead response unmodulated, but leaves the lag suppression modulated to some degree. This result points to a dissociation between the mechanisms that mediate excitation in response to the lead and those mediating suppression of the lag response. While excitation depends primarily on ITD, suppression is mediated by a combination of ITD and other cues. Similar manipulations of directional cues were conducted in 33 neurons. For one-third of the neurons, the most potent cues for lead excitation and lag suppression were distinct. For these neurons, the lead excitation and the lag suppression may be pro-

duced by inputs from different subcollicular neural populations that would be sensitive to different directional cues.

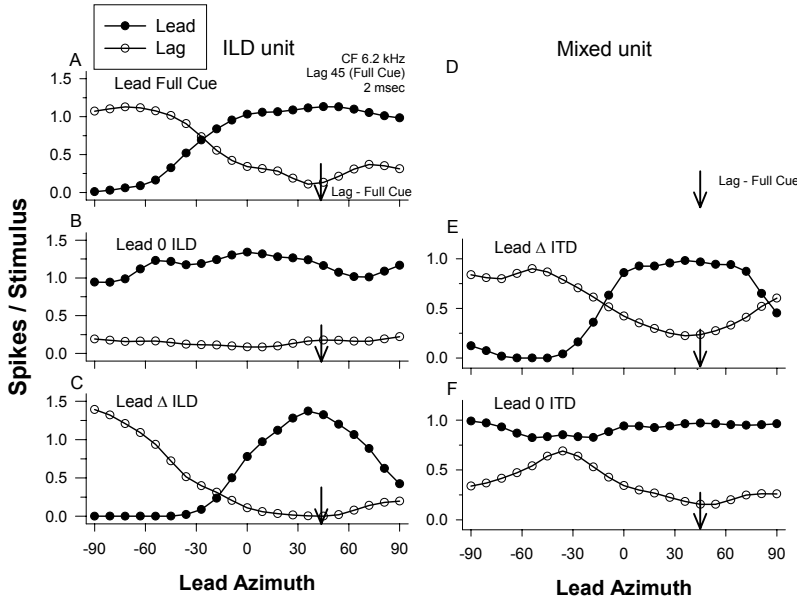


Fig. 3: Responses of lead (●) and lag (○) plotted as a function of leading source azimuth. Lag stimulus was always at +45° and contained all directional cues. The lead was varied from ±90°. For one neuron (CF 6.2 kHz, left) lead stimulus either contained all cues (A), or had fixed ILD (B) or only ILD varied (C). For a second neuron (CF 1.9 kHz, right) lead was either full cue (D), or only ITD varied (E) or had ITD fixed (F).

3.2 Spatial Release from Masking

We measured masked threshold as a function of noise azimuth for 19 neurons with CFs ranging from 1 to 40 kHz. We studied neurons that responded to either a click train or a chirp train and whose responses were masked by intense noise. Most units met these criteria. The signal was held at a particular azimuth (usually in front at 0°), while noise azimuth was varied from -90° to +90° in the frontal hemifield. For each neuron, the signal level was fixed, and the noise level adjusted to determine masked threshold. The signal was 200 ms in duration and delivered with a repetition rate of 2.5 Hz.

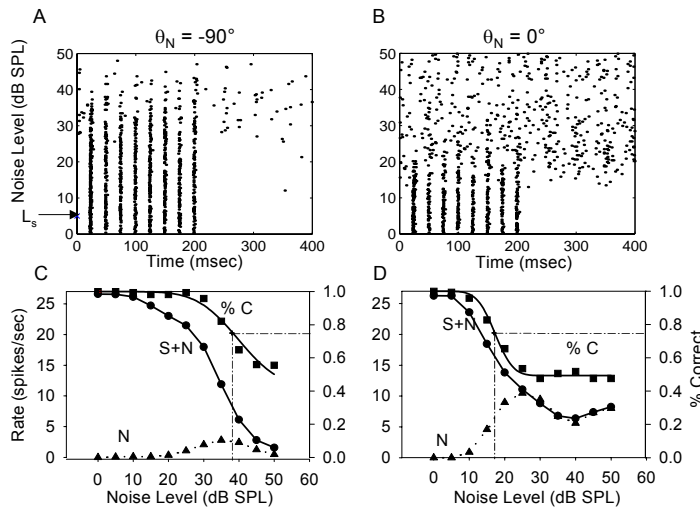


Fig. 4: For this unit (CF = 6 kHz), an 8-dB SPL chirp train (0-200 ms) was held at 0° with continuous noise at -90° (A,C) or 0° (B,D). A, B: Dot rasters for increasing noise levels in 5 dB steps with 20 stimulus presentations for each level. C, D: Rate for the signal and noise (S + N, ●), rate for the noise alone (N, ▲), and percent of stimulus presentations for which the number of spikes in S + N is greater than in N (%C, ■). The smooth line is a fit cumulative Gaussian function. Masked threshold is the noise level where the fitted curve crosses 75% (dash-dot line)

Masked threshold, L_n^{Th} , was defined in a manner similar to the psychophysical two-interval, two-alternative forced-choice paradigm where threshold is the level for which the subject gives the correct response 75% of the time. Here masked threshold was the noise level where the number of spikes in the signal-and-noise interval (0-200 ms) is greater than in the noise-only interval (200-400 ms) for 75% of the stimulus presentations (Fig. 4). In Fig. 4, there is a masking release of 21 dB when the noise is moved from 0° ($L_n^{Th} = 17$ dB) to -90° ($L_n^{Th} = 38$ dB).

Figs. 5A-C show three neurons' masked thresholds as a function of noise azimuth. The most effective masking azimuths, or the azimuths for which the lowest-level masker could mask the signal, are 18°, 54°, and 54° for these neurons, respectively. For the neurons studied, the most effective masking azimuths ranged from -54° to +54°. Most (17/19) neurons had most effective masking azimuths contralateral to the recording site, the side that usually excites IC neurons. The distribution of most effective masking azimuths has a broad maximum around 36°, roughly corresponding to the acoustic pinna axis (Phillips, 1982).

The maximum release from masking is the difference in masked threshold for the least effective and most effective masking azimuths. In Fig. 5A, these thresholds are 16 and 38 dB, respectively; therefore, the release from masking is 22 dB. The release from masking for all the neurons studied ranged from 3 to 44 dB, with a mean of 17 dB. These values are comparable to the 15-18 dB range seen psychophysically for broadband stimuli (Gilkey and Good, 1995; Saberi et al., 1991).

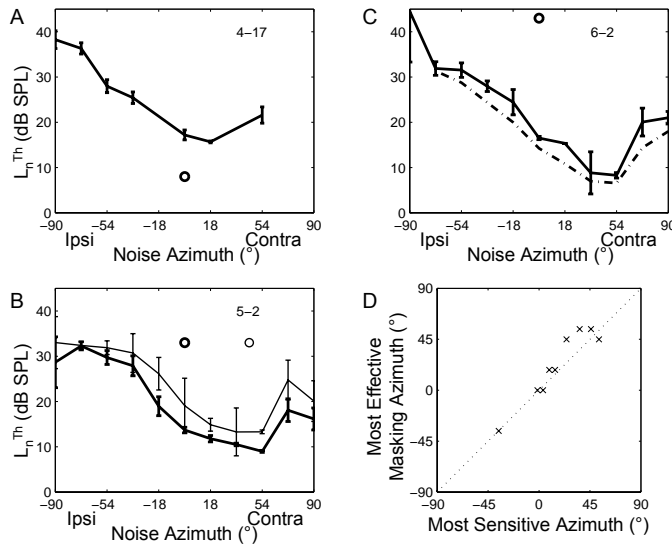


Fig. 5: A-C: Thick lines show masked threshold vs. noise azimuth for the signal at 0°. Signal level is indicated by the bold circle. B: Thin line shows masked threshold for signal at 45° contralateral (thin circle). C: Dashed lines show threshold in quiet for noise. The threshold in quiet is the noise level that elicits one spike every other stimulus presentation. D: Most effective masking azimuth vs. most sensitive masking azimuth for all the neurons that could be excited by continuous noise.

Because psychophysics shows that a masker is most effective when it is co-located with a signal, we measured masked thresholds for different signal azimuths in some neurons. For the neuron in Fig. 5B, changing the signal azimuth from 0° (thick line) to 45° (thin line) shifts the threshold curve vertically without changing its shape. Therefore, masking in this neuron did not depend on the separation between the signal and the masker, as seen in psychophysics; such dependence might be found by considering a population of neurons (see Discussion).

As discussed above, suppression studied with the PE paradigm could be predicted from the lead excitation in many neurons. Masking could also be predicted from the excitation produced by the noise masker for some neurons. For example, Fig. 5C shows both a neuron's masked threshold for the signal at 0° and the noise threshold in quiet as a function of noise azimuth. For this neuron, the threshold in quiet parallels the masked threshold curve. The most sensitive azimuth (the azimuth where threshold in quiet is lowest) coincides with the most effective masking azimuth at 54°. Fig. 5D shows the most sensitive azimuth against the most effective masking azimuth for all the neurons that were excited by continuous noise (9/19). For this group, the most sensitive noise azimuth is strongly correlated with the most effective masking azimuth ($\rho = .95$, $p < .001$), indicating that the excitation pattern is related to the masking pattern.

The continuous noise masker did not excite the other half of the neurons. For these neurons, the response to the signal was masked through a suppressive mechanism, possibly synaptic inhibition. We also observed suppression of the signal response by the noise for some of the neurons that were excited by the noise (e.g. Fig. 4D). Therefore, the correlation between excitation and masking may not indicate that excitation is responsible for the masking per se, but instead that the excitation and suppression have similar directionality, not unlike the Group A neurons in the PE experiments (Fig. 2A).

In the simplest case, masking patterns might be explained by the effective signal-to-noise ratio in one ear. To test this idea, we created a model of the auditory periphery consisting of a Gamma-tone filter centered at the CF of each unit followed by an envelope detector. We passed the noise through the HRTF for one ear (the one that most excited the unit studied) and then through the model. The noise level was adjusted to obtain a certain energy level at the model output, and this energy level was chosen to best fit the data. The resulting noise level was the model masked threshold. Fig. 6 shows masked threshold curves for two neurons and the model. The model predicts the masked threshold curve for the neuron in Fig. 6A, but greatly underestimates the change in masked threshold for the neuron in Fig. 6B. Overall, we were able to predict about one-third of the neurons' directional masking patterns with this simple monaural model. Presumably, binaural and temporal processing must be considered to predict masking patterns in the other two-thirds of the neurons.

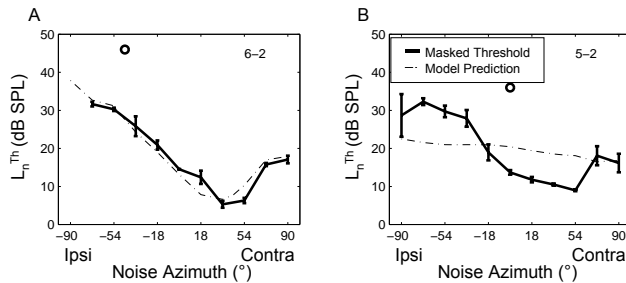


Fig. 6: Masked thresholds for two units with model predictions. The masked threshold in A is predicted by the energy in the contralateral ear, but the masked threshold in B is not predicted by this model.

4. Discussion and Conclusion

We found that for a vast majority of IC neurons, lag suppression and masking vary with lead or masker azimuth, and this directionality depends in part on binaural interactions. Inhibitory mechanisms are likely to play an important role in these effects.

There are further similarities between the PE results and the masking results if we make analogies between the lead and the masker on the one hand, and the lag and the signal on the other hand. For some neurons, masking and suppression appear to be fairly simple. For about half the neurons studied, suppression of the lag response was linearly predictable from the response to the lead (Fig. 2A, 3A). Similarly, directional masking patterns resembled noise thresholds in quiet for neurons that responded to the continuous masker (Fig. 5C,D). For these neurons, there is a close relationship between a sound's ability to excite and its directional pattern of masking or suppression. Because the directional pattern of excitation depends only on the masker azimuth, this result suggests that masking does not depend on the *interaction* between the signal and the masker, but on each of the two sounds independently. This prediction was verified in a few neurons in which masking patterns were measured for different signal azimuths (e.g., Fig. 5B), and similar results were observed with PE stimuli when the lag azimuth was varied (not shown). This separability of the effects of the signal and masker contrasts with psycho-physical data, where masking is strongest when the signal and the masker occupy the

same location (Gilkey and Good, 1995; Saberi et al., 1991). Similarly, several aspects of the PE are stronger when the two stimuli are close together (Litovsky and Shinn-Cunningham, 2000). For the simple neurons showing these effects, neural correlates of PE and SRM should be sought in the response of a *population* of neurons that respond preferentially to different spatial locations. A minimum requirement for such a population code is that the entire range of azimuths must be represented in the population. The wide range of best azimuths in Fig. 5D (when combined with responses from the opposite IC) suggests that a population code may be viable in the IC.

We also found neurons showing more complex forms of masking and echo suppression. For about half the neurons tested, the directional pattern of lag suppression could not be predicted from the response to the lead (Fig. 2B-D). Further decoupling between lead excitation and lag suppression was revealed by manipulating localization cues (Fig 3D-F). Distinct directional patterns of lead excitation and lag suppression might be produced by excitatory and inhibitory inputs from different subcollicular nuclei onto IC cells. The observation of multiple phases of excitation and inhibition in intracellular recordings from the IC (Kuwada et al., 1998) is consistent with this hypothesis. Regardless of the underlying mechanisms, these complex directional patterns show that neural suppression is not simply a form of forward masking and may provide a directional component for the PE. Neurons of this type were less common in the masking study, but half the neurons in our small sample did not respond to continuous noise so that the relationship between excitation and masking was not tested.

In summary, we have begun examining the neural mechanisms underlying the PE and SRM, two psychophysical phenomena likely to play a role when listening in complex acoustic environments. Both of these perceptual phenomena are known to depend on the relative directions of the first and second stimuli. Here we have shown that IC neurons show directional patterns of masking and suppression that correlate with these phenomena. However, a full understanding will require quantitative models for masking and suppression in populations of neurons.

5. References

- Bronkhorst, A.W., and Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J. Acoust. Soc. Am.*, 83, 1508-1516.
- Delgutte, B., Joris, P., Litovsky, R. and Yin, T.C.T. (1995). Relative importance of different acoustic cues to the directional sensitivity of inferior-colliculus neurons. In: G.A. Manley, G.M. Klump, C. Koppl, H. Fastl, and H. Oeckinghaus (Eds.), *Advances in Hearing Research*. World Scientific, London.
- Delgutte, B., Joris, P.X., Litovsky, R.Y., and Yin, T.C.T. (1999). Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. *J. Neurophys.* 81, 2833-2851.
- Fitzpatrick, D.C., Kuwada, S., Batra, R., and Trahiotis, C. (1995). Neural responses to simple, simulated echoes in the auditory brainstem of the unanesthetized rabbit. *J. Neurophys.*, 74, 2469-2486.
- Gilkey, R.H. and Good, M.D. (1995) Effects of frequency on free-field masking. *Hum. Factors* 37, 835-853.
- Irvine, D.R.F. (1992). Physiology of the Auditory Brainstem. In: A.N. Popper and R.R. Fay (Eds.), *The Mammalian Auditory Pathway: Neurophysiology*. Springer-Verlag, New York.
- Jiang, D., McAlpine, D., and Palmer, A.R. (1997). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-level functions. *J. Neurophys.* 77, 3085-3106.
- Litovsky, R.Y. and Yin, T.C.T. (1998a). Physiological studies of the precedence effect in the inferior colliculus of the cat: I. Correlates of psychophysics. *J. Neurophys.* 80, 1302-1316.
- Litovsky, R.Y. and Yin, T.C.T. (1998b). Physiological studies of the precedence effect in the inferior colliculus of the cat: II. Neural Mechanisms. *J. Neurophys.* 80, 1285-1301.
- Litovsky, R.Y., Colburn, H.S., Yost, W.A., and Guzman, S. (1999). The precedence effect. *J. Acoust. Soc. Am.* 106, 1633-1654.

- Litovsky, R.Y. and Shinn-Cunningham, B.G. (2000, in press). Investigation of the relationship between three common measures of precedence: fusion, localization dominance and discrimination suppression. *J. Acoust. Soc. Am.*
- Musicant, A.D., Chan, J.C.K., and Hind, J.E. (1990). Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons. *J. Acoust. Soc. Am.* 87, 757-781.
- Phillips, D.P., Calford, M.B., Pettigrew, J.D., Aitkin, L.M., and Semple, M.N. (1982). Directionality of sound pressure transformation at the cat's pinna. *Hearing Res.* 8, 13-28.
- Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., and Perrot, D.R. (1991). Free-field release from masking. *J. Acoust. Soc. Am.* 90, 1355-1370.
- Yin, T.C.T. (1994). Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J. Neuroscience*, 14, 5170-5186.

Research supported by Grant DC00119 from the NIH-NIDCD.