



PERGAMON

Neural Networks 12 (1999) 31–42

Neural
Networks

Contributed Article

Binaural cross-correlation and auditory localization in the barn owl: a theoretical study

Michele Rucci*, Jonathan Wray

The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

Received 28 October 1997; revised 15 July 1998; accepted 15 July 1998

Abstract

The barn owl is a nocturnal predator that is able to capture mice in complete darkness using only sound to localize prey. Two binaural cues are used by the barn owl to determine the spatial position of a sound source: differences in the time of arrival of sounds at the two ears for the azimuth (interaural time differences (ITDs)) and differences in their amplitude for the elevation (interaural level differences (ILDs)). Neurophysiological investigations have revealed that two different neural pathways starting from the cochlea seem to be specialized for processing ITDs and ILDs. Much evidence suggests that in the barn owl the localization of the azimuth is based on a cross-correlation-like treatment of the auditory inputs at the two ears. In particular, in the external nucleus of the inferior colliculus (ICx), where cells are activated by specific values of ITD, neural activation has been recently observed to be dependent on some measure of the level of cross-correlation between the input auditory signals. However, it has also been observed that these neurons are less sensitive to noise than predicted by direct binaural cross-correlation. The mechanisms underlying such signal-to-noise improvement are not known. In this paper, by focusing on a model of the barn owl's neural pathway to the optic tectum dedicated to the localization of the azimuth, we study the mechanisms by which the ITD tuning of ICx units is achieved. By means of analytical examinations and computer simulations, we show that strong analogies exist between the process by which the barn owl evaluates the azimuth of a sound source and the generalized cross-correlation algorithm, one of the most robust methods for the estimate of time delays. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Auditory localization; Barn owl; Inferior colliculus; Binaural cross-correlation; Interaural time difference

1. Introduction

The capability of localizing the position of a sound source through audition is common to many animal species including humans. Accurate and fast spatial localization of a sound source is crucial for the successful capture of prey or escape from predators. In the case of audition, spatial information is not explicitly represented at the input stages, but needs to be extracted by the brain. A signature of the position of a sound source is usually present in several features of auditory signals: some of these cues are monaural, such as the spectral modifications induced by the configuration of the pinnae, others are binaural. The latter cues are based on comparisons of the two input signals, such as differences in the amplitude and in the time of arrival of sounds at the two ears (interaural level differences (ILDs) and interaural time differences (ITDs)).

In most species studied, the ongoing ITD appears to be an important cue for localizing the azimuth of the sound source (Yin et al., 1987, Middlebrooks and Green, 1991, Konishi, 1993). Simple geometrical considerations reveal that, in the frontal field, ITD is linked to the azimuth of a sound source by a monotonic relationship (Woodworth and Schlosberg, 1962), so that the source position along the azimuthal axis is determined by the measured ITD. It has long been hypothesized that ITDs could be derived in the brain by cross-correlating the signals at the two ears (Licklider, 1951, Sayers and Cherry, 1957, Stern et al., 1988). Several psychophysical experiments in humans and theoretical models are in agreement with this hypothesis (Sayers, 1964, Lindemann, 1986, Saberi, 1996). According to this method, ITDs are estimated by measuring the time at which binaural cross-correlation reaches a maximum (see, for example, Carter, 1981). Since the cross-correlation function is based on a global comparison of the two waveforms, and does not depend on ad hoc parameters, this is a robust method for estimating the ITD between two auditory signals originating from the same sound source.

* Corresponding author. Tel.: +1-619-6262069; Fax: +1-619-6262099; E-mail: rucci@nsi.edu

In the last two decades, auditory localization has been carefully studied in the barn owl, a nocturnal predator that relies heavily on audition for hunting (for a review, see Konishi, 1993). Behavioral studies have shown that the barn owl uses ongoing ITDs and ILDs almost independently for estimating the azimuth and the elevation of the sound source (Moiseff and Konishi, 1981).

Neurophysiological and anatomical studies have shown that two separate neural pathways are specialized for analyzing ITDs and ILDs (Konishi et al., 1988). These pathways start from the cochlea and converge at the level of the central nucleus of the inferior colliculus (ICc). Neurons in the external nucleus of the inferior colliculus (ICx) are activated by auditory stimuli originating from sources located in specific positions of the surrounding space. It has been observed that the spatial selectivity of ICx neurons is the result of a selectivity for specific values of ITD and ILD which unequivocally determine the location of their receptive fields (Moiseff and Konishi, 1981, Knudsen, 1984). The locations of receptive fields vary systematically with the position of cells within the ICx, so as to give rise to an auditory map of space (Knudsen and Konishi, 1978a, Konishi, 1986). Several experimental observations suggest that the ITD tuning of ICx cells emerges as a result of a cross-correlation-like treatment of the signals at the two ears. In particular, it has recently been shown that space-specific neurons within the ICx are sensitive to some measure of the level of correlation of the input signals (Albeck and Konishi, 1995) and that binaural cross-correlation predicts the response of ICx neurons under conditions simulating summing localization (Keller and Takahashi, 1996a). However, it has also been observed that space-specific ICx neurons are less affected by noise than would be predicted by direct binaural cross-correlation (Keller and Takahashi, 1996a). Although intrinsic connections in the inferior colliculus are believed to play an important role in this final result, the exact mechanisms by which such a noise rejection is achieved are not known.

In this paper, we analyze from a theoretical perspective the origins of the azimuthal map of space in the ICx. By means of analytic examinations, as well as computer simulations, we show that the processes by which the barn owl determines the azimuth of a sound source bear a strong resemblance to the so-called generalized cross-correlation method (Knapp and Carter, 1976), a more robust algorithm than direct binaural cross-correlation. Generalized cross-correlation is based on a prefiltering of the input signals before the cross-correlation, so as to compensate for the errors introduced by a finite time window of observation, and by the presence of multiple sound sources or echos. By focusing on a model of the neural pathway dedicated to ITDs, we show that several elements cooperate to improve the signal-to-noise ratio in the activation of ICx units. In particular, we illustrate that the characteristic response to different ITDs of units in the nucleus laminaris together with the intrinsic connections in the inferior colliculus

operate a peculiar frequency weighting of the input power spectral densities, that improve the reliability of ITD estimates in noisy environments. This operation is similar to the one performed by several versions of the generalized cross-correlation algorithm.

In Section 2 we briefly review the generalized cross-correlation method for the estimate of time delays. In Section 3 we describe a mathematical and computational model of the tectal pathway dedicated to azimuth localization in the barn owl. The results of the analysis of this model are presented in Section 4. Finally, a brief discussion is included in Section 5.

2. The generalized cross-correlation method

Let $s_L(t)$ and $s_R(t)$ be the auditory signals at the left and right ears respectively. In the case of a single remote sound source, $s(t)$, a common way to model these signals is

$$s_L(t) = s(t) + n_L(t) \quad (1)$$

$$s_R(t) = ks(t + \Delta) + n_R(t)$$

where Δ is the time delay which we want to evaluate, k is a scalar, and $n_L(t)$ and $n_R(t)$ are noise contributions.

A well-known method for estimating Δ is to search the maximum of the cross-correlation function $c(t)$ between the two input signals:

$$\max_T c(t) = \max_T (s_L(t) \times s_R(t)) \quad (2)$$

where the cross-correlation is defined as

$$s_L(t) \times s_R(t) = \int_{-\infty}^{\infty} s_L(\tau) s_R(\tau + t) d\tau \quad (3)$$

This method is based on the observation that when $s_R(t) = s_L(t - \Delta)$, the cross-correlation $s_L \times s_R$ reaches its global maximum at time Δ (see, for example, Bracewell, 1986).

If the noise terms in Eq. (1) are uncorrelated with the signal $s(t)$, the cross-correlation of the left and right signals yields

$$c(t) = kc_s(t - \Delta) + c_n(t) \quad (4)$$

where $c_s(t)$ is the autocorrelation of $s(t)$, and $c_n(t)$ is the cross-correlation of the two noise terms. In the case in which n_L and n_R are white processes, their cross-correlation is zero and $c(t)$ is a translated copy of $c_s(t)$.

Since binaural cross-correlation implies a global comparison of the input signals, this method is robust with respect to the superposition of noise. This is important, since it guarantees that the time at which the maximum of the cross-correlation function occurs is a good estimate of the ITD, even if the two signals are not exact translated copies. In practice, however, the accuracy of time measurements provided by the direct application of Eq. (2) is limited by a number of factors. A first obvious problem is related to the

fact that real systems operate on a time window of finite extent and $c(t)$ can only be approximated. If W is the length of the observation interval, then the estimated cross-correlation at time t , $\hat{c}_t(\tau)$, is given by

$$\hat{c}_t(\tau) = \frac{1}{W - \tau} \int_{t-W}^{t-\tau} s_L(\lambda) s_R(\lambda + \tau) d\lambda \quad (5)$$

and in general will be different from $c(\tau)$.

A second problem arises in the presence of multiple sound sources. In this case, which can be modeled as

$$s_L(t) = s_0(t) + s_1(t) + \dots + n_L(t) \quad (6)$$

$$s_R(t) = k_0 s_0(t + \Delta) + k_1 s_1(t + \Delta_1) + \dots + n_R(t)$$

multiple peaks, corresponding to the various delays, are present in the cross-correlation function. Assuming n_L and n_R to be white processes and the signals $s_i(t)$ to be uncorrelated with each other, the cross-correlation $c(t)$ is equal to a weighted superposition of the autocorrelations of the various signals, c_i , each translated by its own delay:

$$c(t) = \sum_i k_i c_i(t - \Delta_i) \quad (7)$$

It is clear from Eq. (7) that, in order to be able to discern the different delays in $c(t)$, each autocorrelation c_i must be characterized by a narrow peak. If c_i has multiple or broad peaks, then false target locations may arise even under the ideal assumption of uncorrelated noise and signals.

In order to compensate for these problems, the generalized cross-correlation method has been proposed (Knapp and Carter, 1976), in which the two input signals s_L and s_R are prefiltered before calculating their cross-correlation. The goal of this method is to ensure large and narrow peaks in the terms in Eq. (7), while at the same time ensuring stability with respect to the problem of a finite time window of observation.

It is convenient to describe the operation of the filters, by using the frequency formulation of the cross-correlation (see, for example, Bendat and Piersol, 1993)

$$c(t) = \int_{-\infty}^{\infty} G_{LR}(f) e^{j2\pi ft} df \quad (8)$$

where $G_{LR}(f)$ is the cross-power spectrum of s_L and s_R . If $H_L(f)$ and $H_R(f)$ are the transfer functions of the filters in the two input channels, the cross-power spectrum of the two

filtered signals is equal to

$$G_{LR}^g = H_R(f) H_L^*(f) G_{LR}(f) \quad (9)$$

and their cross-correlation

$$\begin{aligned} c_g(t) &= \int_{-\infty}^{\infty} H_R(f) H_L^*(f) G_{LR}(f) e^{j2\pi ft} df \\ &= \int_{-\infty}^{\infty} \Psi_g(f) G_{LR}(f) e^{j2\pi ft} df \end{aligned} \quad (10)$$

where

$$\Psi_g(f) = H_R(f) H_L^*(f) \quad (11)$$

Some popular preprocessing functions are shown in Table 1 (see Knapp and Carter (1976) for a detailed analysis of these and other methods). A strategy common to several of these filters is to operate on the input signals so as to ‘prewhiten’ them by equalizing the amplitude of the cross-power spectrum in Eq. (10). For example, in the case of a single delay, Δ , and absence of noise, the phase transform method gives

$$c_{PT}(t) = \int_{-\infty}^{\infty} \frac{G_{LR}(f)}{|G_{LR}(f)|} e^{j2\pi ft} df = \int_{-\infty}^{\infty} e^{j\theta(f)} e^{j2\pi ft} df = \delta(t - \Delta) \quad (12)$$

where $\theta(f) = 2\pi f\Delta$ is the phase of the cross-power spectrum. Thus, the cross-correlation of the filtered signals is a Dirac function centered on Δ .

In practice, however, the application of the phase transform method is limited by the effect of noise which alters $\theta(f)$, particularly in the frequency bands in which the signals have less power or the noise is high. Due to the action of noise, the linear dependence of $\theta(f)$ on frequency is corrupted, and Eq. (12) is no longer valid. A way to compensate for such a problem is to weight the amplitude of the cross-power spectrum by taking into account the level of noise at different frequencies. In the maximum likelihood method Ψ_g is dependent on the coherence function, γ_{LR}

$$\gamma_{LR}(f) = \frac{G_{LR}(f)}{\sqrt{G_{LL}(f) G_{RR}(f)}} \quad (13)$$

so that frequency bands with lower coherence values contribute less to the final estimate. It can be shown (Knapp and Carter, 1976) that for the maximum likelihood method the amplitude of the cross-power spectrum at a particular frequency is normalized by a factor proportional to the variability of the estimate of the phase at that frequency:

$$c_{ML}(t) \propto k \int_{-\infty}^{\infty} \frac{e^{j\hat{\theta}}(f)}{\sigma_{\hat{\theta}}^2(f)} e^{j2\pi ft} df \quad (14)$$

In this way, different frequencies give different contributions to the estimate of cross-correlation depending on the amount of noise present. An example of the application of the phase transform and maximum likelihood methods is illustrated in

Table 1
Frequency weightings according to method

Method	Frequency weighting
Phase transform	$1/ G_{LR}(f) $
Smoothed coherence	$1/\sqrt{G_{LL}(f)G_{RR}(f)}$
Maximum likelihood	$ \gamma_{LR}(f) ^2/ G_{LR}(f) (1 - \gamma_{LR}(f) ^2)$

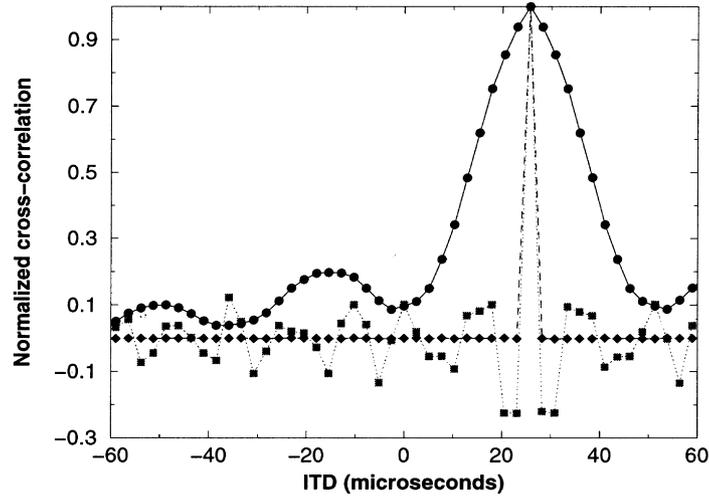


Fig. 1. Application of two different frequency weightings on the cross-correlation between two signals. The input signals were delayed by 25 μ s and no noise was superimposed. (Circles) Direct cross-correlation; (diamonds) phase transform method; (squares) maximum likelihood method.

Fig. 1. Both methods give rise to a reduction of the width of the ITD peak with respect to direct cross-correlation.

3. Modeling the pathway for auditory azimuth localization

The overall organization of the tectal pathways for azimuth localization in the barn owl is schematically illustrated in Fig. 2. The auditory pathway starts with the magnocellular cochlear nuclei, and meets in the optic tectum the visual pathway composed of direct retino-tectal projections. In the tectum, visual and auditory maps of space are aligned with a motor map, which participates in the production of orienting behavior (Knudsen and Brainard, 1995). As shown by electrophysiological recordings, an auditory map of space is first found at the level of the ICx, where the contributions of different frequency laminae at the level of the central nucleus of the inferior colliculus (ICc) are brought together. The neural structures modeled in this paper are the ones included in the dashed box of Fig. 2. This model, which is illustrated in Fig. 3, is a subsection of a larger computational model that we have recently developed for studying the plasticity of map alignment in the optic tectum (Rucci et al., 1997). It includes the nucleus laminaris (NL), which is the first stage where the signals from the two ears are brought together, the ICc and the ICx. Each structure in the model was composed of a collection of units, each implemented as a leaky integrator. The output of a unit can be viewed as representing the average firing rate of a collection of cells and its response properties is representative of a typical cell within such a group.

3.1. Nucleus laminaris

The nucleus laminaris was modeled as a bidimensional array in which unit sensitivity varied systematically with respect to frequency along one axis and interaural time delays

along the other (Carr and Boudreau, 1993). In this way, the unit at location (i,j) in the map was characterized by a preferred mean interaural delay \bar{T}_i and by a preferred frequency \bar{f}_i . The activation of NL units was empirically designed to fit as closely as possible the physiological data available in the literature. The particular anatomical and physiological

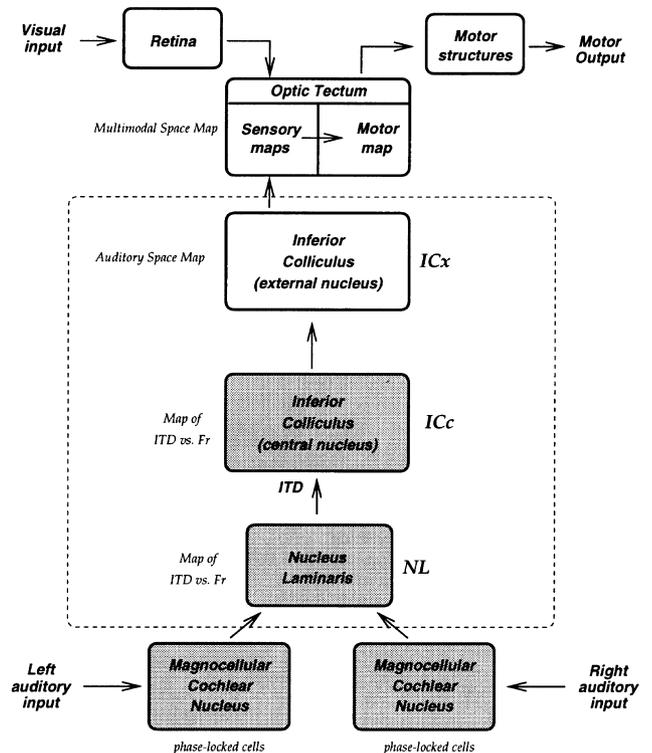


Fig. 2. Schematic representation of the neural pathways to the optic tectum involved in the process of azimuth localization in the barn owl. The auditory and the visual pathways converge at the level of the tectum, and project to motor structures. The opaque boxes indicate auditory structures organized in different frequency laminae. The frequency organization disappears at the level of the external nucleus of the inferior colliculus, where an auditory map of space is first synthesized. The dashed box illustrates the structures that have been considered in this paper.

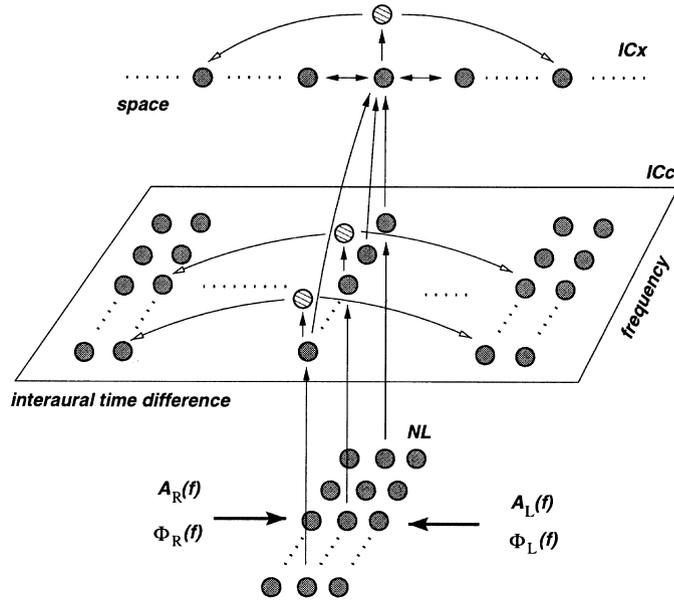


Fig. 3. The considered model of the tectal pathway for auditory azimuth localization in the barn owl. Three stages sequentially process the input signals. In the nucleus laminaris and in the central nucleus of the inferior colliculus units are organized in frequency laminae and respond periodically to the input ITD. In the ICx each unit responds maximally to a specific value of ITD. Unit-preferred ITDs shift systematically with the position of units in the array, so as to create an auditory map of the azimuthal space. Filled and empty arrows (circles) represent excitatory and inhibitory connections (units) respectively.

mechanisms underlying such neural responses in the barn owl (for a review see Konishi et al., 1988) were not modeled explicitly. For binaural stimulation, when the two signals $s_L(t)$ and $s_R(t)$ with Fourier transforms $A_L(f)e^{j\phi_L(f)}$ and $A_R(f)e^{j\phi_R(f)}$ (evaluated over the time-window of observation W) were applied as inputs, the activation of NL units was dependent on the unit characteristic parameters \bar{T}, \bar{f} as follows:

$$U_{\bar{T}, \bar{f}}^{NL} = \mathcal{F}(A_L(\bar{f})A_R(\bar{f})) [\cos(\phi_{LR}(\bar{f}) - 2\pi\bar{f}\bar{T}) + 1]G_{\bar{f}\sigma}(f) \quad (15)$$

where \mathcal{F} is a squashing function that controls the change in firing rate for different amplitudes of the input signals (in the simulations described in this paper a logarithmic function was used), $\phi_{LR}(\bar{f}) = \phi_L(\bar{f}) - \phi_R(\bar{f})$ is the difference

of phase between the left and right input signals at frequency \bar{f} , and $G_{\bar{f}\sigma}(f)$ is a Gaussian function with mean \bar{f} and variance σ . In the barn owl, the firing of NL neurons has been observed to be tuned to frequency and be periodic with respect to ITD (Carr and Konishi, 1990). In addition, it has recently been observed that the activity of NL neurons is proportional to the average intensity of the binaural stimulation (Peña et al., 1996). The activation of three units in the NL, as well as the global pattern of activation in the NL map are shown in Fig. 4.

3.2. Central nucleus of the inferior colliculus

As shown in Fig. 3, the central nucleus of the inferior colliculus (ICc) was modeled as a bidimensional map of

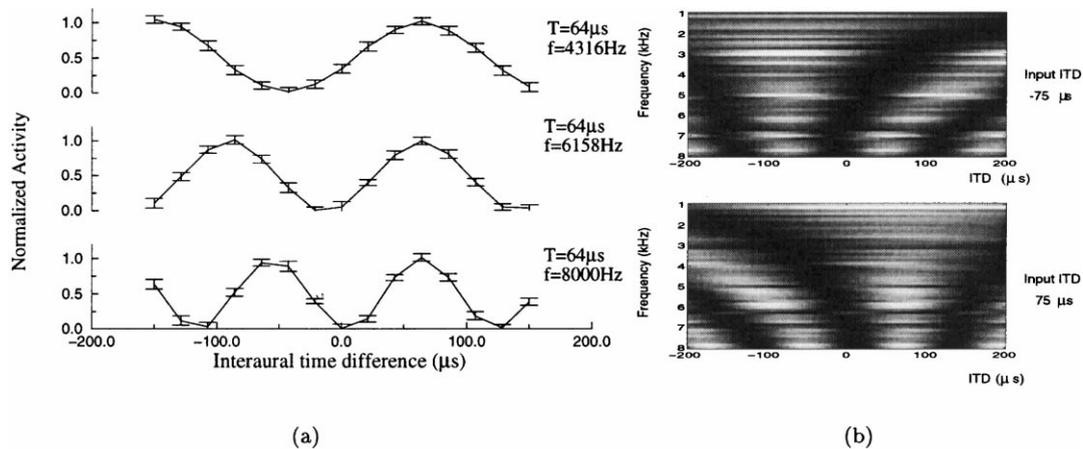


Fig. 4. Characterization of the modeled NL. (a) ITD tuning function of three units with different \bar{f} . (b) Two patterns of activation in the NL (top) ITD = $-75 \mu s$; (bottom) ITD = $+75 \mu s$): units are aligned in frequency laminae on the y axis, and according to their ITD sensitivity along the x axis. The NL included 50×100 units (a) reprinted with permission from Rucci et al. (1998), © 1998 IEEE).

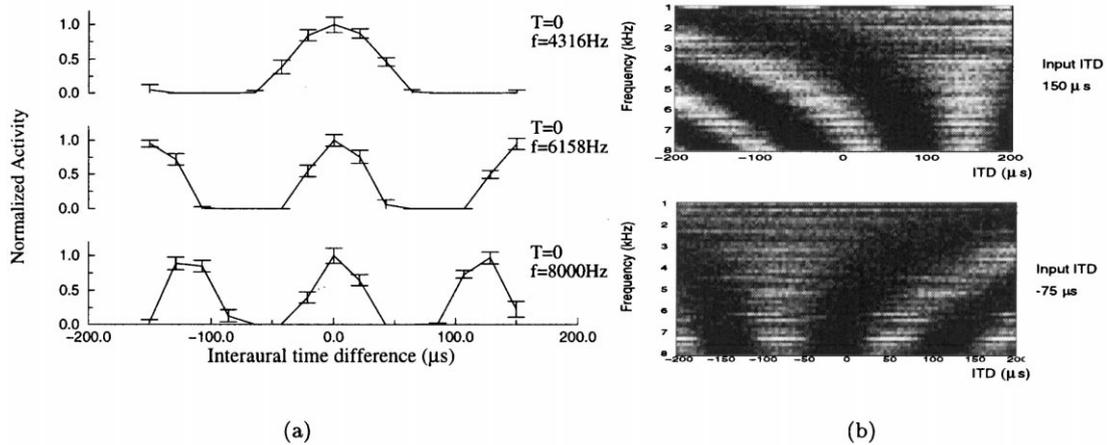


Fig. 5. Characterization of the modeled ICc. (a) ITD tuning function of three units in different frequency laminae. (b) Two patterns of activation in the ICc ((top) $ITD = 150\ \mu\text{s}$; (bottom) $ITD = -75\ \mu\text{s}$). Units are aligned in frequency laminae on the y axis, and according to their ITD sensitivity along the x axis. The ICc included 50×100 units ((a) reprinted with permission from Rucci et al. (1998), © 1998 IEEE).

frequency vs ITD. The map contained both excitatory and inhibitory units. Excitatory units received topographically organized projections from the NL and projected both to the ICx and to surrounding units in the ICc. They also received connections from inhibitory units in more distant regions of the ICc map located in neighboring frequency laminae. This model followed from both anatomical studies on the connections from the NL to the ICc (Knudsen, 1984, Takahashi and Konishi, 1988), and physiological investigations of the ICc (Wagner et al., 1987). The sensitivities of three ICc units to different values of ITD are shown in Fig. 5a. Unit activation was characterized by a narrow frequency range and showed a periodicity with respect to ITD, with the period determined by the neuron preferred frequency. This is similar to what occurs in the ICc of the barn owl (Konishi et al., 1988). Given the layout of unit sensitivities in the ICc map, each array of neurons perpendicular to the frequency axis (often indicated as a characteristic array) had an array-specific ITD, which represented the value of ITD that activated all the neurons of the array at the same relative

response level. Two typical maps of activation for sounds emitted by sources located in two different positions are shown in Fig. 5b. Note that the pattern of activation is centered on the characteristic array that has the same array specific ITD as the input ITD.

3.3. External nucleus of the inferior colliculus

The external nucleus of the inferior colliculus was modeled as a one-dimensional array of excitatory and inhibitory units. Following the data on the barn owl (Knudsen, 1983, Knudsen, 1984, Wagner et al., 1987, Fujita and Konishi, 1991), excitatory ICx units received afferents from ICc units in all the isofrequency laminae and within a limited range of ITD. They also received connections from adjacent excitatory units and from inhibitory units in more distant regions of the ICx map. Excitatory units projected to both excitatory and inhibitory units located nearby. As a result of this pattern of connectivity ICx units responded maximally to specific values of ITD. The unit's preferred ITD varied

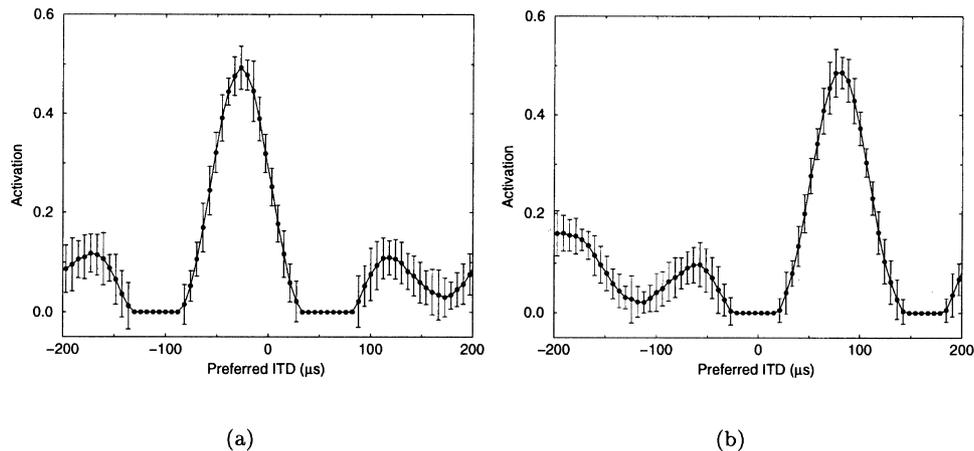


Fig. 6. Patterns of activation of the modeled ICx when binaural signals with ITD equal to (a) $-27\ \mu\text{s}$ and (b) $80\ \mu\text{s}$ are applied as inputs. The graphs show the mean and standard deviation of unit activity evaluated over 10 repetitions. In this example, the ICx map included 100 units.

systematically with the position of the unit within the map, so as to give rise to a representation of the auditory space. This is similar to what has been observed in the ICx of the barn owl (Knudsen et al., 1977, Knudsen and Konishi, 1978b, Takahashi and Konishi, 1986). Patterns of activation in the ICx when binaural stimuli with different ITDs are applied as input are shown in Fig. 6.

4. Model analysis

In order to investigate the contributions of different components of the system to the final activity of ICx units, we have analyzed the model in two distinct steps. First, to study the effect of the feed-forward connectivity we simplified the network by removing the intrinsic connections in the inferior colliculus. When this simplification is made an analytic examination of the activation of ICx units is possible. The results of this analysis are shown in the first part of this section. In the second part of this section we focus on the analysis of the effects of lateral interactions. Due to the analytical complexity of the model in the presence of reciprocal connections, we used computer simulations.

4.1. Removal of lateral connections

We assume, for simplicity, that all the neurons are linear and do not possess a memory of previous levels of activation, and each frequency lamina in the model is sensitive only to a specific frequency value, so $G_{\bar{f}\sigma}(f) = \delta(f - \bar{f})$ in Eq. (15), where $\delta(f - \bar{f})$ is the Dirac delta function centered on \bar{f} . We also assume, for the moment, that $\mathcal{F}(x) = x$ in Eq. (15). Under these conditions, it is possible to formally evaluate the activation of ICx units in the presence of the input signals $s_L(t)$ and $s_R(t)$, when the effect of the intrinsic connections has been removed in the ICx and in the ICc. This is a procedure similar to the one performed experimentally by Fujita and Konishi (1991).

Consider an ICx unit with preferred ITD value T . Let Δt be the spread of projections between ICc and ICx, that is each ICx unit receives connections from the ICc units with interaural time delay in the range $(T - \Delta t, T + \Delta t)$, and ω_{if}^T be the strength of the connection between the ICc unit characterized by the parameters t and f and the considered ICx unit. By approximating the summation over a large number of ICc units with an integral, the activation of this unit is

$$U_T^{\text{ICx}} \approx \int_{T - \Delta t}^{T + \Delta t} \int_{f_a}^{f_b} \omega_{if}^T A_L(f) A_R(f) [\cos(\phi_{\text{LR}} - \phi) + 1] df dt \quad (16)$$

where $\phi = 2\pi ft$ and (f_a, f_b) is the frequency range of ICx units. Assuming the case of random connection strengths with mean value $\langle \omega_{if}^T \rangle = \eta$, the mean value of activation

of an ICx unit is given by

$$\langle U_T^{\text{ICx}} \rangle \approx \eta \int_{T - \Delta t}^{T + \Delta t} \int_{f_a}^{f_b} A_L(f) A_R(f) [\cos(\phi_{\text{LR}} - \phi) + 1] df dt \quad (17)$$

For input signals with power in the range (f_a, f_b) , the estimated cross-correlation over the time window W is equal to,

$$\begin{aligned} \hat{c}(\tau) &= \frac{1}{W - \tau} \int_{t - W}^{t - \tau} s_L(\lambda) s_R(\lambda + \tau) d\lambda \\ &= \frac{2}{W - \tau} \int_{f_a}^{f_b} A_L(f) A_R(f) \cos(\phi_{\text{LR}} - \phi) df \end{aligned} \quad (18)$$

By substituting Eq. (18) in Eq. (17), we get

$$\langle U_T^{\text{ICx}} \rangle \approx \eta \int_{T - \Delta t}^{T + \Delta t} \hat{c}(t) dt + \int_{f_a}^{f_b} A_L(f) A_R(f) df \quad (19)$$

where the second term is a constant factor related to the power of the input signals. Eq. (19) shows that, in this simplified case, the mean activation of units in the ICx is proportional to a smoothed version of the cross-correlation function. The time integral in Eq. (19) has the contrasting effects of reducing the noise contribution from one side, and broadening the peaks of the cross-correlation function from the other.

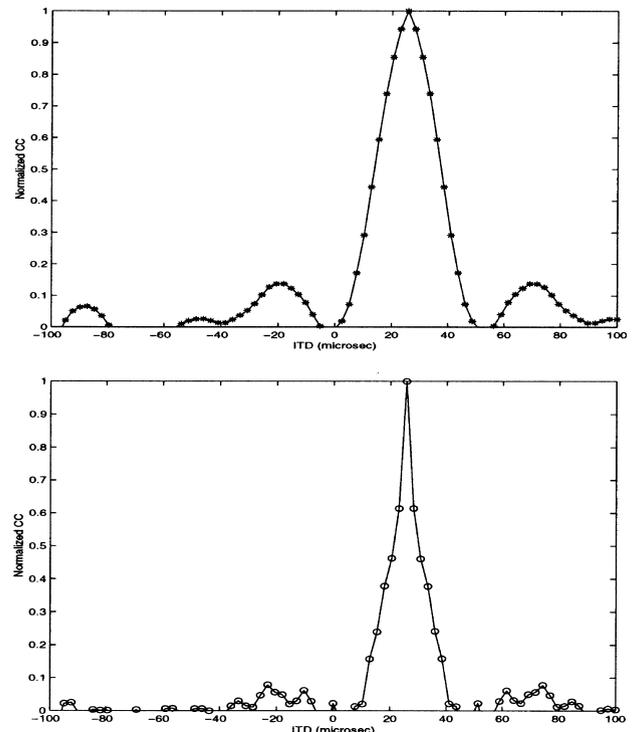


Fig. 7. Effect of logarithmic filtering of the input signals on binaural cross-correlation. The two signals are delayed copies of the same signal as in Eq. (1) ($\Delta = 25 \mu\text{s}$). (Top) direct cross-correlation of the input signals. (Bottom) cross-correlation after logarithmic filtering.

In the more realistic case in which $\mathcal{F}(x)$ in Eq. (15) is a nonlinear function, it can be shown with a similar procedure that the average activation of ICx units is equal to

$$\begin{aligned} \langle U_T^{\text{ICx}} \rangle &\approx \int_{T-\Delta t}^{T+\Delta t} \int_{-\infty}^{\infty} \Psi_{\mathcal{F}}(f) \hat{G}_{\text{LR}}(f) e^{j2\pi ft} df dt + \tilde{n} \\ &= \frac{\eta}{2} \int_{T-\Delta t}^{T+\Delta t} c_g(t) dt + \tilde{n} \end{aligned} \quad (20)$$

where $G_{\text{LR}}^g = 2A_L A_R e^{j\phi_{\text{LR}}}/W$ is the estimate of the cross-power spectrum in the window of observation and

$$\Psi_{\mathcal{F}}(f) = \frac{\mathcal{F}(A_L(f)A_R(f))}{A_L(f)A_R(f)}$$

If $\mathcal{F}(x)$ is a squashing function as experimental observations seem to indicate, the processor $\Psi_{\mathcal{F}}(f)$ operates in a similar way to the phase transform shown in Table 1. It introduces a

‘prewhitening’ of the signals by reducing the range of variability of the spectral amplitudes of the input signals. As a result, the peaks of the cross-correlation become sharper. An example in which a logarithmic compression, $\mathcal{F}(x) = \log(x)$, is applied to the input signals is shown in Fig. 7. As illustrated in the figure, the cross-correlation of the log-filtered signals presents a sharper peak than the direct cross-correlation of the inputs.

4.2. Computer simulations

The behavior of the network in the presence of effective lateral connections was studied by means of computer simulations (see also Rucci et al., 1997, 1998). In this section we consider results obtained with a system composed of 100×50 units in both the NL and ICc, and 100 units in the ICx (the number of interneurons was equal to the number of excitatory units). When lateral interaction among units was in place, the activation of ICx units was no longer

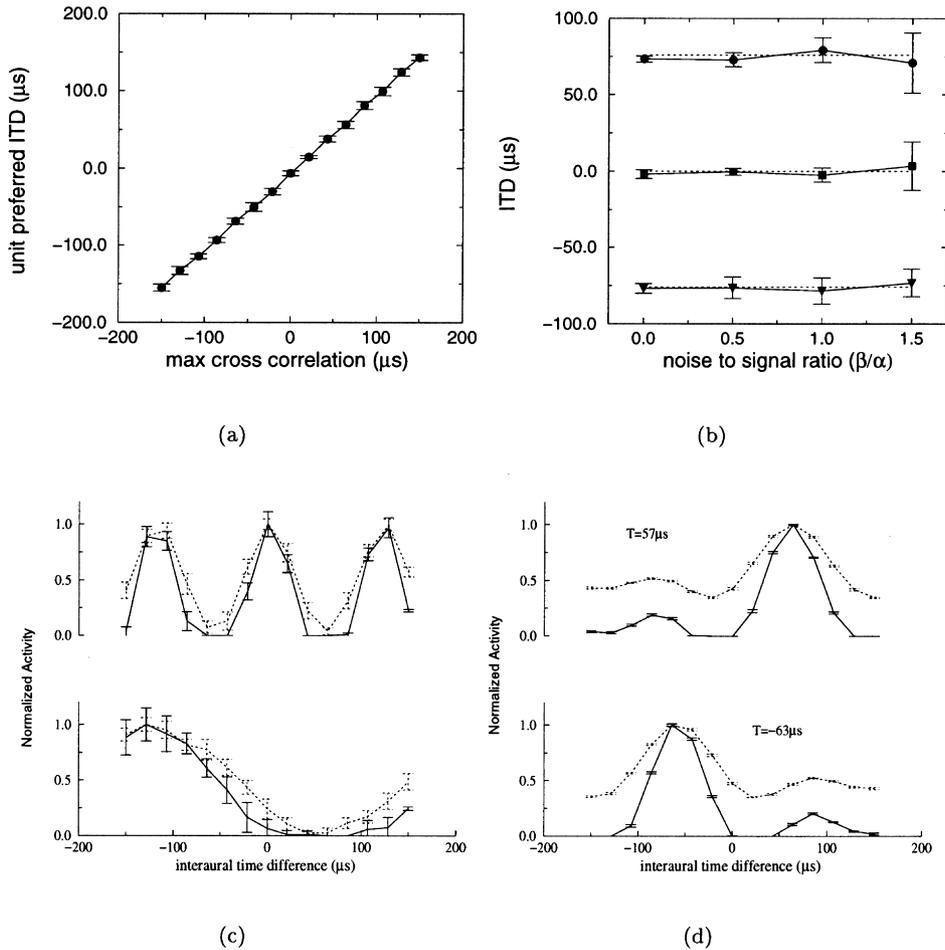


Fig. 8. (Top row) Correspondence between direct cross-correlation of the input signals and ICx activation in the case of a single delay. (a) Analysis for different locations of the sound source. The data are averages of 10 repetitions of 15 different values of ITD, $S/N = 3$. (b) Analysis for different noise-to-signal ratios of the inputs. The measurements for three different positions of the sound source are shown (ITD = $76 \mu\text{s}$ (top), $0 \mu\text{s}$ (middle), $-76 \mu\text{s}$ (bottom)). The dashed lines are the measurements produced by the cross-correlation algorithm, the continuous lines are the results of the model. (Bottom row) Unit activation in the modeled ICc (c) and ICx (d) with (continuous line) and without (dashed line) the action of lateral connections. The graphs show mean values and standard deviation of unit activity of two units in the ICc (c) and in the ICx (d) evaluated on 10 repetitions, when binaural signals with different values of ITD are applied as input.

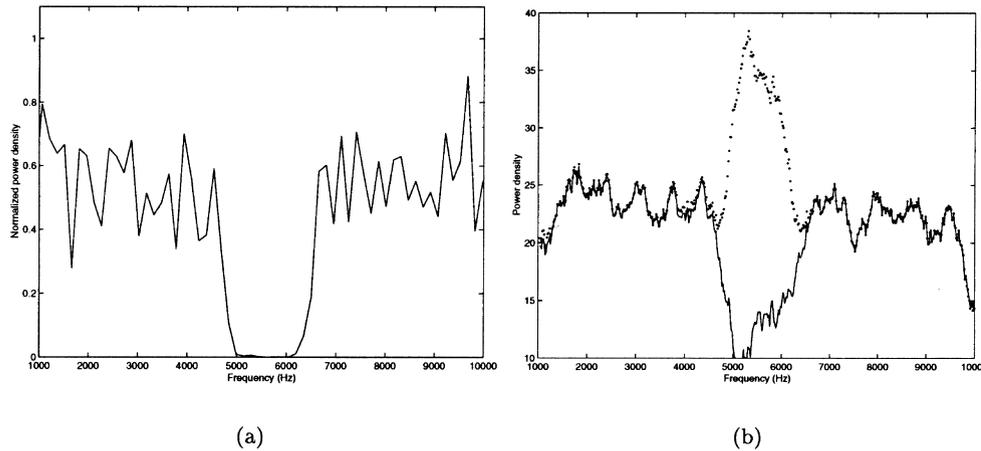


Fig. 9. (a) Spectral power density of the filtered signals in the maximum likelihood method. (b) Spectral power density in the frequency laminae of the ICc with (solid line) and without (dotted line) the action of lateral connections in the ICc. Both in (a) and (b) the inputs were broad band signals with noise superimposed in the range 5–7 kHz and with a single interaural delay.

proportional to the cross-correlation between the input auditory signals as in Eq. (20), even when the units were characterized by a linear activation function. Nevertheless, in the case of a single delay, direct binaural cross-correlation reliably predicted the preferred ITD of the ICx unit with maximum value of activation. As illustrated in Fig. 8a,b, a close correspondence between the delay given by the cross-correlation algorithm and the location of the peak of activity in the ICx was present for different positions of the sound source, even in the presence of noise. Such a correspondence followed from the fact that the effect of inhibition in the model was to sharpen the ITD tuning of ICc and ICx units, without altering the preferred ITDs of these units. This is illustrated in Fig. 8c,d, where the activation of ICc and ICx units for binaural inputs with different delays is plotted both with and without the action of inhibitory connections. As shown by these graphs, the peaks of the tuning functions in the two conditions are in good alignment. As illustrated in Fig. 8d, at the level of the ICx, inhibition also removed the tonic activation of the units which appears

when inhibition is not present. These results closely match experimental observations in the barn owl (Fujita and Konishi, 1991; Albeck and Konishi, 1995).

Lateral connections in the model had beneficial effects for auditory localization both at the level of the ICc and ICx. In the ICc, due to the reciprocal inhibition of areas tuned to different interaural delays and to the local excitation of neurons sensitive to similar delays in different frequency laminae, particularly active ensembles of cells emerged only in adjacent frequency laminae with similar estimates of the interaural phase ϕ_{LR} (see Eq. (15)). In the laminae corresponding to frequencies in which the signal to noise ratio is low, the variability of the estimate of the interaural phase is larger, and the activity of neurons tuned for similar delays at different frequencies are more dissimilar than in frequency bands with higher signal-to-noise ratio. The global result of the interaction in the ICc is a frequency weighting in which the contribution of different frequency bands to the final estimate of the cross-correlation has a relationship of inverse dependence with the variability of the estimate of

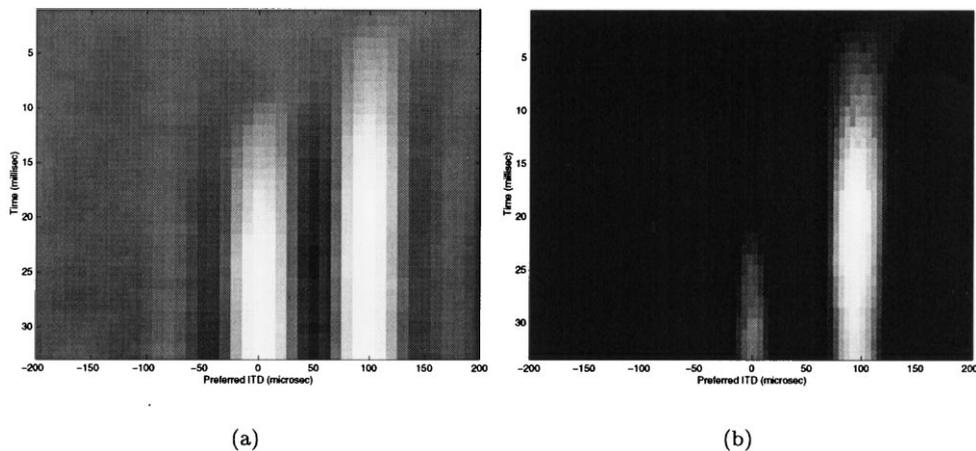


Fig. 10. Echo suppression operated by lateral connectivity in the model. Course of activation in the ICx without (a) and with the action of lateral connections (b) in the presence of two sound sources are activated one after the other. In both figures, the activity of ICx units tuned to different ITDs (horizontal axis) is represented at different times (vertical axis) with an intensity scale. The brighter a point, the more active the corresponding unit at the considered time.

the phase. This is similar to the operation of the maximum likelihood method (see Eq. (14)). Fig. 9 illustrates an example in which input signals had an interaural delay equal to 40 μ s and a large noise contribution ($S/N = 0.5$) superimposed in the frequency band 5–7 kHz. The graphs illustrate how the action of the intrinsic connections in the ICc reduce the spectral power in the frequency band affected by noise in a similar manner to the maximum likelihood method.

An additional effect of lateral connections in the inferior colliculus can be considered in the temporal domain. In the presence of multiple sound sources activated asynchronously, so as to simulate the presence of echos, late occurring peaks in the cross-correlation were suppressed in the modeled ICx, due to the action of inhibitory connections. An example of this phenomenon is illustrated in Fig. 10 in which two signals with different ITDs (0 and 100 μ s) were applied one after the other. The trace of activity in the ICx corresponding to the second source, that is present when lateral connections are ineffective (Fig. 10a), disappears when inhibition is in place (Fig. 10b). This effect is reminiscent of the precedence effect in humans (Blauert, 1983) and is similar to results recently obtained in the barn owl (Keller and Takahashi, 1996b). Such a rejection of echos further improves the reliability of auditory localization with respect to direct binaural cross-correlation.

5. Conclusions

A large amount of experimental evidence suggests that the barn owl localizes the azimuth of a sound source by means of a cross-correlation-like treatment of the signals at the two ears. Space-specific neurons in the inferior colliculus have been reported to be sensitive to the level of binaural cross-correlation, so that the cross-correlation at different time-lags appears to be represented by the activation of units tuned for different ITDs in the ICx map. Two main components seem to be crucial for the achievement of this result: the characteristic response of NL units to ITD and the anatomical pattern of connectivity between ICc and ICx. Evidence has been collected illustrating that the physiological and anatomical characteristics of the NL are compatible with the Jeffres (1948) place theory of sound localization (Konishi et al., 1988, Carr and Konishi, 1990). It is known (Licklider, 1951) that Jeffress' model implements time-delayed comparisons between the various spectral components of the input signals, which provide a basis for evaluating binaural cross-correlation. The projections between the ICc and ICx, with the convergence of multiple frequency laminae in the ICc onto single neurons in the ICx (Wagner et al., 1987) seems to operate so as to combine the results of the spectral comparisons in the NL.

In this paper, by analyzing a simplified model of the ITD pathway, we have shown that these two elements (the response profile of NL neurons and the pattern of connectivity between the ICc and the ICx) give rise to a

correspondence between binaural cross-correlation and pattern of activation in the ICx that can be described analytically. When the effect of lateral connections is eliminated, the average activation of ICx units is equivalent to a smoothed function of the binaural cross-correlation estimated at a time equal to the unit-preferred ITDs. It should be noted that this demonstration relies on a particular activation function of NL units, that explicitly includes the recently observed modulation by signal amplitude, a term not present in previous similar attempts to model the activation of ICx units (Albeck and Konishi, 1995).

The formal correspondence between ICx activity and binaural cross-correlation no longer holds when lateral connections are effective in the modeled inferior colliculus and units are considered to be non-linear. In this case, the activation of ICx units is still linked to binaural cross-correlation, but a more complex relationship than linear proportionality exists. Using computer simulations, we have shown that several concomitant factors contribute to improve auditory localization over direct binaural cross-correlation. Such improvements result in better ITD discrimination in the presence of echos and multiple sound sources, and in a more reliable localization in the presence of high level of noise. An important component is the particular dependence of NL activation on the spectral amplitudes of the auditory inputs. Experimental evidence seems to indicate that the activity of NL cells is modulated by a squashed function of the product of the signal amplitudes, rather than the product of the amplitudes itself (Peña et al., 1996). We have shown that in the model a similar intensity compression in the activation of NL units is responsible for sharpening the peaks of the cross-correlation, thus improving ITD discrimination. However, when this element is considered by itself, the increase in ITD resolution comes at the expense of a greater sensitivity to noise. This follows from the assumption that modulation of NL units is independent of the signal-to-noise ratio at the considered frequencies, an operation similar to the noise-independent amplitude normalization of the phase transform method. In the signal processing literature, other methods, such as the maximum likelihood filter, were designed to overcome this limitation. In the model, other elements at subsequent neural stages operate together with the amplitude modulation so as to reduce the effect of noise. It is not known whether in the barn owl the activity of NL neurons is also characterized by a similar independence to the signal-to-noise ratio, or whether the mechanisms underlying ITD sensitivity of NL units, not modeled in this study, could contribute to a rejection of noise.

Lateral connections in the model further improved auditory localization in several ways. In agreement with what has been observed in the barn owl (Albeck and Konishi, 1995), inhibition at the level of the ICx contributed to discount the net background input from the ICc, and lower secondary peaks of activation in the ICx map. These effects perform a first reduction of the sensitivity of the system with

respect to noise, as it becomes less likely that a saccade is performed toward a location corresponding to a non-primary peak of the binaural cross-correlation. The mutual inhibition of units tuned to different ITDs in the ICc and ICx was also responsible for the suppression of activity in the ICx corresponding to simulated echos, a phenomenon which has also been observed in the barn owl (Keller and Takahashi, 1996b). An important effect mediated by lateral connections in the model was the differential modulation of units tuned for different frequency in the ICc. Given the pattern of intrinsic connectivity in the modeled ICc, by which nearby units are mutually excited and inhibit units tuned to different ITDs, particularly active ensembles of cells emerge only in frequency bands with high signal-to-noise ratio. By means of computer simulations, we have shown that lateral interactions decrease the contribution to the activation of ICx units of frequency bands with high variability in their estimate of the interaural phase difference. In addition, in the bands where noise is low, lateral interactions further reduce the sensitivity of ICc units with respect to the intensity of the input signals. Taken together, the differential modulation of frequency laminae in the ICc and the amplitude sensitivity of NL units operate in the model in a way which is very similar to the maximum likelihood method, the best filter for generalized cross-correlation. Different spectral contributions of the input signals are treated in different ways, depending on the signal-to-noise ratio, so as to reduce the dependence of the output on the frequency bands more heavily affected by noise.

Although the model presented in this paper is a simplification of the real neural pathways in the barn owl, the results of this paper support a close analogy between generalized cross-correlation and the neural processes underlying auditory localization in the barn owl. Generalized cross-correlation is one of the most robust and accurate methods for the estimate of the ITD between the signals at the two ears, and provides good estimates in the presence of high levels of noise and multiple echos. It is intriguing that human ingenuity has discovered a method for sound localization which appears to be analogous to the one implemented by evolution in a species that heavily relies on audition. It will be interesting, as more data are gathered, to apply similar theoretical analysis to other species, and compare them to the barn owl and to other methods for sound localization.

Acknowledgements

This work was carried out as part of the theoretical neurobiology program at The Neurosciences Institute, which is supported by Neurosciences Research Foundation. Support for this program is received in part from the van Ameringen Foundation and Novartis Pharmaceutical Corporation.

References

- Albeck, Y., & Konishi, M. (1995). Responses of neurons in the auditory pathway of the barn owl to partially correlated binaural signals. *J. Neurophysiol.*, *74*, 1689–1700.
- Bendat, J.S., & Piersol, A.G. (1993). *Engineering applications of correlation and spectral analysis*. New York: Wiley.
- Blauert, J. (1983). *Spatial hearing*. Cambridge, MA: MIT Press.
- Bracewell, R.N. (1986). *The Fourier Transform and its applications*. New York: McGraw-Hill.
- Carr, C.E., & Boudreau, R.E. (1993). Organization of the nucleus magnocellularis and the nucleus laminaris in the barn owl: encoding and measuring interaural time differences. *J. Comparative Neurol.*, *334*, 337–355.
- Carr, C.E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.*, *10*, 3227–3246.
- Carter, C.G. (1981). Special issue on time delay estimation. *IEEE Trans. Acoustic Speech Signal Processing*, *ASSP29*, 461–623.
- Fujita, I., & Konishi, M. (1991). The role of GABAergic inhibition in processing of interaural time difference in the owl's auditory system. *J. Neurosci.*, *11*, 722–739.
- Jeffres, L.A. (1948). A place theory of sound localization. *J. Comparative Physiol. Psychol.*, *41*, 35–39.
- Keller, C.H., & Takahashi, T.H. (1996a). Binaural cross-correlation predicts the responses of neurons in the owl's auditory space map under conditions simulating summing localization. *J. Neurosci.*, *16*, 4300–4309.
- Keller, C.H., & Takahashi, T.T. (1996b). Responses to simulated echos by neurons in the barn owl's auditory space map. *J. Comparative Physiol.*, *178*, 499–512.
- Knapp, C.H., & Carter, G.C. (1976). The generalized correlation method for estimation of time delay. *IEEE Trans. Acoustics Speech Signal Processing*, *24*, 320–327.
- Knudsen, E.I. (1983). Subdivision of the inferior colliculus in the barn owl (*Tyto alba*). *J. Comparative Neurol.*, *218*, 174–186.
- Knudsen, E.I. (1984). Synthesis of a neural map of auditory space in the owl. In: Edelman, G.M., Cowan, W.M., & Gall, W.E. (Eds.), *Dynamic aspects of neocortical function* (pp. 375–396). New York: Wiley.
- Knudsen, E.I., & Brainard, M.S. (1995). Creating a unified representation of visual and auditory space in the brain. *Annu. Rev. Neurosci.*, *18*, 19–43.
- Knudsen, E.I., & Konishi, M. (1978a). A neural map of auditory space in the owl. *Science*, *200*, 793–795.
- Knudsen, E.I., & Konishi, M. (1978b). Space and frequency are represented separately in auditory midbrain of the owl. *J. Neurophysiol.*, *41*, 870–884.
- Knudsen, E.I., Konishi, M., & Pettigrew, J.D. (1977). Receptive fields of auditory neurons in the owl. *Science*, *198*, 1278–1280.
- Konishi, M. (1986). Centrally synthesized maps of sensory space. *Trends Neurosci.*, *9*, 163–168.
- Konishi, M. (1993). Listening with two ears. *Sci. Am.*, *268*, 34–41.
- Konishi, M., Takahashi, T.T., Wagner, H., Sullivan, W.E., & Carr, C.E. (1988). Neurophysiological and anatomical substrates of sound localization in the owl. In: Edelman, G.M., Gall, W.E., & Cowan, W.M. (Eds.), *Auditory function: neurobiological bases of hearing* (pp. 721–744). New York: Wiley.
- Licklider, J.C.R. (1951). A duplex theory of pitch perception. *Experientia*, *7*, 128–134.
- Lindemann, W. (1986). Extension of a binaural cross-correlation model by contralateral inhibition. I. simulation of lateralization for stationary signals. *J. Acoustical Soc. Am.*, *80*, 1608–1622.
- Middlebrooks, J.C., & Green, D.M. (1991). Sound localization by human listeners. *Annu. Rev. Psychol.*, *42*, 135–159.
- Moiseff, A., & Konishi, M. (1981). Neural and behavioral sensitivity to binaural time difference in the owl. *J. Neurosci.*, *1*, 40–48.
- Peña, J.L., Viete, S., Albeck, Y., & Konishi, M. (1996). Tolerance to sound

- intensity of binaural coincidence detection in the nucleus laminaris of the owl. *J. Neurosci.*, *16*, 7046–7054.
- Rucci, M., Tononi, G., & Edelman, G.M. (1997). Registration of neural maps through value-dependent learning: modeling the alignment of auditory and visual maps in the barn owl's optic tectum. *J. Neurosci.*, *17*, 334–352.
- Rucci, M., Edelman, G.M., & Wray, J. (1998). Adaptation of orienting behavior: from the barn owl to a robotic system. *IEEE Trans. Robotics Automation* (in press).
- Saberi, K. (1996). An auditory illusion predicted from a weighted cross-correlation model of binaural interaction. *Psychol. Rev.*, *103*, 137–142.
- Sayers, M.M. (1964). Acoustic-image lateralization judgements with binaural tones. *J. Acoustical Soc. Am.*, *36*, 923–926.
- Sayers, B.M., & Cherry, E.C. (1957). Mechanisms of binaural fusion in the hearing of speech. *J. Acoustical Soc. Am.*, *29*, 973–987.
- Stern, R.M., Zeiberg, A.S., & Trahtotis, C. (1988). Lateralization of complex binaural stimuli: a weighted-average model. *J. Acoustical Soc. Am.*, *84*, 156–165.
- Takahashi, T.T., & Konishi, M. (1986). Selectivity for interaural time difference in the owl's midbrain. *J. Neurosci.*, *6*, 3413–3422.
- Takahashi, T.T., & Konishi, M. (1988). Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. *J. Comparative Neurol.*, *274*, 190–211.
- Wagner, H., Takahashi, T.T., & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J. Neurosci.*, *7*, 3105–3116.
- Woodworth, R.S., & Schlosberg, H. (1962). *Experimental psychology*. New York: Holt, Rinehard and Winston.
- Yin, T.C.T., Chan, J.C.K., & Carney, L.H. (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J. Neurophysiol.*, *58*, 562–583.