

A physiologically-based population rate code for interaural time differences (ITDs) predicts bandwidth-dependent lateralization

Kenneth E. Hancock^{1,2}

¹ Eaton-Peabody Laboratory, Massachusetts Eye & Ear Infirmary, Boston MA USA,
Ken_Hancock@meei.harvard.edu

² Department of Otology and Laryngology, Harvard Medical School, Boston MA USA.

1 Introduction

Interaural time difference (ITDs) are the most important cue to the location of sounds containing low-frequency energy (Wightman and Kistler, 1992). ITDs are encoded centrally in the medial (MSO) and lateral (LSO) superior olives which transmit the code to the inferior colliculus (IC) (Batra et al., 1997; Goldberg and Brown, 1969). Each ITD-sensitive neuron is characterized by its best ITD (BD), the one producing maximal discharge rate. It is a longstanding view that these neurons are conceptually arranged in an array with best frequency (BF) on one axis and BD on the other to form a *labeled-line code*. According to this model, the stimulus ITD corresponds to the BD (i.e. the label) of the most active neuron in the array (Jeffress, 1948).

The labeled-line model is challenged by physiological data from guinea pig (confirmed in cat and gerbil) showing that the distribution of BD is highly dependent on BF, and in general does not correspond to the range of naturally-occurring ITDs (Brand et al., 2002; Hancock and Delgutte, 2004; McAlpine et al., 2001). Instead, best interaural *phase* ($BP = BD \times BF$) is more nearly independent of BF, such that the steepest slopes of neural rate-ITD curves tend to occur near the midline. Because the slopes, not the peaks, align near the midline (where perceptual ITD acuity is finest), it has been suggested that ITD is encoded by the discharge rate itself rather than by the locus of maximal activity (McAlpine et al., 2001). Thus, ITD may be represented by a *population rate code*, in which the activity of many neurons pool to form monolithic ITD channels on each side of the brain, and the stimulus ITD may be inferred by comparing the relative activity of the two channels (van Bergeijk, 1962; von Békésy, 1960).

Though the physiological data suggest the existence of a rate code, analysis of its viability has barely begun (Marquardt and McAlpine, 2001). Here, we demonstrate that a population rate code model can account for the dependence of perceived laterality on stimulus bandwidth.

2 Methods

Model neurons are arranged into four arrays, one representing each MSO and LSO (Fig. 1C). Individual neurons are modeled by the cross-correlation operation depicted in Fig. 1A. The sounds at each ear are filtered using identical gammatone filters (center frequency CF and time constant τ). The contralateral filter output is both delayed and phase-shifted (CD and CP , respectively), then multiplied by the ipsilateral filter output. The cross-correlator output is converted to a firing rate by a quadratic function with coefficients A and B .

The single neuron model is thus specified by the six parameters $\{CF, \tau, CD, CP, A, B\}$, whose values were previously constrained by fitting the model to cat IC data (Hancock and Delgutte, 2004). For all model neurons, the coefficients A and B

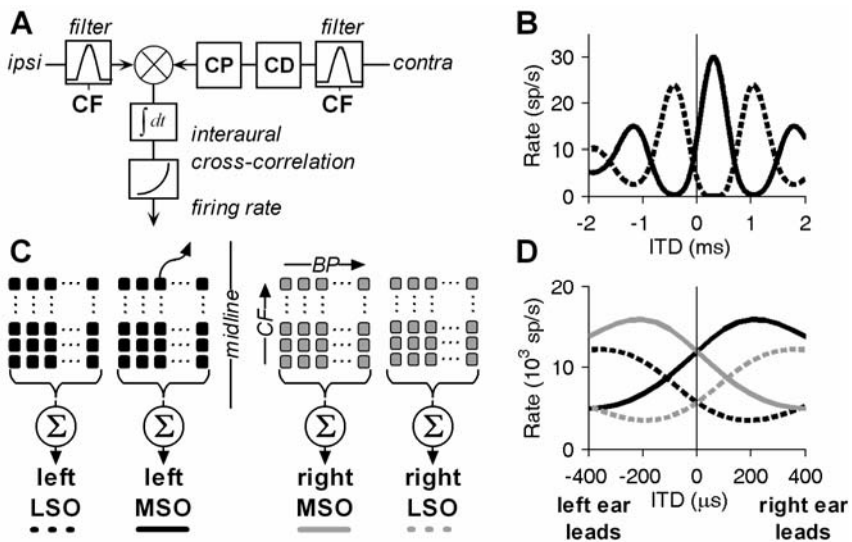


Fig. 1. Population rate model of ITD coding. A) Model for ITD-sensitive neurons. Acoustic inputs are bandpass-filtered, then cross-correlated after applying delay (CD) and phase shift (CP) to one side. Quadratic function converts cross-correlation to neural firing rate. B) Broadband noise rate-ITD curves for two model neurons. *Solid line*, MSO neuron ($CP=0$). *Dashed line*, LSO neuron ($CP=0.5$ cycles). C) Model neurons are grouped into four arrays, representing each MSO and LSO. CF is distributed along one dimension of array, best phase (BP) along the other. Output of each array is the sum of its neural rates. D) Array outputs as a function of ITD in response to broadband noise.

were assigned the mean physiological values. The filter time constant varies inversely with CF according to $\tau = Q/CF$, where $Q = 0.3$. The remaining parameters were assigned as described below.

We have made the simplest possible assumption that MSO and LSO have similar CF and BP distributions, and differ primarily in characteristic phase. The CP was set to zero for all neurons in the MSO arrays, and set to 0.5 cycles in the LSO arrays. The CF parameter was represented along one dimension of each array according to a log-normal distribution with a mean of about 600 Hz, and ranging from 50 Hz to 1500 Hz. The best phase BP was represented along the other dimension following a Gaussian distribution with mean and standard deviation each equal to 0.3 cycles. Each value of BP was used to assign the characteristic delay: $CD = BP/CF$.

Responses of two model neurons to broadband noise varied in ITD are shown in Fig. 1B. For both neurons, $CF = 650$ Hz and $BP = 0.2$ cycles ($CD = 308$ μ s). The *solid line* represents the response of a model MSO neuron ($CP = 0$), and shows a peak firing rate at CD . The *dashed line* represents a model LSO neuron ($CP = 0.5$ cycles), and exhibits a null at CD .

The output of each array is the sum of its individual firing rates. The outputs are illustrated in Fig. 1D as functions of ITD for broadband noise stimulation. Each MSO is maximally active when the stimulus is in the contralateral hemisphere, while each LSO is most strongly activated by ipsilateral stimulation.

3 Results

3.1 Summary of psychophysical results to be modeled

This section summarizes psychophysical data which illustrate a dependence of laterality on stimulus bandwidth (Trahiotis and Stern, 1994), and which represent a nontrivial test of lateralization in the model. The stimulus is bandpass noise centered at 500 Hz, with $ITD=1.5$ ms. Figure 2A shows the pattern of activity produced by this stimulus in the BF-BD plane. The full display consists of a series of peaks and valleys, but for clarity we show only the two peaks closest to the midline (*solid black lines*). The straight contour is at 1.5 ms, corresponding to the ITD of the stimulus. The secondary contour is separated from the main contour by the CF period, and hence is curved.

When the stimulus is narrowband (*dark gray shading*), it is perceived on the left. Trahiotis and Stern (1994) explained this

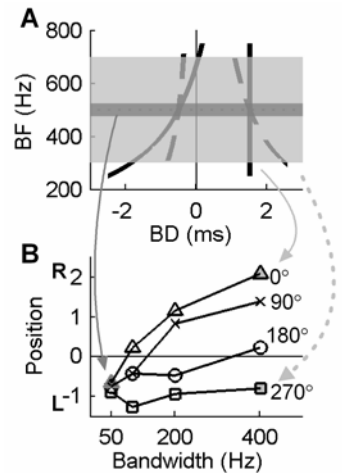


Fig. 2. A) Contours of peak activity in CF-BD plane produced by noise with $ITD=1.5$ ms. *Solid lines*, $IPD=0^\circ$. *Dashed lines*, $IPD=270^\circ$. B) Image heard to the left for narrow bandwidths and/or large IPDs. Image heard to the right for wide bandwidths.

percept as “centrality” dominated, because it favors the secondary peak in the BF-BD plane which occurs nearer the midline. In contrast, broadband stimuli (*light gray shading*) are perceived on the right side. This was described as “straightness” dominated, because it favors the peak for which the ITD is consistent across CF.

Laterality also depends on the stimulus interaural phase difference (IPD). The *dashed lines* in Fig. 2A show the contours corresponding to a 270° phase shift. The ITD was adjusted so that the contours always passed through the center frequency at constant ITD-values. Shifting the phase straightens the left contour and curves the right one. In the broadband condition, this stimulus is perceived on the left because that contour is favored by both straightness and centrality.

3.2 Model

The dependence of lateralization on bandwidth and IPD can be explained using models that incorporate both straightness- and centrality-weighting (Stern et al., 1988). Weighting that reflects the overall tendency for physiological BD values to occur within the naturally-occurring ITD range is the most straightforward realization of centrality (Shackleton et al., 1992; Stern et al., 1988), and is an explicit property of the model described in this paper. One method of implementing straightness-weighting is to integrate across BF along constant values of BD (Shackleton et al., 1992). This fundamentally requires a labeled-line representation because the inputs to the integration stage must be segregated according to BD. We show here that the bandwidth-dependent lateralization data can also be simulated using a simple population rate code model, without explicit straightness-weighting and the resulting need for labeled lines.

Figure 3A shows the output of each of the four channels as a function of bandwidth (for IPD=0°). We consider first a two-channel model comprising only MSO activity, and argue that it is insufficient to predict the psychophysical data. We assume that the position estimate is simply a vector sum of the MSO rates, and note that activity in one MSO corresponds to an image position in the contralateral hemifield. As bandwidth gets larger, the activity in the right MSO (R_{MSO}) decreases (*solid gray line*)

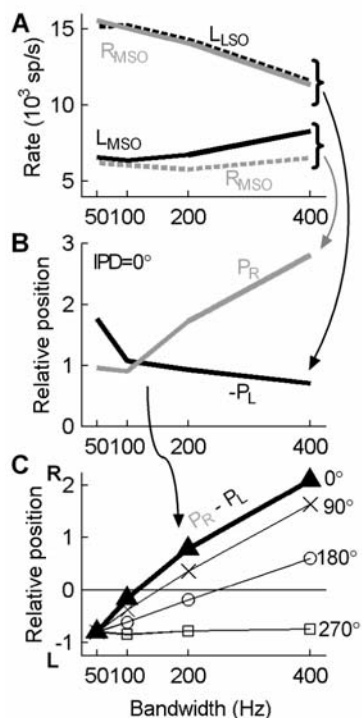


Fig 3. A) Individual channel responses to noise (ITD=1.5ms, IPD=0°) vs. bandwidth. B) Left and right components of model position estimates. C) Model fit to psychophysical data.

while the activity in the left MSO (L_{MSO}) increases (*solid black line*). This correctly predicts that the lateral position moves rightward with increasing bandwidth. But the image can never actually cross to the right of the midline because R_{MSO} is *always* more active than L_{MSO} . This reflects the fact that the secondary peak produced by the stimulus is always more central than the main peak.

A four-channel model incorporating both MSO and LSO, however, can account for both the trends and magnitudes of the psychophysical data. Lateral position estimates were generated by linear combination of the channel outputs:

$$P = \alpha(L_{MSO} - R_{MSO}) + \beta(L_{LSO} - R_{LSO}) \quad (1)$$

The parameters $\alpha = 1.55 \times 10^{-3}$ and $\beta = 1.45 \times 10^{-3}$ were chosen to minimize the sum of squared error between the model position estimates and the psychophysical data. The resulting model fit (Fig. 3C) agrees well with the data (Fig. 2B), for the reasons discussed below.

The effect of incorporating LSO channels into the model is illustrated in Fig. 3B. The position estimates derived from the channel outputs of Fig. 3A are shown decomposed into left and right components:

$$P_R = \alpha L_{MSO} - \beta R_{LSO} \quad (2)$$

$$P_L = \alpha R_{MSO} - \beta L_{LSO}$$

Including LSO channels preserves the essential trend exhibited by the MSO outputs (position increasingly favors the right as bandwidth increases), but shifts the balance such that the position estimate crosses the midline and attains realistic magnitudes. Because the main peak occurs outside the natural range of ITD, it evokes unnaturally large activity in L_{LSO} . In this model, it is heightened L_{LSO} activity, rather than straightness, that trades with the centrality manifested in the R_{MSO} activity to shift the image across the midline.

As IPD increases from 0° to 270° , the main stimulus peak curves away from the array of BDs comprising the L_{MSO} channel. At the same time, the secondary peak straightens, becoming more closely aligned with the R_{MSO} channel. Consequently, R_{MSO} activity increases with respect to L_{MSO} activity, and the image position shifts to the left. As discussed in Sec. 3.1, this is purely a reflection of centrality.

4 Discussion

4.1 Advantages to a code without labeled lines

Neural best ITDs are determined by several factors, including axonal propagation delays, inhibition, and perhaps peripheral mechanical delays due to interaural CF mismatches (Beckius et al., 1999; Brand et al., 2002; Joris et al., 2004). A labeled-line code demands a stable and relatively precise distribution of best ITDs from these combined factors. In contrast, the population rate model requires only a hemifield bias in the best ITD distribution of each channel. The details of the distribution are not necessarily important, especially if the ITD processor is part of

a larger feedback control system that guides orienting movements by restoring balance among the outputs of the four sensory channels.

4.2 A motor interpretation of the population rate code

Rotation of the head about the vertical axis involves four muscles, two on each side of the head. Rotation to the right is accomplished primarily by the sternocleidomastoid (SCM) muscle on the left side (i.e. contralateral to the direction of motion) with assistance from the splenius muscles on the right (ipsilateral to the motion).

The outputs of the population rate model may be suitable for generating motor commands that orient the head toward a sound source (Fig. 4). Specifically, because the MSO is broadly responsive to sounds originating in the contralateral hemisphere, its output is appropriate for driving the muscles which turn the head toward the contralateral hemisphere.

The LSO channels in Fig. 4 inhibit motion toward the ipsilateral side, consistent with the function illustrated by Eq. 2 and Fig. 3. Low-frequency LSO neurons, which presumably comprise the ITD-sensitive component of the LSO, make inhibitory projections to ipsilateral IC (Glendenning and Masterton, 1983).

The broken lines in Fig. 4 emphasize that the auditory brainstem does not have direct control over motor action. Rather, the intention is only to suggest a functional strategy underlying the coding of ITD. It is conceivable, however, that the population rate code described here is an evolutionary remnant of a primitive, more direct coupling between sensory stimulation and motor response. In that context, it is interesting to note that the saccule responds to moderately intense acoustic stimulation and projects to SCM motoneurons by way of the vestibular nucleus (McCue and Guinan, 1997).

4.3 Conclusion

The population rate code and the motor interpretation represent a simple, alternative framework to the conventional labeled-line view of ITD coding. It may prove useful in considering such issues as the development of ITD processing, comparison of ITD processing across species, and binaural hearing in reverberant environments (Devore et al., 2006).

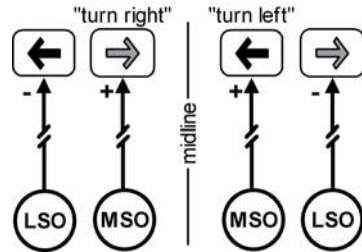


Fig. 4. Motor control suggests a possible strategy underlying the population rate code. Broken lines indicate that the sensory-motor connections are conceptual; no physical connections are implied.

Acknowledgements

This work was supported by NIH grants DC07353 and DC002258.

References

- Batra, R., Kuwada, S., Fitzpatrick, D.C. 1997. Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. *J Neurophysiol* 78, 1222-36.
- Beckius, G.E., Batra, R., Oliver, D.L. 1999. Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci* 19, 3146-61.
- Brand, A., Behrend, O., Marquardt, T., McAlpine, D., Grothe, B. 2002. Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417, 543-7.
- Devore, S., Ihlefeld, A., Shinn-Cunningham, B.G., Delgutte, B. 2006. Neural and behavioral sensitivities to azimuth degrade similarly in reverberant environments, *International Symposium on Hearing*.
- Glendenning, K.K., Masterton, R.B. 1983. Acoustic chiasm: efferent projections of the lateral superior olive. *J Neurosci* 3, 1521-37.
- Goldberg, J.M., Brown, P.B. 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32, 613-36.
- Hancock, K.E., Delgutte, B. 2004. A physiologically based model of interaural time difference discrimination. *J Neurosci* 24, 7110-7.
- Jeffress, L.A. 1948. A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41, 35-39.
- Joris, P.X., van der Heijden, M., Louage, D., Van de Sande, B., Van Kerckhoven, C. 2004. Dependence of binaural and cochlear "best delays" on characteristic frequency. In: Pressnitzer, D., de Cheveigne, A., McAdams, S., Collet, L., (Eds.), *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. Springer-Verlag, New York.
- Marquardt, T., McAlpine, D. 2001. Simulation of binaural unmasking using just four binaural channels, *Assoc. Res. Otolaryngology Abstracts*.
- McAlpine, D., Jiang, D., Palmer, A.R. 2001. A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4, 396-401.
- McCue, M.P., Guinan, J.J., Jr. 1997. Sound-evoked activity in primary afferent neurons of a mammalian vestibular system. *Am J Otol* 18, 355-60.
- Shackleton, T.M., Meddis, R., Hewitt, M.J. 1992. Across frequency integration in a model of lateralization. *J Acoust Soc Am* 91, 2276-2279.
- Stern, R.M., Zeiberg, A.S., Trahiotis, C. 1988. Lateralization of complex binaural stimuli: a weighted-image model. *J Acoust Soc Am* 84, 156-65.
- Trahiotis, C., Stern, R.M. 1994. Across-frequency interaction in lateralization of complex binaural stimuli. *J Acoust Soc Am* 96, 3804-6.
- van Bergeijk, W. 1962. Variation on a theme of von Békésy: a model of binaural interaction. *J. Acoust. Soc. Am.* 34, 1431-1437.
- von Békésy, G. 1960. *Experiments in Hearing*. McGraw-Hill Book Company, Inc., New York. pp. 272-301.
- Wightman, F.L., Kistler, D.J. 1992. The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91, 1648-61.