Sound localization and delay lines – do mammals fit the model?

David McAlpine[1] and Benedikt Grothe[2]

[1]Department of Physiology, University College London, London WC1E 6BT, UK
[2]Max-Planck-Institute of Neurobiology, Martinsried D-82452, Germany

The current dominant model of binaural sound localization proposes that the lateral position of a sound source is determined by the position of maximal activation within an array of binaural coincidence-detector neurons that are tuned to different interaural time differences (ITDs). The tuning of a neuron for an ITD is determined by the difference in axonal conduction delay from each ear – the so-called 'delay line' hypothesis. Although studies in birds appear to support this model, recent evidence from mammals suggests that the model does not provide accurate descriptions of how ITDs are encoded in the mammalian auditory brainstem or of how ITD-sensitive neurons contribute to mammalian sound localization.

Sound localization is a fundamental attribute of the way that animals perceive their environment, enabling them to determine where their prey or a potential mate is, or from where a predator is approaching. A major cue for localizing sounds is the difference in arrival of a sound at the two ears – the interaural time difference (ITD). Since the late 1940s, after Jeffress' seminal paper [1], the dominant model of localization has consisted of an array of coincidence detectors, fed from each ear by a series of delay lines. Crucial aspects of the model are that: (1) neurons signal spatial position by virtue of their peak firing rates, responding maximally when phase-locked excitatory inputs (action potentials whose timing is restricted to a particular phase of the stimulus waveform) arrive in coincidence from each ear; (2) differences in axonal conduction delay from the two ears encode ITDs within the physiological range of delays that would be experienced by the animal (a few hundreds of microseconds), with ITDs largely around zero where spatial acuity is greatest; and (3) all positions in azimuth (the horizontal plane) are represented by different coincidence detectors, with a full representation of azimuthal space encoded in each frequency channel. This theory has formed the basis of several influential computer models (e.g. Ref. [2]) and is sufficiently dominant to be classed as the paradigm. One reason for its dominance, apart from its inherent elegance, has been the seminal investigations into the sound-localization abilities of barn owls by Konishi and colleagues over the past three decades [3–5]. Initiated from an ethological perspective, the

study of barn-owl localization behaviour, and its neurophysiological basis, has become the *de facto* model for understanding sound localization. An integral aspect of this model lies in its synthesis of a neural representation of auditory space in the form of a topological map [3,6] (Fig. 1a).

However, the extent to which this model of binaural hearing extends to species other than the barn owl and, in particular, its validity for mammalian hearing is increasingly a matter of conjecture. Recent electrophysiological evidence from the brainstem and midbrain of small mammals suggests that the means by which neural sensitivity to ITDs is realized in the mammalian brain might be very different to that envisaged by Jeffress, and very different to the way in which models of binaural hearing have been developed by many psychophysicists, physiologists and computational neuroscientists in the ensuing half-century.

Timed inhibition determines neural tuning for ITD

Physiological observations in both barn owls and mammals confirm that action potentials phase-locked to the stimulus waveform converge from each ear onto single neurons in the brainstem to generate interaural-delay sensitivity. The responses of ITD-sensitive neurons at multiple levels of the mammalian auditory pathway are usually described in terms of the output of such a cross-correlation process and are qualitatively similar across a range of avian and mammalian species. However, this does not preclude a role for additional or alternative neural mechanisms for generating ITD sensitivity that account for behavioural and psychological observations. Our latest physiological evidence from single-neuron recordings in the medial superior olivary nucleus (MSO) of the gerbil – a species that, like humans, uses ITDs to localize low-frequency (<1500 Hz [7]) sounds – indicates that temporally precise inhibition plays a crucial, perhaps even singular, role in determining the tuning of binaural neurons for their favoured ITD [8]. It has been known for some time that, in addition to bilateral excitatory innervation, MSO neurons also receive inhibitory (glycinergic) inputs from the brainstem medial and lateral nuclei of the trapezoid body (MNTB and LNTB, respectively) [9–11]. In gerbils, following a period of developmental refinement [12] these glycinergic inputs synapse almost exclusively on the somata of MSO principal neurons. However, their contribution to ITD processing had not previously been investigated. It now transpires

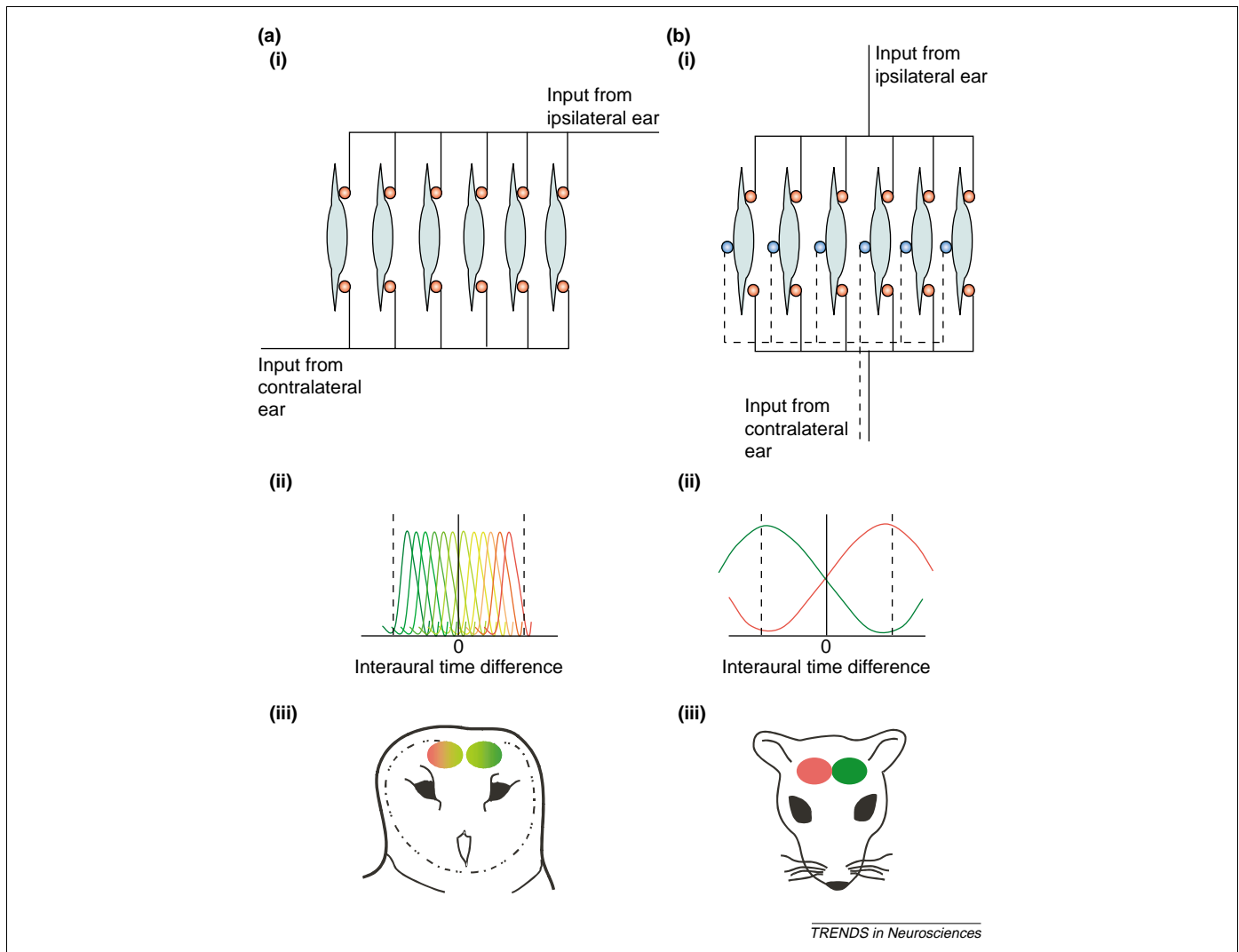*Corresponding author:* David McAlpine (d.mcalpine@ucl.ac.uk).

Fig. 1. Different encoding strategies for interaural time differences (ITDs) in birds and mammals. (a) Fundamentals of the Jeffress model as it might be realized in barn owls. (i) The Jeffress matrix of coincidence detectors and delay lines. Each neuron in an iso-frequency lamina is tuned to a different ITD by virtue of differences in the axonal conduction delay from each ear. (ii) Owing to the relatively high-frequency tuning of binaural neurons in the barn owl, neurons are sharply tuned for ITDs relative to the width of the head (dotted lines). The lateral position of a sound source is read out as the position within the array that is maximally active – a form of local coding. Different ITD tuning of single neurons is indicated by different colours. (iii) Neurons in each brain hemisphere are tuned to different lateral positions in contralateral space. (b) A potential model for ITD-sensitive neurons in mammals. (i) In the absence of inhibitory inputs, axonal conduction delays are distributed around zero ITD. Addition of glycinergic input from the contralateral ear (dotted lines) shifts the peaks of ITD functions towards longer ITDs. (ii) The distribution of peak responses is positioned beyond the physiological range (dotted lines), centred on +45° interaural phase difference with respect to neural tuning for sound frequency. The sensitive slope of the broadly tuned functions is positioned within the physiological range. (iii) The relative activation of the two brain hemispheres could provide a code of lateral position.

that this glycinergic input is more than simply an adjunct to the binaural cross-correlation process presumed to account for ITD sensitivity in the auditory nervous system. Its crucial role in ITD processing was demonstrated most vividly by iontophoresis of the glycine antagonist strychnine through one barrel of a multi-barrel pipette electrode, while simultaneously recording the responses of single MSO neurons to interaurally delayed pure tones [8]. During this inhibitory block, two related effects were observed: (1) discharge rates increased compared with control conditions, in an ITD-specific manner, and (2) the ITD that evoked the peak response shifted from outside to inside the physiological range of the gerbil, to peak at, or very close to, zero ITD. The implication from this experiment is clear: inhibition by glycine determines the tuning of a neuron for ITD because without this inhibition the axonal conduction delay is effectively zero. A computer

simulation using an established MSO model [13,14] with internal delay set to zero and a fast inhibitory conductance verified the contribution of inhibition to the ITD tuning, indicating that the peak ITD shifts to longer and longer values as the magnitude of the inhibitory conductance is increased [8]. Note that the inhibition itself, although phase-locked, need not be ITD-sensitive. Rather, its specificity is a result of its timing relative to the excitatory inputs, possibly related to the restriction of inhibitory synapses to the cell soma [12].

A restricted range of ITD detectors exists in the mammalian brain

What, then, is the consequence for the current dominant model of binaural hearing of this apparently crucial role for glycine-mediated inhibition in shaping ITD sensitivity? *In extremis*, it suggests that the systematic arrangement of

axonal conduction delays that has been demonstrated to exist in the barn-owl brainstem does not exist in the mammalian brainstem. Evidence for delay lines in mammals rests largely on the reported ITD sensitivity of binaural neurons; anatomical evidence for a systematic arrangement of axonal delay lines is equivocal [15]. Therefore, it could be contested that, although glycine-mediated inhibition is important in tuning neurons for their peak ITD, it is of relatively little consequence to how we understand the operation of binaural hearing in mammals. If timed inhibition by glycine creates the appearance of delay lines other than by means of axonal conduction delays, then the basic tenet of the Jeffress model still holds (i.e. individual neural elements encode the lateral position of a sound source by virtue of being tuned to a particular ITD). However, we are dissuaded from this more conservative view by the second major outcome of our most recent study. Over much of the relevant frequency range for mammalian ITD sensitivity (< 1500 Hz), MSO neurons were tuned to ITDs well beyond the range that the head of a gerbil can create under natural listening conditions. In fact, the majority of neurons responded maximally at ITDs roughly equivalent to an interaural phase difference (IPD) of 45° with respect to their tuning for sound frequency [8], consistent with observations made from a much larger population of neurons in the guinea-pig midbrain nucleus of the inferior colliculus [16]. This relatively restricted range of peak ITDs in each frequency channel suggests that the Jeffress matrix of internal delays is not realized in the mammalian brain. Even in humans, with their larger interaural distance, 45° IPD corresponds to ITDs close to, or beyond, the physiological range for a significant proportion of the ITD-sensitive frequency range.

What value might there be in positioning the peaks of ITD functions at such long ITDs? A likely reason concerns the underlying response statistics that determine the ability of binaural neurons to signal small differences in ITD. Similar to neurons coding other sensory modalities and sub-modalities, ITD-sensitive neurons show greatest resolution where the variance, relative to the mean firing rate, is relatively low. This corresponds to a position on the slope of the ITD function [17,18] and it therefore makes sense to position these slopes where greatest acuity is required and, indeed, observed – within the physiological range of ITDs and close to the midline. Although this does not of itself necessitate that the peaks of ITD functions be positioned beyond the physiological range, a single crucial factor makes this more likely in mammals than in the barn owl, the species on which much of the support for the Jeffress model is based. This factor is the range of sound frequencies over which ITDs are utilized, and it is well-established, although less well promulgated, that barn owls differ markedly from mammals, and indeed from other birds, in this respect. Owls can utilize ITDs for localizing high-frequency carriers, of up to 9 kHz [4,19], far above the range at which the temporal information required for binaural hearing is lost in mammals [20,21]. This has important implications for ITD processing because a necessary outcome of the binaural cross-correlation process is that the ITD-tuning sharpness of individual

binaural neurons is related to their tuning for sound frequency – the primary feature encoded by the auditory sensory epithelium. Because barn-owl ITD sensitivity operates over a much higher frequency range than that of mammals (many studies report ITD sensitivity for neurons with frequencies of 5–7 kHz [4]), ITD tuning of single neurons in barn owls is inherently sharper than ITD tuning of single neurons in mammals, where relatively few neurons tuned to frequencies > 1500 Hz are ITD-sensitive [22]. It is generally assumed, therefore, that the exquisite sensitivity of the barn owl to spatial cues is provided by populations of sharply tuned ITD-sensitive neurons within each (high) frequency channel and that these signal sound-source position by virtue of their peak firing rates – a local coding strategy (Fig. 1aii).

Timed inhibition enhances sensitivity to naturally occurring ITDs

In contrast to high-frequency (> 3 kHz) tuning of most ITD-sensitive neurons recorded in barn owls, the low-frequency tuning of ITD-sensitive neurons that is observed in mammals can provide only a very coarse representation of ITD, based on their peak responses. For a hypothetical low-frequency neuron tuned to an ITD within the physiological range, a near-maximal response could be elicited by ITDs spanning the entire width of the head, depending on the exact frequency and interaural distance. Because resolution of ITDs by the peaks of low-frequency neurons is likely to be poor, it makes sense that mammals should position the most sensitive region of their ITD functions – the slopes – within the physiological range to perform accurate sound localization (Fig. 1bii). To achieve this, gerbil MSO neurons employ a fast glycine-mediated inhibitory conductance, shifting the peak of many ITD functions out of the physiological range in the process. Notably, under conditions of strychnine block in the MSO, the proportion of the ITD-modulated spike output that lay within the physiological range of ITDs fell from ~ 80% to ~ 20%. Furthermore, the response was rendered non-monotonic within this range and, therefore, ambiguous for ITD [8]. Although this suggests a difference in the means by which mammals and barn owls utilize the output of binaural neurons to determine the lateral position of a sound source, there is no reason *a priori* why barn owls might not also use a coding strategy based on the slopes of ITD functions, the peaks of which reside within the physiological range. It is unlikely that the response statistics of binaural neurons in the barn owl are significantly different to those in other species or in other sensory modalities.

One might imagine, then, that in the absence of precisely timed inhibition by glycine, the distribution of internal delays in the mammalian MSO is roughly normally distributed around zero ITD, reflecting essentially random differences in axonal path length from the two ears. By adding glycine-mediated inhibition and adjusting the magnitude of its conductance, the position of the peak response, or more probably the position of maximum slope of the response, can be shifted under dynamic control. Fine-tuning of delay sensitivity during development could take the form of adjusting the relative

synaptic weights of excitation and inhibition, rather than pruning (or lengthening) axon collaterals.

A population code for sound localization in mammals

Because only a relatively restricted range of peak tuning for ITD is observed within each frequency channel, often beyond the limits of the physiological range, mammalian sound localization must be achieved by some means other than the local-coding strategy suggested by the Jeffress model and, apparently, adopted by the barn owl. One possibility is that mammals localize sounds by means of a population code, in which the lateral position of a sound source is determined by the relative activation of just two broadly tuned binaural channels, one in each brain hemisphere [16], beginning in the left and the right MSO (Fig. 1b). Examples of population codes based on a small number of rather broadly tuned receptors are abundant in nature. One well-known example is mammalian colour vision, in which only three discrete channels form a continuous colour space [23]. The possibility that interaural information is encoded by two such orthogonal (separated by 90° in phase space) channels by means of the non-redundant coding of the Fourier Transform is the focus of ongoing work in our laboratories, as is the more difficult task of designing psychophysical experiments capable of discriminating between the local-coding strategy currently favoured by the Jeffress model and the population-coding strategy suggested by the restricted range of internal delays observed physiologically.

In summary, we suggest that the Jeffress model, originally developed to account for human localization abilities, accounts for neither the generation of tuning for ITD in mammals nor the contribution of ITD-sensitive neurons to mammalian sound localization. We suggest a new model of mammalian binaural hearing, one in which neural tuning for ITD is determined by a fast inhibitory conductance rather than by a systematic arrangement of delay lines, and in which the lateral position of a sound source is determined by the relative activity within two populations of neurons that are broadly tuned for ITD, rather than by the peak responses of individual neurons that are sharply tuned to ITDs within the physiological range.

References
1 Jeffress, L.A. (1948) A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41, 35–39
2 Stern, R.M. *et al.* (1988) Lateralisation of complex binaural stimuli: a weighted image model. *J. Acoust. Soc. Am.* 84, 156–165
3 Konishi, M. (1973) How the owl tracks its prey. *Am. Sci.* 61, 414–429
4 Carr, C.E. and Konishi, M. (1990) A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.* 10, 3227–3246
5 Saberi, K. *et al.* (1999) Neural bases of an auditory illusion and its elimination in owls. *Nat. Neurosci.* 2, 656–659
6 Takahashi, T.T. and Keller, C.H. (1994) Representation of multiple sound sources in the owl's auditory space map. *J. Neurosci.* 14, 4780–4793
7 Heffner, R.S. and Heffner, H.E. (1988) Sound localization and use of binaural cues by the gerbil (*Meriones unguiculatus*). *Behav. Neurosci.* 102, 422–428
8 Brand, A. *et al.* (2002) Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417, 543–547
9 Cant, N.B. and Hyson, R.L. (1992) Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil. *Hear. Res.* 58, 26–34
10 Kuwabara, N. and Zook, J.M. (1992) Projections to the medial superior olive from the medial and lateral nuclei of the trapezoid body in rodents and bats. *J. Comp. Neurol.* 324, 522–538
11 Grothe, B. and Sanes, D.H. (1993) Bilateral inhibition by glycinergic afferents in the medial superior olive. *J. Neurophysiol.* 69, 1192–1196
12 Kapfer, C. *et al.* (2002) Experience-dependent refinement of inhibitory inputs to auditory coincidence-detector neurons. *Nat. Neurosci.* 5, 247–253
13 Carney, L.H. (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. *J. Acoust. Soc. Am.* 93, 401–417
14 Brughera, A.R. *et al.* (1996) A model with excitation and inhibition for cells in the medial superior olive. *Aud. Neurosci.* 2, 219–233
15 Smith, P.H. *et al.* (1993) Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J. Comp. Neurol.* 331, 245–260
16 McAlpine, D. *et al.* (2001) A neural code for low-frequency sound localization in mammals. *Nat. Neurosci.* 4, 396–401
17 Skottun, B.C. (1998) Sound localization and neurons. *Nature* 393, 531
18 Skottun, B.C. *et al.* (2001) The ability of inferior colliculus neurons to signal differences in interaural delay. *Proc. Natl. Acad. Sci. U. S. A.* 98, 14050–14054
19 Knudsen, E.I. and Konishi, M. (1978) Center-surround organization of auditory receptive fields in the owl. *Science* 202, 778–780
20 Kettner, R.E. *et al.* (1985) Postnatal development of the phase-locked response to low frequency tones of auditory nerve fibers in the cat. *J. Neurosci.* 5, 275–283
21 Palmer, A.R. and Russell, I.J. (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear. Res.* 24, 1–15
22 McAlpine, D. *et al.* (1996) Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. *Hear. Res.* 97, 136–152
23 Daw, N.W. (1973) Neurophysiology of colour vision. *Physiol. Rev.* 53, 571–611

Letters to *TINS*

If you wish to comment on any article published in *TINS*, or would like to discuss issues of broad interest to neuroscientists, then why not write a **Letter to the Editor?**

Letters should be up to 700 words. Letters are often sent to the author of the original article for their response, in which case both the letter and reply will be published together.

Please note: submission does not guarantee publication.