

# The Calyx of Held: A Hypothesis on the Need for Reliable Timing in an Intensity-Difference Encoder

Philip X. Joris<sup>1,\*</sup> and Laurence O. Trussell<sup>2</sup>

<sup>1</sup>Laboratory of Auditory Neurophysiology, Department of Neurosciences, University of Leuven, Leuven B-3000, Belgium

<sup>2</sup>Oregon Hearing Research Center and Vollum Institute, Oregon Health and Science University, Portland, OR 97239, USA

\*Correspondence: [philip.joris@kuleuven.be](mailto:philip.joris@kuleuven.be)

<https://doi.org/10.1016/j.neuron.2018.10.026>

The calyx of Held is the preeminent model for the study of synaptic function in the mammalian CNS. Despite much work on the synapse and associated circuit, its role in hearing remains enigmatic. We propose that the calyx is one of the key adaptations that enables an animal to lateralize transient sounds. The calyx is part of a binaural circuit that is biased toward high sound frequencies and is sensitive to intensity differences between the ears. This circuit also shows marked sensitivity to interaural time differences, but only for brief sound transients (“clicks”). In a natural environment, such transients are rare except as adventitious sounds generated by other animals moving at close range. We argue that the calyx, and associated temporal specializations, evolved to enable spatial localization of sound transients, through a neural code congruent with the circuit’s sensitivity to interaural intensity differences, thereby conferring a key benefit to survival.

## Introduction

The calyx of Held is a giant axon terminal on neurons in the medial nucleus of the trapezoid body (MNTB). This terminal has played a pivotal role in the identification (Guillery, 2005) and understanding of the structure and function of chemical CNS synapses (Borst and Soria van Hoeve, 2012; Schneggenburger and Forsythe, 2006). Compelling images of light and EM reconstructions of the hand-like calyx holding the soma of an MNTB neuron have adorned many journal covers in the past decades (Figure 1). Clearly, strong selection pressure must have led to the evolution of this morphologically and physiologically specialized structure. Ironically, the nature of that selection pressure is not understood.

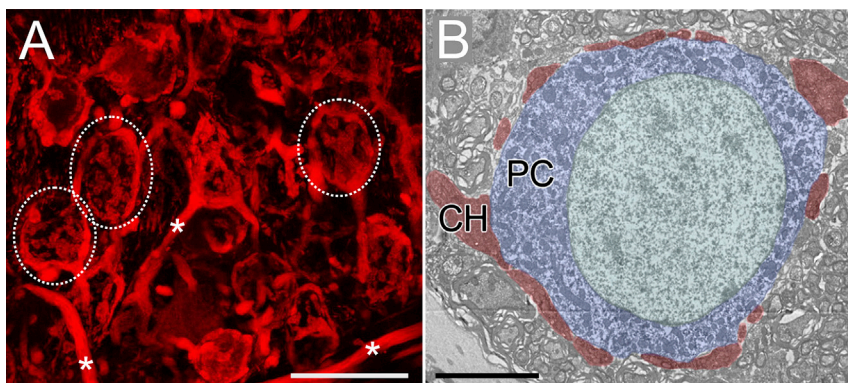
We first briefly review the connectivity and main features of the circuit in which the calyx is embedded and highlight salient properties of the synapse revealed by *in vitro* studies, with a focus on the “timing” aspects at high frequencies that have been stressed in the literature. After reviewing behavioral and physiological binaural sensitivity, we argue that the temporal properties of this circuit are unnecessary and even ill-suited for the functional role traditionally assigned to it—the creation of sensitivity to interaural intensity differences. A brief review of temporal coding in the monaural afferents and of the sensitivity to interaural time differences in this circuit brings us to the one form of binaural sensitivity in which it excels: sharp tuning to interaural time differences of high-frequency, transient sounds. Finally, consideration of natural sources of high-frequency sounds and their acoustic propagation leads to our hypothesis: the calyx of Held and associated circuit specializations enable spatial localization of sound transients caused by movement of a nearby animal. We end by briefly exploring alternative hypotheses.

## General Circuit Structure

The calyx is a giant glutamatergic terminal formed by the main axon of globular bushy cells (Figure 2; GBCs). These cells have their cell body in the cochlear nucleus contralateral to the MNTB, and receive large axosomatic terminals from the auditory nerve (endbulbs of Held). GBCs have limited dendritic trees, have a large-diameter myelinated axon, and project to several other targets besides the MNTB.

The calyx innervates MNTB principal cells: these are glycinergic and project to several targets in the auditory brainstem. The main target is a binaural nucleus: the lateral superior olive (LSO), on which we focus here as it is the target most likely to reveal the function of the calyx. Other targets, not shown in Figure 2 but briefly discussed in a later section, include both binaural and monaural structures (Banks and Smith, 1992; Smith et al., 1998; Sommer et al., 1993; Spangler et al., 1985), including the medial superior olive (MSO), which is specialized for the processing of interaural time differences. Together, the MNTB, MSO, and LSO form the main nuclei of the superior olivary complex.

While the general connectivity of the LSO is well known, many details still need to be filled in. The contralateral input seems to be largely dominated by the MNTB. The ipsilateral pathway to the LSO is less well characterized than the contralateral pathway and is simplified in Figure 2, which only shows the input from spherical bushy cells (SBCs) and omits smaller sources (Cant and Casseday, 1986; Doucet and Ryugo, 2003; Gómez-Álvarez and Saldaña, 2016). Also, morphological and *in vitro* physiological studies show that the LSO is not homogeneous: it contains several cell classes (Helfert and Schwartz, 1986; Rietzel and Friauf, 1998).



**Figure 1. The Giant Calyx of Held**

(A) Calyces of Held (dashed outlines) and large preterminal axons (asterisks) labeled in an Atoh1-cre  $\times$  Ai9 mouse, expressing the red fluorophore tdTomato in GBCs and their axons. Image from G. Romero.

(B) Electron micrograph of calyx terminal (red) surrounding MNTB soma (blue). Image from Hruskova et al. (2012). Scale bars, 20  $\mu$ m for (A) and 5  $\mu$ m for (B).

### Synaptic Physiology of the Calyx

Many features of the calyx of Held-MNTB synapse have been described as specializations for auditory processing. Here we will briefly review three aspects of this synaptic connection: its morphology, presynaptic physiology, and postsynaptic intrinsic properties. The MNTB functions as a sign-inverting relay, where “relay” implies reliability. An enemy of reliable response transmission is the time required for recovery after transmission, specifically the times required to recover from spike refractory period and synaptic depression. Such biologically imposed recovery periods not only limit response rates, but also impact a system’s ability to respond consistently to brief but unpredictable stimuli. Therefore, along with speed of conduction through the relay synapse, the MNTB has also minimized the need for, and duration of, such recovery times.

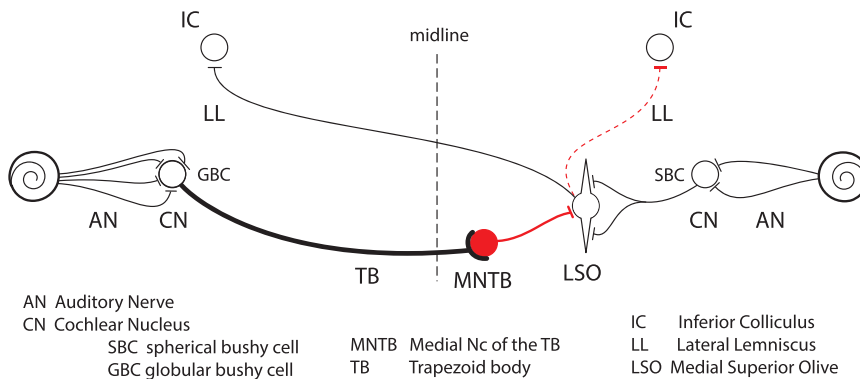
The first and most obvious characteristic of the calyx is the size of the terminal and pre-terminal axon. By using one of the largest diameter, and therefore fastest conducting, axons in the CNS (Kuwabara et al., 1991; Spirou et al., 1990), the GBC-to-MNTB pathway transmits signals across the brainstem as quickly as possible. The MNTB principal cell, although possessing a small dendrite, restricts the excitatory input from the GBC almost entirely to the cell body (Smith et al., 1991, 1998). The somatic synaptic location and large size of the EPSP minimize the time required for the excitatory postsynaptic current to charge the membrane potential of the axon to spike threshold. The added time for the synaptic delay intrinsic to all chemical synapses apparently is brief and reproducible enough not to detract from the overall speed of the sign inverting relay (Joris, 1996; Joris and Yin, 1998). Although the calyx “grips” the postsynaptic cell, the terminal structure is sufficiently fenestrated in adult animals to permit escape of transmitter during high-frequency transmission (Ford et al., 2015). While a fast, electrical synaptic contact using gap junctions might be expected to be faster than the calyceal chemical synapse, such an arrangement would likely present the problem of impedance mismatch between presynaptic axon terminal and the postsynaptic soma, possibly slowing transmission or introducing jitter. Thus, the morphology of the calyx makes it readily apparent that its function is to minimize delays in sign inversion and relay of signals from the contralateral ear.

The physiology of transmitter release from the calyx supports this conclusion. In order to function as a relay, the excitatory

postsynaptic potential (EPSP) must be large enough for the principal cell membrane potential to rapidly reach spike threshold, and must do so reliably upon arrival of each presynaptic action potential.

This is ensured by the wedding of several features. The calyx harbors many hundreds of transmitter release sites. For example, in rats, roughly 600 release sites (Meyer et al., 2001; Sätzler et al., 2002) are supplied with a readily releasable pool of up to several thousand vesicles (Schneggenburger et al., 2002). These release sites feature a tight, nanodomain coupling between the vesicles and the calcium channels that trigger their release, thus facilitating rapid transmission (Stanley, 2016). However, the reason for having so many release sites is not only to summate hundreds of individual release events, but to ensure that depletion of transmitter does not lead to synaptic failure. Rather than having a large release probability, each release site has a small chance of exocytosis during each presynaptic spike, releasing only about 20 vesicles (de Lange et al., 2003; Lorteije et al., 2009). Thus, from trial-to-trial, release sites that lost a vesicle may have time to recover their vesicle content and be ready to respond again. Initial *in vitro* studies of transmission in the MNTB highlighted profound synaptic depression (Borst et al., 1995; Forsythe et al., 1998). This depression was all the more puzzling given that GBC fibers are spontaneously active and could therefore lead to a substantial level of depression even prior to acoustically evoked activity (Hermann et al., 2007). However subsequent work revealed that under more “*in vivo*-like” conditions, including physiological  $\text{Ca}^{2+}$  levels, temperature, and possibly enhanced negative feedback mechanisms, the calyx is able to sustain transmission under high rates of activity (Lorteije et al., 2009). Presumably, reduced  $\text{Ca}^{2+}$  and elevated temperatures lower release probability and enhance the rate of vesicle replenishment sufficiently to support ongoing transmission. Moreover, within this population of vesicles, it is likely that release probabilities vary widely (reviewed in Sakaba, 2018). Such heterogeneity may be essential to sustain transmission to different patterns and durations of sensory signal.

Presynaptic ion channels allow consistency in the presynaptic spike trigger for exocytosis, regardless of spike interval. Voltage-gated  $\text{K}^+$  (Kv) channels quickly repolarize spikes (Ishikawa et al., 2003; Nakamura and Takahashi, 2007), thus limiting  $\text{Na}^+$  channel inactivation (Leão et al., 2005). Recovery from inactivation is also quite rapid, and contributes to an amazingly stable spike height across different stimulus rates (Sierksma and Borst, 2017). An additional factor is the interaction of spike-triggered ionic conductances with the conductances that set the resting potential



**Figure 2. Simplified Rendition of the LSO Circuit**

The trapezoid body (TB) and lateral lemniscus (LL) are fiber bundles containing some of the relevant axons. Besides the calyx of Held on cell bodies of MNTB neurons, the circuit also features large axosomatic terminals on spherical bushy cells (SBCs) and globular bushy cells (GBCs) of the cochlear nucleus (CN), referred to as the endbulbs and modified endbulbs of Held. Red color indicates inhibitory (glycinergic) projection. The LSO contains several cell types: the projection to the contralateral inferior colliculus (IC) is excitatory while that to the ipsilateral IC is mixed excitatory and inhibitory (dashed red line, for ease shown as if originating from same neuron).

(Huang and Trussell, 2008, 2011, 2014; Kim and von Gersdorff, 2012), an interaction that results in the membrane potential before and after the presynaptic spike being almost identical (Sierksma and Borst, 2017). As a result, the electrical starting point for each spike is nearly independent of its history.

There also exists a balance of ion channels that ensure the responsiveness of the postsynaptic cell to calyceal activity. Subunits of the AMPAR glutamate receptors are optimized to ensure rapid channel gating (Geiger et al., 1995; Yang et al., 2011). In this way, the lifespan of the synaptic current does not limit the duration of the EPSP, particularly during repetitive activity. Several types of Kv channel also conspire to shape the postsynaptic response. Activation of Kv2 and Kv3 channels rapidly repolarizes the action potential (Brew and Forsythe, 1995; Dodson et al., 2002; Johnston et al., 2008; Wang et al., 1998). Kv1 channels, by contrast, are active at modest depolarizations from rest and have two functions: first to help the neuron refrain from spiking more than once during an EPSP, and second to allow the cell to recover quickly from an EPSP-spike sequence, so that it can again respond in case a second EPSP is triggered soon after. While Kv3 channels are associated with high-frequency firing, other neurons in the auditory system that do not fire at short latency, encode auditory timing, or have calyceal synapses also express these channels and can also fire at high rates (Perney and Kaczmarek, 1997; Rusznák et al., 2008). Thus, high firing rate alone is not a unique hallmark of the MNTB.

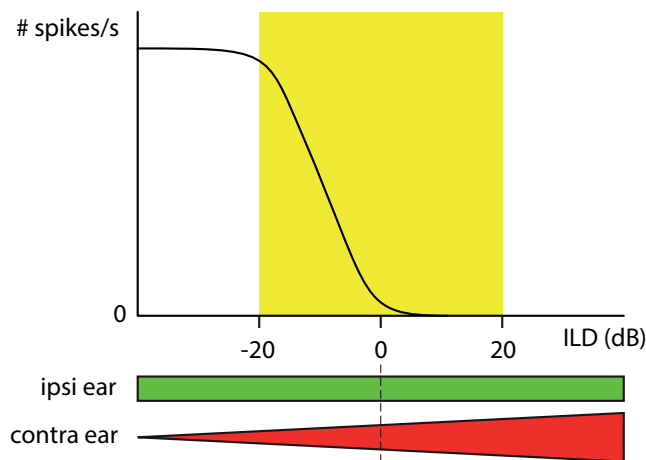
It is of interest to compare Kv channel expression in MNTB to that of neurons of the MSO and cochlear nucleus (in the latter, bushy cells and octopus cells). In those neurons, the expression of the Kv1 population, and the resulting electrical “leakiness” of the cell membrane, is greater than in MNTB (Bal and Oertel, 2001; Cao et al., 2007; Scott et al., 2005). Indeed, MSO, bushy, and octopus cells are so leaky that even the somatically recorded action potential is severely attenuated. This high degree of Kv channel expression is matched by high expression levels of HCN channels as well, and together these channels ensure that synaptic integration by spatial summation can occur successfully only over a narrow time window (Golding and Oertel, 2012). For the monoinnervated MNTB, such synaptic integration does not occur, and thus there is less need for such brief time membrane time constants. Instead, the level of expression of Kv channels need only be sufficient to limit the number of

postsynaptic spikes generated per presynaptic spike (Klug and Trussell, 2006; Scott et al., 2005). This spike limiting mechanism, combined with the short synaptic latencies and reliability of the EPSP-spike sequence, allows the MNTB to function reliably.

The monoinnervation and huge size of the calyx suggest a view that the MNTB is a simple sign inverter, delivering inhibitory copies of GBC spiketrains to the LSO. Although there is some debate, and likely some species dependence, regarding the faithfulness of synaptic transmission in the MNTB (Englitz et al., 2009; Kopp-Scheinflug et al., 2003; Lorteije et al., 2009; Mc Laughlin et al., 2008), there is perhaps no other neuron in the brain with a spike output that so faithfully resembles that of its input, but with an inversion of sign. However, under some stimulus conditions, particularly when high, sustained firing rates are evoked by electrical pulse trains (Guinan and Li, 1990), transmission failures can occur.

### Behavioral and Physiological Sensitivity to Interaural Intensity Differences

The functional role classically associated with the LSO circuit is sensitivity to interaural intensity differences, described in all major neuroscience textbooks. The importance of interaural intensity differences for spatial hearing has been known for more than a century and goes back to early theoretical and experimental work such as that of Lord Rayleigh (Strutt, 1907), who reasoned from physical considerations that this binaural cue would be particularly relevant for the horizontal (azimuthal) localization of high-frequency sounds. Due to their small wavelength relative to the dimensions of external ears, head, and torso, high-frequency sounds are “shaded” by these structures so that sound intensity is larger at the eardrum closest to the sound source. In contrast, the other main binaural cue, interaural time difference, was thought to only operate at low frequencies. The spatial percept of a pure tone depends on the interaural time difference, but only if the tone is below  $\sim 1,500$  Hz. These early considerations and observations on the perception of pure tones resulted in the duplex theory, which holds that azimuthal localization depends on interaural time differences at low frequencies and interaural intensity differences at high frequencies. Importantly, the duplex theory breaks down for more natural sound stimuli, environments, and listening situations. Of particular relevance here is that there is also sensitivity to interaural time differences at high frequencies.



**Figure 3. Schematic of the Sensitivity to Interaural Intensity Difference in LSO Neurons**

The function is obtained by holding the sound intensity constant at the ipsilateral, excitatory ear (green) and increasing the sound intensity at the contralateral, inhibitory ear (red). Equal intensity occurs at interaural intensity difference (IID) = 0 dB; at positive values the contralateral (inhibitory) ear is more intense than the ipsilateral (excitatory) ear. The yellow rectangle roughly indicates the range of maximal interaural intensity differences commonly found in acoustic measurements at the two ears.

Human sensitivity to interaural intensity differences has been studied mostly via headphones, which allow easy presentation of stimuli that are identical in all respects except their intensity. The smallest detectable differences are typically somewhat below 1 dB, and the behavioral evaluation of this cue can be reasonably well approximated as a “level meter” that calculates and compares the sound power at both ears averaged over a time window with a duration of a few hundred ms (Hartmann and Constan, 2002).

Physiologically, the neural computation underlying the extraction of interaural intensity differences seems exceedingly simple and follows from the “IE” property (inhibited by contra and excited by ipsi) of LSO neurons. If a sound is closer in space to the excitatory (ipsilateral) ear than the inhibitory (contralateral) ear, the sound intensity is higher at the excitatory than at the inhibitory ear. This is usually referred to as a negative interaural intensity difference, and results in net excitation of LSO neurons (Figure 3). By contrast, for a positive interaural intensity difference, i.e., if the sound is closer to the inhibitory (contralateral) than the excitatory (ipsilateral) ear, there is net inhibition of LSO neurons. This simple push-pull operation only requires that the monaural inputs from the two ears are opposite in sign, that they have a monotonic relationship between sound intensity and spike rate, and that they are similar in the temporal pattern of their response. Indeed, LSO neurons show a rather stereotyped sigmoidal relationship between firing rate and interaural intensity difference, first documented in classical studies (Boudreau and Tsuchitani, 1968; Tsuchitani and Boudreau, 1969) and sketched in Figure 3.

The MNTB to LSO projection shows a developmentally refined tonotopy (Kim and Kandler, 2003). Studies of connectivity and physiology indicate that MNTB, LSO, and their afferents from cochlear nucleus are biased toward the representation of high

frequencies (“high” is defined relative to the limit of phase-locking to pure tones; see below) (Guinan et al., 1972; Tsuchitani, 1977, 1997). Unlike in MSO, the tonotopy in MNTB and LSO neurons has a compressed representation of low frequencies but spans the full high-frequency end of cochlear sensitivity. A second manifestation of the high-frequency bias is across species, simply in terms of nuclear volume. MNTB and LSO covary in size and tend to be better developed in small mammals with high-frequency hearing than in larger mammals with low-frequency hearing, such as humans, where the MSO is better developed (Adams, 1996; Moore, 1987). The bias of the LSO circuit toward high frequencies and the opposite bias in the MSO circuit, combined with their different binaural sensitivity (for interaural intensity and time differences, respectively) are traditionally interpreted as an embodiment of the “duplex theory of sound localization.”

High-frequency hearing has been pushed to a considerably higher limit in mammals than in other vertebrates, whose hearing is mostly restricted to frequencies below ~10 kHz (Fay, 1992). It has been argued that sound localization based on interaural intensity differences was one of the evolutionary forces driving mammalian high-frequency hearing (Heffner and Heffner, 2008; Manley, 2010; Masterton et al., 1969).

#### **Beyond the “Level Meter”**

Binaural studies show that some LSO neurons can support discrimination of interaural intensity differences with a resolution near behavioral threshold (~1 dB) (Tollin et al., 2008). However, while human discrimination thresholds are invariant with overall stimulus intensity, this is not the case for LSO responses—a problem that was identified in initial studies (Tsuchitani and Boudreau, 1969) and was recently reexamined (Tsai et al., 2010). These studies suggest that some type of pooling across LSO neurons is needed to account for behavior. However, the view that the LSO circuit acts as the proposed “level meter” is unsatisfactory for other, quite distinct reasons.

First, nothing in this operation calls for a giant synapse. Physical intensity is not an instantaneous quantity. For adequate measurement, some averaging over time is required (e.g., over several cycles of the sound waveform’s fine-structure or envelope; see next section). As discussed below, bushy cells excel at the coding of certain temporal features, over a frequency range similar to the auditory nerve (Joris and Yin, 1998; Joris et al., 1994a), but this ability is contrary to the need for integration required to measure intensity. At low frequencies, they show very limited dynamic range, and at frequencies above a few kHz they produce variable and irregular “primary-like” spike trains that lock on to fast changes in amplitude but are a poor choice to supply monaural intensity information. Computational models lead to similar conclusions (Bures and Marsalek, 2013; Yue and Johnson, 1997). Indeed, other cell types in the cochlear nucleus are superior in the coding of intensity because of longer integration times (reducing variability in spike rate) and/or wider dynamic range (Rhode and Smith, 1986; Shofner and Dye, 1989). In fact, at all synaptic stages between inner hair cell and LSO neuron, the circuit stands out—even relative to other auditory circuits at the brainstem level—by its limited neuronal convergence (even monoinnervation in case of the calyx!) and fast membrane and synaptic properties (Krächan et al., 2017; Young and Oertel, 2004), which are factors contrary to temporal integration.



A second argument against the LSO being a simple “level meter” may seem rather conceptual but has experimental support. It is easy to appreciate the benefit of computing interaural time differences as “early” as possible in the ascending auditory system—where there is still full access to temporal features coded in monaural channels. Indeed, as is true in other sensory systems, the frequency range over which spikes are locked to temporal stimulus features becomes more restricted when ascending in these systems (e.g., Joris et al., 2004). But it is less obvious to see the benefit of computing interaural intensity difference at a very low anatomical level rather than, e.g., at the level of the midbrain. The inferior colliculus is a key target structure for virtually all auditory main brainstem nuclei including LSO, and there are many combinatorial possibilities of ascending excitatory and inhibitory pathways that would allow creation or modification of sensitivity to interaural intensity difference. Such “neural design” considerations do not necessarily provide much insight, but there are empirical observations that suggest that the role of LSO must indeed be sought beyond the computation of interaural intensity difference. Most strikingly, sensitivity to interaural intensity differences is still observed in the inferior colliculus after neurotoxic lesioning of the binaural nuclei of the superior olivary complex, including the LSO circuit (Li and Kelly, 1992), but such lesions nevertheless impair sound localization (van Adel and Kelly, 1998; Kavanagh and Kelly, 1992). Mice with genetic deletion of MNTB were impaired in the detection of rapid spatial (left-right) shifts of sounds, but spatial acuity was not affected for longer integration times (Jalabi et al., 2013). Experiments in which various subcollicular inputs were lesioned or pharmacologically inactivated, as well as intracellular recordings allowing direct observation of inhibition and excitation, have revealed that sensitivity to interaural intensity differences in the midbrain does not just reflect input from LSO, but is a robust property that comes about through a diversity of pathways and mechanisms, including *de novo* creation in the inferior colliculus (Kuwada et al., 1997; Park and Pollak, 1993; Pollak, 2012). Finally, in birds, sensitivity to interaural intensity differences first arises in lemniscal nuclei, at a level closer to the midbrain than the LSO in mammals. Moreover, this sensitivity shows physiological features similar to that in the mammalian LSO, but without involving giant axosomatic synapses (Curry and Lu, 2016; Mogdans and Knudsen, 1994).

In short, we argue that characterization of the LSO circuit as the instantiation of a “level meter” fails to clarify the need for its calyx and other structural and physiological specializations and thereby misses an essential function that the circuit is performing. A recent study (Brown and Tollin, 2016) provides a powerful illustration of the above two lines of argument. Consistent with earlier studies (Hartmann and Constan, 2002), the authors show that the sensitivity of humans to interaural intensity differences is little affected by the temporal structure of sounds at the two ears. Even when the sounds to the two ears are significantly decorrelated, as happens naturally due to, e.g., reverberations of sound off physical objects or through superposition with other sounds, human subjects are still able to detect interaural intensity differences and to lateralize sources in azimuth. This strongly suggests that fast temporal structure in ongoing sounds is discarded in the evaluation of interaural intensity differ-

ences, e.g., by temporal averaging. Indeed, this was also observed in single neurons in the inferior colliculus, and Brown and Tollin (2016) make a convincing case that this ability requires temporal integration of sound over a window (of ~3 ms) that is much wider than observed in the LSO (~1 ms).

### Timing in the Afferent Circuits

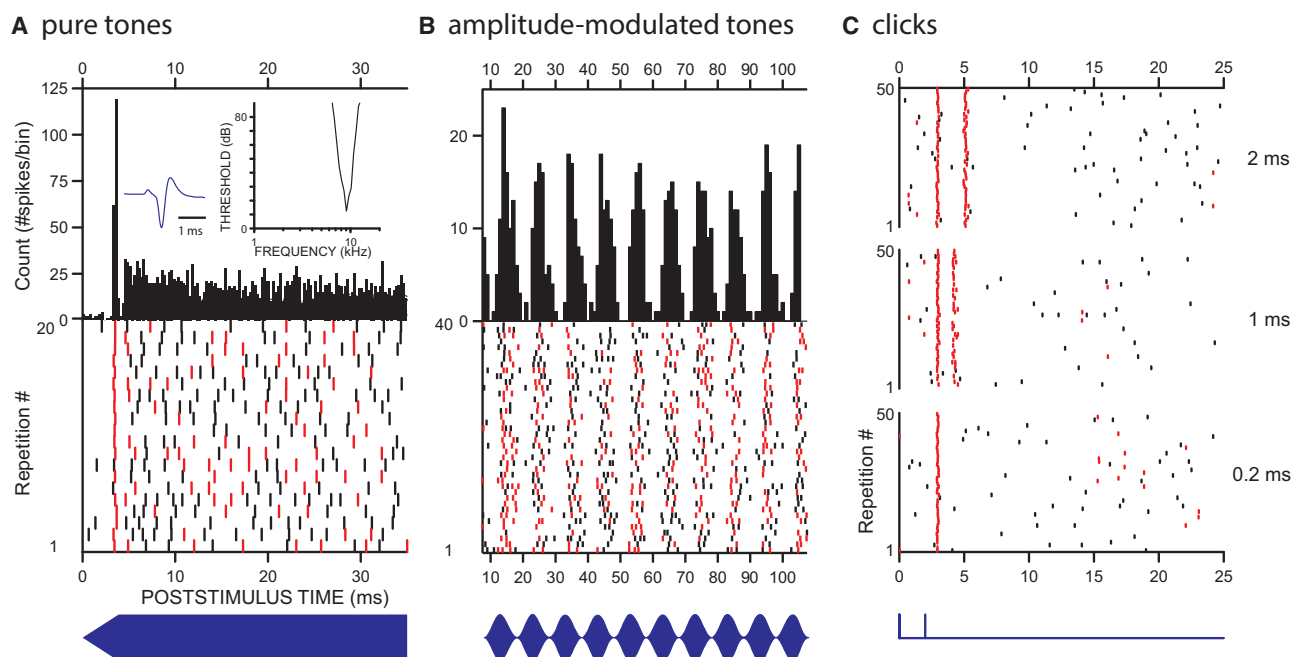
The previous sections suggest that temporal factors are somehow key in the LSO circuit. An obvious (but not the only) possibility is a role in the processing of interaural time differences. Such processing necessarily relies on the comparison of temporal features that are peripherally coded and transmitted to the binaural nuclei. In this section, we review these temporal features and their coding.

At each point along its length, the cochlear basilar membrane vibrates in response to a limited range of frequencies. To most natural sounds, this motion takes the form of fast vibrations near the frequency to which the basilar membrane is most sensitive at the point examined, and the amplitude of these fast vibrations shows a slower waxing and waning. These fast and slow components are referred to as temporal fine-structure and envelope, respectively. Some sounds are too brief to be meaningfully described in terms of fine-structure and envelope; we refer to them as transients or clicks.

The fast temporal component—fine-structure—is coded in the auditory nerve via so-called phase-locking in which fibers tend to fire spikes at a certain phase of successive stimulus cycles. There is an upper limit above which this process of nerve phase-locking disappears; this limit is species dependent (Weiss and Rose, 1988) but is around a few kHz in most mammals and decreases at each synaptic node in the ascending auditory system. In humans, the upper limit is likely rather low (Joris and Verschooten, 2013) since their sensitivity to interaural time differences of pure tones has an abrupt upper limit just below 1.5 kHz (Hartmann and Macaulay, 2014). In view of the limited representation of low frequencies and therefore of temporal coding of fine-structure in the LSO circuit, the role of the calyx must be sought in other forms of temporal coding.

In sounds other than pure tones, such as natural signals, temporal information is present at higher frequencies in the form of envelopes and transients (Figure 4), and there is ample psychophysical evidence that these cues enable sensitivity to interaural time differences. These are also the temporal cues most likely to be relevant to the LSO circuit, given its high-frequency bias. In a later section, we review that LSO neurons indeed show sensitivity to interaural time differences of envelopes and transients. Such sensitivity obviously requires that these cues are coded in the monaural afferent chains converging on the LSO. It is worth examining whether morphological or physiological features of these afferent chains have special relevance toward such coding.

A striking property of the LSO circuit is that the main neurons of the monaural pathways converging on the LSO (i.e., the SBC, GBC, and MNTB) share the same morphological and physiological “bushy cell features.” Contrasted against other neurons at this brainstem level, these features include compact (“bushy”) dendritic trees, large axosomatic synapses, and fast membrane properties (Young and Oertel, 2004). The temporal behavior of



**Figure 4. Responses of a High-Frequency MNTB Neuron in Cat to Three Classes of Sound that Constrain Spiking to Narrow Time Windows**

(A) Responses to a pure tone at 9 kHz. The raster (bottom) shows spikes to 20 repetitions of the tone. Red spikes indicate occurrence of another spike at the same time (within a 50  $\mu$ s window) in another repetition. Clearly, only at stimulus onset does the cell consistently fire a spike within a narrow time window. Spikes after stimulus onset are essentially random in their timing: there is no phase-locking of these spikes to the “fine-structure” of the tone (not illustrated). The post-stimulus time histogram (upper panel), which consists of spike counts as a function of post-stimulus time, shows the typical “primary-like-with-notch” pattern of MNTB neurons. The response peak at stimulus onset reflects the consistent firing of an accurately timed spike in nearly all tone presentations (here: in 195 out of 200 presentations). The 50 ms tone was presented 200 times at 60 dB SPL; only 35 ms of the response is shown, for the first 20 repetitions in the lower panel and for all repetitions in the upper panel.

(B) Response to a 9 kHz tone but now sinusoidally amplitude-modulated at 100 Hz. The raster (bottom) and post-stimulus time histogram (top) show a portion of the sustained response. There is clear temporal patterning in the response, locked to the stimulus envelope.

(C) Responses to pairs of 20  $\mu$ s rarefaction clicks, presented 50 times and separated by 0.2 ms (bottom), 1 ms (middle), and 2 ms (top). Clicks clearly evoke spikes at consistent times, as evidenced by the red color of most occurrences. Moreover, the neuron quickly recovers from a preceding spike and fires a spike to most instances of a click preceded by another click only 1 ms earlier (middle panel). The neuron had high spontaneous activity (103 spikes/s); minimum threshold (12 dB) and characteristic frequency (9,060 Hz) are illustrated by the threshold tuning curve (insert in A). The extracellularly recorded spike shows a prepotential (insert in A). The stimuli are schematically illustrated under the rasters (for C, only a click pair with a 2 ms interval is shown). Note that the spikes that are coincident across repetitions (marked in red) give a simple prediction of coincident activity in convergent MNTB inputs to the postsynaptic LSO neuron (Joris, 2003).

SBC, GBC, and MNTB neurons is remarkable, but is often misstated. These neurons are extremely well phase-locked to the fine-structure of low-frequency tones (Joris et al., 1994a, 1994b; Smith et al., 1998) and broadband noise (Louage et al., 2005), but these properties are more relevant for the MSO than the LSO, given the limited low-frequency representation of LSO. The temporal behavior of SBC, GBC, and MNTB at high frequencies is less unusual: the physiological classification of these cells is “primary-like” (SBC) or “primary-like-with-notch” (GBC and MNTB; see Figure 4A), indicating that these responses are rather auditory nerve-like in their firing and tuning properties. Compared to other auditory neurons at a similar anatomical level, high-frequency SBC, GBC, and MNTB neurons do not show unusually high firing rates, or the most accurate timing to envelopes. Neurons in another subnucleus of the cochlear nucleus (posteroventral) are certainly more notable in these regards, but these neurons have wider frequency tuning and/or higher thresholds, and do not project to the main nuclei of the superior olivary complex (Godfrey et al., 1975; Rhode and Smith, 1986; Smith et al., 2005). Although the three bushy-type cells

(SBC, GBC, and MNTB) show good phase-locking to stimulus envelope (Figure 4B), it is not the strength of this phase-locking that is exceptional, but rather the wide range of modulation frequencies they transmit, which is almost as wide as in the auditory nerve (Joris and Yin, 1998). Also, GBC and MNTB neurons show a reliable and well-timed spike to stimulus onset (Blackburn and Sachs, 1989; Tsuchitani, 1994; Young et al., 1988) (Figure 4A). In summary, high-frequency neurons carrying monaural information to the LSO are characterized by low thresholds, rather nerve-like frequency tuning and sustained responses, phase-locking to envelopes over a wide range of modulation frequencies, and a well-timed and reliable onset response in GBC and MNTB neurons. From a binaural point of view, another relevant property is that responses of SBC neurons are quite similar to those of GBC and MNTB neurons (Joris and Yin, 1998). Thus, the inhibitory and excitatory inputs to LSO are rather well matched in threshold, frequency tuning, and modulation (Tsuchitani, 1997).

The matching of the inputs of both ears extends to a most remarkable property of this circuit: the close match in time delay,

despite the extra path length and synapse for the contralateral input. This property was already apparent in the first single-unit studies of LSO, which showed that the first spike elicited by an ipsilateral sound could be inhibited by a simultaneously gated stimulus to the contralateral ear, taken as “an indication that the latency of inhibition is about as fast as the latency of excitation” (Boudreau and Tsuchitani, 1968). This match in delay has since been reported using a variety of techniques and preparations, both for stimulus onset and for the ongoing portion of the response. There is some variability between LSO neurons, so that in some neurons inhibition reaches LSO before excitation, but overall a variety of measures agree quite well that contralateral inhibition is effectively delayed relative to ipsilateral excitation by only a fraction of a ms. In cat, the average delay of inhibition relative to excitation is  $\sim 200 \mu\text{s}$ , with about a third of the neurons showing earlier inhibition than excitation (Joris, 1996; Joris and Yin, 1998; Tollin and Yin, 2005). This matching critically depends on the “speeding up” by the large caliber axons of the neurons in the contralateral limb (particularly of the GBCs). There also appears to be a “slowing down” ipsilaterally by thin (but myelinated) axons of high-frequency SBCs, but this has been less studied. The location of the inputs (somatic for contralateral inhibition and dendritic for ipsilateral excitation) may also be relevant for relative timing.

### Sensitivity of LSO to Interaural Time Differences

If “temporal” features of the LSO circuit are unnecessary for the computation of interaural intensity difference, a reasonably hypothesis is that these features evolved to enable sensitivity to interaural time differences. Given the high-frequency bias of the circuit, and given that behavioral sensitivity to interaural time differences of high-frequency sounds can be remarkably acute, an obvious hypothesis is that temporal features of the LSO circuit evolved to enable this acuity.

This hypothesis was tested with a stimulus popular in early psychophysical studies: high-frequency tones amplitude-modulated by a low-frequency sine wave (Henning, 1974). Within certain limits, the envelope of such sounds is coded by the auditory nerve and the ipsi- and contralateral monaural pathways converging on the LSO (Joris and Yin, 1992, 1998), and indeed both MSO and LSO neurons are sensitive to interaural time differences of such sounds (Batra et al., 1997; Joris, 1996; Joris and Yin, 1995; Yin and Chan, 1990). However, Joris and Yin (1995) argued that the sensitivity of LSO neurons is not commensurate with the extreme extent of specializations present in this circuit, most prominently the calyx of Held. The core of the argument is that the envelope interaural time difference sensitivity of LSO neurons is easily swamped by superimposed interaural intensity differences. The physiological data are broadly consistent with psychophysical experiments showing that despite excellent discrimination of high-frequency interaural time differences in amplitude-modulated sounds, such interaural time differences are a rather weak cue to lateralize sound, particularly when pitted against other cues (Tollin and Yin, 2002).

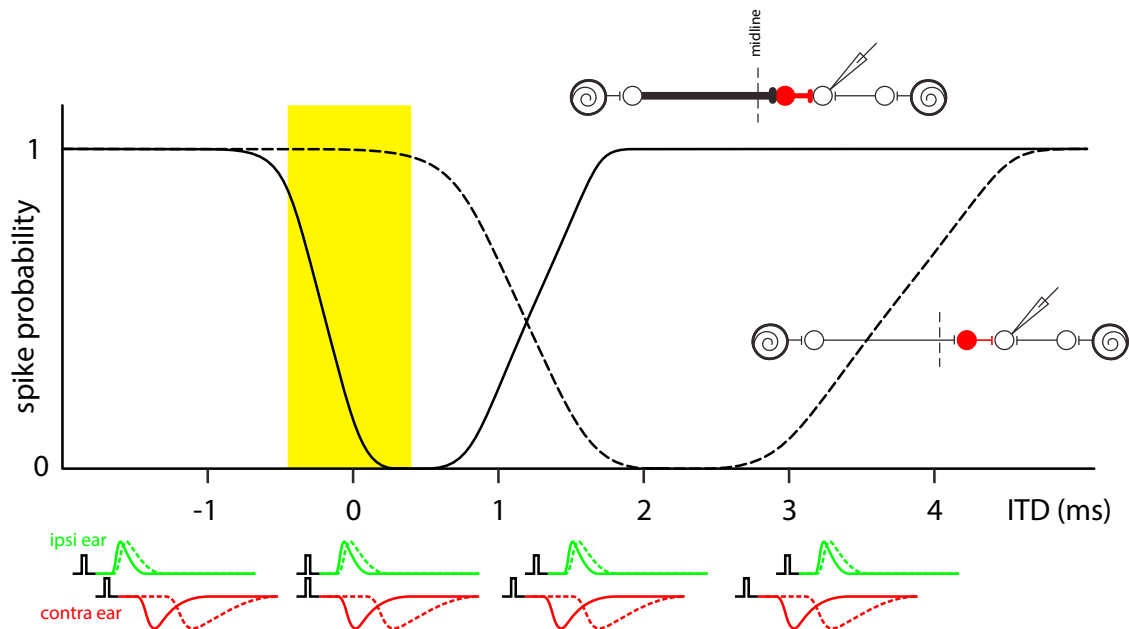
A related, but somewhat different take on the temporal specializations in the LSO circuit is that they are not there for sensitivity to interaural time differences per se, but to compute interaural intensity differences on corresponding parts of the

waveform (Joris and Yin, 1998). However, ultimately this view suffers from the same limitation: if effects of interaural time differences in AM stimuli are weak, this also means that the alignment is not very critical, and that again this requirement is not commensurate with the enormity of the calyx.

Another possibility that was explored is sensitivity to interaural time difference in fine-structure. For reasons that are unclear, classical studies of LSO reported a lack of binaural sensitivity in LSO neurons tuned to low frequencies (Tsuchitani, 1977). Later studies not only revealed that low-frequency LSO neurons can show sensitivity to interaural intensity differences, but also demonstrated that these neurons feature sensitivity to interaural time difference of pure tones (Caird and Klinke, 1983; Finlayson and Caspary, 1991; Joris and Yin, 1995; Tollin and Yin, 2005). However, in various ways this sensitivity is less robust than that in low-frequency MSO neurons. Combined with the fact that the low-frequency representation of the LSO is small, so that the vast majority of LSO neurons are tuned to frequencies too high for an effect of fine-structure, it is not plausible that binaural sensitivity to fine-structure “explains” the presence of the calyx and other temporal features.

There is, however, one form of interaural time difference to which LSO neurons are particularly sensitive. *In vitro* studies showed a steep dependence of firing probability when shocks to the inhibitory and excitatory pathways were delivered with different time delays (Sanes, 1990; Wu and Kelly, 1992). A similar sensitivity was found to acoustic clicks (Caird and Klinke, 1983; Irvine et al., 2001; Joris and Yin, 1995). The basic shape of this sensitivity is very simple, as illustrated by the cartoon in Figure 5 (solid line). At large (positive or negative) interaural time differences, there is a reliable response (1 spike/stimulus) to the ipsilateral click. This is as expected, as there is no overlap between the EPSP and IPSP triggered by the ipsi- and contralateral ear, respectively (Figure 5, cartoons at bottom). However, at small interaural time differences, the PSPs triggered by the two ears overlap, and indeed the spike output can be completely inhibited. The striking features of the few responses published are the steepness of the slopes and the narrow width of the interaural time difference range over which the two ears interact.

Somewhat surprisingly, sensitivity to interaural time differences of transient stimuli has received little attention in physiological studies, perhaps partly because these stimuli generate strong field potentials due to the synchronous firing of many neurons, which hamper isolation of spikes from a single neuron. There are no published accounts of responses of MSO neurons to such stimuli. A study of low-frequency neurons in the inferior colliculus shows ITD sensitivity that is very intensity dependent (Carney and Yin, 1989). Preliminary whole-cell recordings from identified MSO and LSO neurons confirm sensitivity with sharp slopes in LSO neurons (cf. Figure 5) and show surprisingly poor sensitivity in MSO neurons, apparently due to a breakdown of the coincidence requirement in these neurons in response to transient stimuli (Franken et al., 2016). MSO neurons are “EE” (excited by ipsi- and contralateral sounds) but typically do not respond well to monaural sustained stimuli: they require near-coincident excitatory inputs triggered by ipsi- and contralateral sound (Franken et al., 2015; van der Heijden et al., 2013). However, monaural transient stimuli appear to cause sufficient



**Figure 5. Schematic of Temporal Specializations of the LSO Circuit toward Sensitivity to Interaural Time Differences (ITDs) of Transients**  
The idealized curves in the main panel show the probability of obtaining a spike output of an LSO neuron in response to a pair of clicks (one click at each ear) of which the ITD is varied. The schematics below these panels illustrate click-pairs and the EPSPs (green) and IPSPs (red) triggered by them in the LSO neuron. Two scenarios are depicted. The dashed lines and bottom circuit cartoon show hypothetical ITD sensitivity in the absence of temporal specializations. The PSPs are prolonged in time and transmission of the contralateral input is slow relative to the ipsilateral input. Effective interaction of the IPSP and EPSP (trough in spike probability, dashed line) requires very large positive ITDs, which exceed the range of acoustic ITDs (yellow rectangle) that can be generated by the animal's head. The solid line and circuit cartoon on the upper right show the binaural sensitivity to transients that we expect in LSO. Fast membrane properties, large axosomatic terminals, and differences in axonal diameter result in fast and reliable EPSPs and IPSPs, which are triggered nearly simultaneously in response to simultaneous clicks to the two ears (ITD = 0 ms). In this cartoon, conduction of the contralateral input results in only a slight lag of the IPSP relative to the EPSP, despite the longer pathlength and extra synapse. For the range of binaural parameters that is acoustically plausible (yellow rectangle here and in Figure 4), this generates a steep dependence of spike probability on interaural time differences consistent with the sensitivity to interaural intensity differences to sustained sounds (Figure 4): increased probability of firing for a source closer to the ipsilateral ear.

coincident activity to bring MSO cells to threshold, so that there is no longer a binaural coincidence requirement and the firing rate of the cell is no longer sharply tuned to interaural time difference. Thus, transients appear to be the only stimulus for which an “anti-coincidence” or IE binaural interaction (as present in LSO) is superior to a coincidence or EE interaction (as present in MSO) to create tuning to interaural time differences.

### Hypothesis: (1) Lateralization of Transients

The acute sensitivity of LSO neurons to interaural time differences of brief sounds suggests a hypothesis regarding the function of the calyx of Held. There are two components to this hypothesis. In this section, we develop the physiological reasoning. In the next section, we tie the physiology to ecological acoustics.

We hypothesize that the presence of the calyx, and other temporal specializations in this circuit, generate coherent codes for interaural intensity and time differences. Figure 5 illustrates a hypothetical case where the temporal specializations are lacking (dashed lines and associated cartoons). In this situation, the IPSP generated by the contralateral ear arrives much later than the ipsilaterally evoked EPSP: only by giving a corresponding acoustic lead (positive interaural time difference) to the contralateral ear can the EPSP be inhibited, but the interaural

time difference required is larger than can be generated by a small head. This mismatch can be offset by speeding up the contralateral signal, and/or slowing down the ipsilateral signal. Moreover, if this offset in speed is tweaked right, it can lead to an interaural time difference curve positioned such that the steep slope corresponding to a leading EPSP being overtaken by a lagging IPSP (slope on the left in Figure 5) is within the range of interaural time differences experienced by the animal. In that case, the neural code signaling a sound source closer to the ipsi- than to the contralateral ear is consistent for interaural intensity and time differences: more firing means closer to the ipsilateral ear. Thus, the evolutionary development of the calyx, high conduction speed, etc. are perhaps an instance of exaptation: addition of temporal precision and speed to the IE circuit conveyed sensitivity to another spatial cue, resulting in a common neural code.

Existing data are only partially aligned with this idea. For example, in an LSO neuron extensively tested with click-interaural time differences in cat (Joris and Yin, 1995), interaural intensity difference and interaural time difference sensitivity were not consistent. However, in many cases of a larger sample of rat LSO neurons, the “correct” slope was centered within the relevant range of interaural time differences (Irvine et al., 2001), and this was also the case in an *in vitro* study in mouse (Wu and Kelly, 1992). Also, as mentioned, in the majority of LSO



neurons, the inhibitory input from the contralateral ear is effectively slightly (on average  $\sim 200 \mu\text{s}$ ) delayed relative to the excitatory ipsilateral input (Joris, 1996; Joris and Yin, 1998; Tollin and Yin, 2005), which predicts coherent interaural time difference and interaural intensity difference tuning in the majority of neurons. An additional factor to consider is that intensity changes by themselves introduce latency changes, which will interact with acoustically imposed interaural time differences (the so-called latency hypothesis). While such changes are expected to have only small and rather complex effects for sustained responses to sounds (Michelet et al., 2012), their effect can be significant for responses to sound transients (Irvine et al., 2001). In a natural environment, interaural intensity and time differences are highly correlated (Gaik, 1993), particularly at sound onset, because a source closer to one ear will stimulate that ear not only earlier in time but also more intensely. For the curve depicted with the solid line in Figure 5, interaural intensity differences will tend to deepen and broaden the trough at positive interaural time differences, and to steepen the slope at negative interaural time differences, increasing the similarity to tuning for interaural intensity differences (Figure 3) (Tollin, 2008).

### Hypothesis: (2) Ecological Acoustics

Taken together, the prominence of the LSO circuit in small mammals, its bias toward high frequencies, its binaurality, and its temporal features suggest that the calyx of Held evolved because it conferred on small mammals a survival benefit that has to do with spatial hearing based on temporal cues at high frequencies. The physiology suggests more specifically that sensitivity to transient sounds is key. We next ask how sound travels in natural space, what the sources are of high-frequency sounds, and how these sources may constitute an evolutionary selection pressure on binaural circuits.

#### High-frequency sounds strongly attenuate with distance

Sound emanating from a point in 3D space decreases in intensity with the square of the distance to the source: the inverse square law. The loss of energy due to such geometric spreading causes a decrease of 6 dB per doubling of distance. The actual attenuation of sound is, however, typically higher ("excess attenuation"). Atmospheric absorption causes a loss that increases with frequency squared. Vegetation in the sound path between source and receiver causes scattering and again particularly affects high frequencies (Römer and Lewald, 1992). Soil can also cause significant attenuation by acoustic cancellation (Aylor, 1972). Particularly for a small mammal living near a "soft" ground plane (forest floor, snow cover, grass), high frequencies are strongly attenuated (Embleton et al., 1976). Wind speed gradients, temperature gradients, and turbulence affect sound propagation and again their effects increase with increasing sound frequency (Römer and Lewald, 1992; Wiley, 2015). Simply put, environmental factors act like a low-pass filter (Römer, 1998), so that high frequencies likely indicate a source at close range. Besides affecting mean energy level, natural environments also increase the variability in the level of sounds traveling over a certain distance, particularly when the sound spectrum is narrowband. Finally, the degradation of sound due to these natural physical factors does not only affect sound intensity, but also the quality of directional cues. With increasing distance

from the source, the sound field becomes increasingly diffuse, and interaural intensity differences appear to degrade more strongly with distance than interaural time differences (Michelsen and Rohrseitz, 1997).

#### High-Frequency Sounds Derive Mainly from Other Animals

The spectrum of non-biogenic environmental sounds (wind, rain, surf, running water, thunder...) is strongly biased toward low frequencies, extending up to only a few kHz (Forrest, 1994). Besides being low frequency, these sounds are not punctate in either time or space. Conversely, high-frequency sounds usually signal the presence of (other) animals. Two broad categories can be distinguished: voluntary sounds emitted for communication, e.g., vocalizations, and involuntary sounds resulting from an animal's activities, e.g., locomotion.

A first class of biogenic, high-frequency sounds are intentional communication sounds. Land animals show a general inverse relationship between body mass and frequency range of vocalization (Fletcher, 2004): having physically smaller vocalization structures than humans, the vast majority of land mammals produce communication sounds much higher in frequency than humans. For example, communication calls of rats and mice extend into the ultrasonic frequency range (Mahrt et al., 2016; Seffer et al., 2014; Willadsen et al., 2014). Of course, this requires a matched hearing range at the receiver. As already mentioned, high-frequency hearing is indeed one of the hallmarks of mammalian hearing (Heffner and Heffner, 2008; Manley, 2010) and is thought to have been essential to the evolutionary success of early mammals (Kermack and Mussett, 1983), which were small and exploited the nocturnal niche ("nocturnal bottleneck hypothesis"; Gerkema et al., 2013). Early primates likely also had good high-frequency but poor low-frequency hearing (Coleman and Boyer, 2012). One impetus to communicate via high-frequency sounds is to minimize masking by low-frequency sounds. There are many examples of how animals have shifted the spectrum of their communication calls to higher frequencies to bypass masking from low-frequency environmental (e.g., anthropogenic) "noise" (Barber et al., 2010; Lesage et al., 1999; Shen et al., 2008; Slabbekoorn and Peet, 2003). A trade-off is that higher frequencies do not propagate as far as low frequencies and therefore limit the distance of communication, though that can be turned into advantage: it allows communication at close range without alerting predators (Stebbins, 1983; Wiley, 2015).

Not all biogenic, high-frequency sounds, however, are communication sounds. A second, significant class of biogenic sounds results (often unwantedly but unavoidably) from motor activity other than vocalization, e.g., locomotion. The causes of such adventitious sounds include physical friction, fracture, impact, and rapid deformation (Lewis and Fay, 2004). Examples are snapping twigs, crumpling leaves, and tumbling stones. Among environmental sounds, humans tend to cluster "discrete impact sounds" in a category by themselves (Gygi et al., 2007). Sounds are an unavoidable byproduct of locomotion and can therefore signal the approach of a predator or the presence of prey (Clark, 2016; Goerlitz and Siemers, 2007; Goerlitz et al., 2008; Magrath et al., 2007; Siemers and Güttinger, 2006). Just as predators have adaptations suited for stealth, natural

selection for the detection and localization of adventitious sounds is expected to be particularly strong in circumstances where vision is limited, e.g., at night or in dense vegetation. There has been much less study of adventitious sounds than of communication sounds, but there are a number of acoustical and behavioral studies that addressed the general category of “rustling” sounds from the viewpoint of sensory ecology. For example, the sounds caused by arthropods crawling on leaves consists of a series of broadband clicks with a spectrum between about 3 and 30 kHz. Behavioral experiments show that gray mouse lemurs detect and orient to these sounds, enabling them to locate prey at distances far greater (several meters) than would be possible with nocturnal vision (Goerlitz and Siemers, 2007; Goerlitz et al., 2008; Siemers et al., 2007). Conversely, mice choose routes that generate minimal rustling to avoid detection by auditory predators (Fitzgerald and Wolff, 1988; Roche et al., 1999).

Putting this all together, high-frequency transient sounds likely indicate a locomoting animal at close range. We surmise that mammals evolved under strong selection pressure to localize such transients, to detect prey and predators, and that temporal specializations in the LSO circuit, including the calyx of Held, are adaptations subserving such localization.

### Alternate Hypotheses

MNTB neurons provide inhibitory inputs to other brainstem targets besides the LSO (Smith et al., 1998; Sommer et al., 1993; Spangler et al., 1985). Much attention has been devoted to the projection to the MSO (Banks and Smith, 1992; Grothe and Sanes, 1993, 1994; Kuwabara and Zook, 1991, 1992). In cat, this projection is sparse and does not extend to very high frequencies (Smith et al., 1998). Application of strychnine caused changes in sensitivity to interaural time differences in MSO neurons of gerbil (Brand et al., 2002), from which it was proposed that the MNTB provides a tunable source of temporal delay of the contralateral excitatory inputs to the MSO. This hypothesis is controversial (Franken et al., 2015; Myoga et al., 2014; Roberts et al., 2013; van der Heijden et al., 2013), and in any case only concerns the role of MNTB at very low frequencies, at which individual stimulus cycles trigger IPSPs without temporal summation (Grothe and Sanes, 1994; Roberts et al., 2013, 2014). The high-frequency bias of MNTB neurons within and across species, mentioned above, necessitates that their main role must be sought elsewhere. Other projection targets of MNTB neurons that have received increasing attention are a peri-olivary nucleus (dorsomedial periolivary nucleus, particularly prominent as the superior paraolivary nucleus in rodents), as well as the ventral nucleus of the lateral lemniscus, which provides massive inhibition to the inferior colliculus. Neurons in these nuclei show only weak binaural sensitivity but have prominent monaural temporal properties (Batra and Fitzpatrick, 2002; Behrend et al., 2002; Berger et al., 2014; Dehmel et al., 2002; Felix et al., 2011; Kadner and Berrebi, 2008; Kopp-Scheinflug et al., 2011; Kuwada and Batra, 1999; Nayagam et al., 2005, 2006; Recio-Spinoso and Joris, 2014). However, it is not obvious that any of these properties demands the hyperspecialized calyx of Held, and it is unclear which functional role is subserved by these nuclei that could have led to

the selection of the calyx and its associated functional and morphological specializations.

The *in vitro* literature puts much emphasis on “high-frequency firing” enabled by the calyx. As pointed out above, high-frequency GBC and MNTB neurons are not unusual in their maximal firing rates: discharge rates of several hundred spikes/s are rather common in neurons at several anatomical levels between auditory nerve and midbrain. It is not so much the high-frequency firing per se that calls for a calyx, but rather the monoinnervation: to sustain high firing rates by a single input is certainly unusual. Monoinnervation provides the advantage of a purely sign-inverting relay: the firing properties established at the level of GBCs (Figure 2) are transmitted with great fidelity and speed as an inhibitory input to the LSO. As illustrated in Figure 4C, MNTB neurons are able to fire to transients in quick succession, enabled by the mechanisms reviewed above, particularly the complement of different Kv channels. Going one synapse back, the GBCs are unique in having a large number of somatic “modified endbulbs of Held,” derived from a dozen or more auditory nerve inputs (Spirou et al., 2005). While individual nerve fibers will not reliably respond to a succession of temporal events, such as a brief series of transients, GBCs detect coincidences across their inputs and thus pick up temporal structure in their inputs while preserving frequency selectivity. Thus, rather than enabling high-frequency firing per se, the calyx provides a means to relay a temporal and frequency-specific pattern of coincidences with high speed and reliability as an inhibitory copy to the LSO.

### Humans

The hearing of primates, and apes and old-world monkeys in particular, is shifted to low frequencies relative to other mammals (Coleman, 2009). Evolutionary changes in middle-ear and cochlear anatomy suggest a gradual shift from the mammalian high-frequency pattern to “reoccupy” low frequencies, with some loss in high-frequency hearing (Coleman and Boyer, 2012; Masterton et al., 1969). This emphasis on low frequencies fits with human brainstem anatomy, which shows a large MSO and small LSO (Moore, 2000). Whether humans have an MNTB has been highly controversial (Adams, 1996; Bazwinsky et al., 2003; Kulesza and Grothe, 2015; Kulesza, 2014; Moore, 1987; Moore and Moore, 1971; Richter et al., 1983). If present, it is certainly in a reduced form, even when compared to macaque (Kulesza, 2014). In any case, humans have excellent sensitivity to interaural intensity differences, over a wide range of frequencies (Yost and Dye, 1988). Also, humans have excellent sensitivity to interaural time differences of high-frequency transient sounds: for example, discrimination thresholds to the type of rustling sounds discussed above can be as low as 30  $\mu$ s (Ewert et al., 2012).

A better understanding of the circuits involved in binaural processing in animals and humans is not just of academic interest. Cochlear implants to remediate profound deafness deliver extremely transient stimuli (electrical pulses) to auditory nerve fibers at cochlear locations normally most sensitive to high frequencies. In the past decade, an increasing number of patients received bilateral implants, in the hope that bilateral stimulation will recruit the binaural system and improve spatial hearing and

speech perception in noisy environments. Unfortunately, interaural time difference thresholds are generally poor in these patients, for reasons that are only partly understood (Laback et al., 2015).

In humans with MRI-defined brainstem lesions due to multiple sclerosis, which are thought to affect fine spike timing, but not spike rate, thresholds for interaural intensity differences were typically not impaired while thresholds for ongoing interaural time difference were (Levine et al., 1993a). Such patients also show poor discrimination to interaural time differences of clicks (van der Poel et al., 1988). According to our hypothesis, the impaired thresholds for interaural time differences of ongoing and transient sounds would be caused by jittered spike patterns to the MSO and LSO, respectively. Due to their tight temporal synchronization, neurons in these pathways generate strong mass potentials which can be measured on the scalp. It may be possible to further refine the correlations that have been observed between scalp potentials, behavioral effects, and lesion location, by improving the interpretation of scalp potentials based on single-unit measurements in animals (Goldwyn et al., 2014; Laumen et al., 2016; Levine et al., 1993b; van der Poel et al., 1988).

### Summary and Outlook

The calyx of Held is such a striking specialization that its functional role in auditory processing would be thought to be obvious. Unlike the muscle endplate or squid giant axon or Mauthner cell or other such hypertrophied model structures, the function of the calyx is not self-evident. This is remarkable because it is a critical component of a well-studied circuit that is thought to create a well-understood sensitivity with clear behavioral significance, namely sound localization based on interaural intensity differences.

We have reviewed key structural and functional properties of the calyx of Held and the circuit of which it is part, and argue that it must have evolved in the context of processing of spatial attributes of high-frequency sounds. The calyx as such provides no known benefits toward the coding of sound intensity or the extraction of interaural intensity differences, and we provide conceptual and empirical arguments against the textbook view of the LSO as only a simple “interaural intensity difference processor.” We hypothesize that the role of the calyx is to generate a specific form of sensitivity to interaural time differences: to transient sounds. We further hypothesize that it evolved because such sensitivity has critical survival value, particularly for small ground-dwelling animals, as it enables lateralization of adventitious sounds generated by movements of other animals at close range. Movement unavoidably creates sounds containing high-frequency transients, which do not carry over large distances and which by their nature may not repeat and require instant detection and lateralization. The ability to detect and lateralize such transients can plausibly make the difference whether to have food or be food. Our hypothesis suggests that the most essential properties of MNTB are its sign inversion, its speed, and its reliability. Sustained high-frequency firing is not particularly relevant, in this scheme, though it should be pointed out that rustling sounds may consist of a short burst of transients rather than a single click.

The broader implication of our hypothesis is that both binaural nuclei, MSO and LSO, are binaural time processors, and that an IE interaction is the more adequate neural operation to extract useful timing information from high-frequency transient sounds. Thus, “interaural time difference sensitivity” should not be equated with coincidence mechanisms as exemplified by the MSO: at high frequencies, IE “anti-coincidence detection” is a more suitable neural operation to lateralize sounds. The role of the LSO should thus be sought along the same lines as the MSO: it is a structure early in the auditory system that applies a basic operation (here subtraction on a short timescale, or anti-coincidence) on inputs that faithfully monitor the nerve input to the CNS, both temporally and spectrally, with a bias toward high frequencies. The MNTB provides a negative image of cochlear activity that can be compared with the positive image of the other side. For pure tones, which are one spectral and temporal extreme of sound, this results in interaural intensity difference sensitivity. For clicks, which are the other spectral and temporal extreme, it results in marked interaural time difference sensitivity. The presence of the calyx does not make sense for tones, but does for clicks.

Our argument should not be misunderstood to imply that the LSO is irrelevant for sensitivity to interaural intensity differences or to interaural time difference of envelopes. Rather, these sensitivities do not “explain” the presence of the calyx and other temporal features in this circuit. The fact that the LSO contains a variety of cell types and that its neurons have sharp frequency selectivity already suggests that this nucleus is not just a lateralizer for transients but that different cell types use IE sensitivity toward different types of processing, over different timescales, where interaural time difference and interaural intensity difference should perhaps be seen more as a continuum than as a dichotomy.

The proposal that the calyx has evolved to enable lateralization of transients generated by other animals at close range is of course highly speculative and only partly refutable. To that end, we identify some general questions raised by the framework sketched above, as well as experiments directly targeted to test the hypothesis. It remains to be tested whether tuning of LSO neurons to interaural time differences of transients, particularly when combined with natural interaural intensity differences, is of a nature that gives plausibility to our hypothesis. *In vivo physiology* can refine our knowledge of binaural sensitivity to transients: is it usually congruent with sensitivity to interaural intensity differences of sustained stimuli? Do combined interaural intensity differences and interaural time differences affect spatial tuning in a beneficial way? A recent *in vivo* intracellular study (Franken et al., 2018) suggests that previous extracellular recordings were strongly biased toward non-principal LSO neurons, and that principal neurons are leaky, MSO-like neurons. Do these different LSO cell types differ in their ITD sensitivity? Almost completely lacking is knowledge of the *acoustic biotopes* of the small mammals that are popular models in auditory neuroscience and that differ much from humans, not only in size but also in other regards. For example, most of these animal models live close the ground (some even burrowing) and have a hearing range much extending above the human range. There is very little documentation of the kinds of soundscape these animals

experience, particularly with regard to adventitious sounds (Barber et al., 2010). Making *selective lesions* has been a long-standing problem in the functional study of the intricate binaural circuits in the brainstem. The new toolkit of optogenetics allows such experiments for the first time. Does lesioning of MNTB impair lateralization and detection of high-frequency transients? How does it affect detection of interaural intensity differences? At the *cellular level*, we need insight into the mechanisms that enable the short inhibitory window and steep sensitivity of LSO neurons to interaural time differences of transients, and into the effects of placement and convergence of inhibitory and excitatory synapses on different cellular compartments.

# ACKNOWLEDGMENTS

This work benefited from discussions with many colleagues over many years, particularly Romain Brette, Andrew Brown, Tom Franken, Dan Tollin, Philip Smith, Marcel van der Heijden, Lutz Wiegrebe, and Tom Yin, for which we are grateful. Gabriel Romero made the image of fluorescently labeled calyces. We thank Geert Callewaert, Karl Farrow, Ralf Schneggenburger, and three anonymous reviewers for critical reading of our manuscript. Support from the Fund for Scientific Research – Flanders (G.091214N) (P.X.J.) and NIH grant DC004450 (L.O.T.) is also gratefully acknowledged.

# AUTHOR CONTRIBUTIONS

P.X.J. conceived the work. P.X.J. and L.O.T. shared in the writing.

# REFERENCES

Adams, J.C. (1996). Neural circuits in the human auditory brainstem. In *Auditory Basis of Speech Perception*, W.A. Ainsworth and S. Greenberg, eds. (Keele University), pp. 39–44.

Aylor, D. (1972). Noise reduction by vegetation and ground. *J. Acoust. Soc. Am.* 51, 197–205.

Bal, R., and Oertel, D. (2001). Potassium currents in octopus cells of the mammalian cochlear nucleus. *J. Neurophysiol.* 86, 2299–2311.

Banks, M.I., and Smith, P.H. (1992). Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body. *J. Neurosci.* 12, 2819–2837.

Barber, J.R., Crooks, K.R., and Fristrup, K.M. (2010). The costs of chronic noise exposure for terrestrial organisms. *Trends Ecol. Evol.* 25, 180–189.

Batra, R., and Fitzpatrick, D.C. (2002). Monaural and binaural processing in the ventral nucleus of the lateral lemniscus: a major source of inhibition to the inferior colliculus. *Hear. Res.* 168, 90–97.

Batra, R., Kuwada, S., and 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, 1222–1236.

Bazwinsky, I., Hilbig, H., Bidmon, H.J., and Rübsamen, R. (2003). Characterization of the human superior olivary complex by calcium binding proteins and neurofilament H (SMI-32). *J. Comp. Neurol.* 456, 292–303.

Behrend, O., Brand, A., Kapfer, C., and Grothe, B. (2002). Auditory response properties in the superior paraolivary nucleus of the gerbil. *J. Neurophysiol.* 87, 2915–2928.

Berger, C., Meyer, E.M.M., Ammer, J.J., and Felmy, F. (2014). Large somatic synapses on neurons in the ventral lateral lemniscus work in pairs. *J. Neurosci.* 34, 3237–3246.

Blackburn, C.C., and Sachs, M.B. (1989). Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. *J. Neurophysiol.* 62, 1303–1329.

Borst, J.G.G., and Soria van Hoeve, J. (2012). The calyx of Held synapse: from model synapse to auditory relay. *Annu. Rev. Physiol.* 74, 199–224.

Borst, J.G., Helmchen, F., and Sakmann, B. (1995). Pre- and postsynaptic whole-cell recordings in the medial nucleus of the trapezoid body of the rat. *J. Physiol.* 489, 825–840.

Boudreau, J.C., and Tsuchitani, C. (1968). Binaural interaction in the cat superior olive S segment. *J. Neurophysiol.* 31, 442–454.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417, 543–547.

Brew, H.M., and Forsythe, I.D. (1995). Two voltage-dependent K<sup>+</sup> conductances with complementary functions in postsynaptic integration at a central auditory synapse. *J. Neurosci.* 15, 8011–8022.

Brown, A.D., and Tollin, D.J. (2016). Slow temporal integration enables robust neural coding and perception of a cue to sound source location. *J. Neurosci.* 36, 9908–9921.

Bures, Z., and Marsalek, P. (2013). On the precision of neural computation with interaural level differences in the lateral superior olive. *Brain Res.* 1536, 16–26.

Caird, D., and Klinke, R. (1983). Processing of binaural stimuli by cat superior olivary complex neurons. *Exp. Brain Res.* 52, 385–399.

Cant, N.B., and Casseday, J.H. (1986). Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei. *J. Comp. Neurol.* 247, 457–476.

Cao, X.-J., Shatadal, S., and Oertel, D. (2007). Voltage-sensitive conductances of bushy cells of the Mammalian ventral cochlear nucleus. *J. Neurophysiol.* 97, 3961–3975.

Carney, L.H.C., and Yin, T.C.T. (1989). Responses of low-frequency cells in the inferior colliculus to interaural time differences of clicks: excitatory and inhibitory components. *J. Neurophysiol.* 62, 144–161.

Clark, C.J. (2016). Locomotion-induced sounds and sonations: mechanisms, communication function, and relationship with behavior. In *Vertebrate Sound Production and Acoustic Communication*, R.A. Suthers, W.T. Fitch, R.R. Fay, and A.N. Popper, eds. (Springer), pp. 83–117.

Coleman, M.N. (2009). What do primates hear? A meta-analysis of all known nonhuman primate behavioral audiograms. *Int. J. Primatol.* 30, 55–91.

Coleman, M.N., and Boyer, D.M. (2012). Inner ear evolution in primates through the Cenozoic: implications for the evolution of hearing. *Anat. Rec. (Hoboken)* 295, 615–631.

Curry, R.J., and Lu, Y. (2016). Synaptic inhibition in avian interaural level difference sound localizing neurons. *eNeuro* 3. ENEURO.0309-16.2016. <https://doi.org/10.1523/ENEURO.0309-16.2016>.

de Lange, R.P.J., de Roos, A.D.G., and Borst, J.G.G. (2003). Two modes of vesicle recycling in the rat calyx of Held. *J. Neurosci.* 23, 10164–10173.

Dehmel, S., Kopp-Scheinflug, C., Dörrscheidt, G.J., and Rübsamen, R. (2002). Electrophysiological characterization of the superior paraolivary nucleus in the Mongolian gerbil. *Hear. Res.* 172, 18–36.

Dodson, P.D., Barker, M.C., and Forsythe, I.D. (2002). Two heteromeric Kv1 potassium channels differentially regulate action potential firing. *J. Neurosci.* 22, 6953–6961.

Doucet, J.R., and Ryugo, D.K. (2003). Axonal pathways to the lateral superior olive labeled with biotinylated dextran amine injections in the dorsal cochlear nucleus of rats. *J. Comp. Neurol.* 461, 452–465.

Embleton, T.F.W., Piercy, J.E., and Olson, N. (1976). Outdoor sound propagation over ground of finite impedance. *J. Acoust. Soc. Am.* 59, 267–277.

Englitz, B., Tolnai, S., Typlt, M., Jost, J., and Rübsamen, R. (2009). Reliability of synaptic transmission at the synapses of Held in vivo under acoustic stimulation. *PLoS ONE* 4, e7014.

Ewert, S.D., Kaiser, K., Kernschmidt, L., and Wiegrebe, L. (2012). Perceptual sensitivity to high-frequency interaural time differences created by rustling sounds. *J. Assoc. Res. Otolaryngol.* 13, 131–143.

Fay, R.R. (1992). Structure and function in sound discrimination among vertebrates. In *The Evolutionary Biology of Hearing*, D.B. Webster, A.N. Popper, and R.R. Fay, eds. (Springer New York), pp. 229–263.



- Felix, R.A., 2nd, Fridberger, A., Leijon, S., Berrebi, A.S., and Magnusson, A.K. (2011). Sound rhythms are encoded by postinhibitory rebound spiking in the superior paraolivary nucleus. *J. Neurosci.* 31, 12566–12578.
- Finlayson, P.G., and Caspary, D.M. (1991). Low-frequency neurons in the lateral superior olive exhibit phase-sensitive binaural inhibition. *J. Neurophysiol.* 65, 598–605.
- Fitzgerald, V.J., and Wolff, J.O. (1988). Behavioral responses of escaping permyscus leucopus to wet and dry substrata. *J. Mammal.* 69, 825–828.
- Fletcher, N.H. (2004). A simple frequency-scaling rule for animal communication. *J. Acoust. Soc. Am.* 115, 2334–2338.
- Ford, M.C., Alexandrova, O., Cossell, L., Stange-Marten, A., Sinclair, J., Kopp-Scheinpflug, C., Pecka, M., Attwell, D., and Grothe, B. (2015). Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat. Commun.* 6, 8073.
- Forrest, T.G. (1994). From sender to receiver: propagation and environmental effects on acoustic signals. *Am. Zool.* 34, 644–654.
- Forsythe, I.D., Tsujimoto, T., Barnes-Davies, M., Cuttle, M.F., and Takahashi, T. (1998). Inactivation of presynaptic calcium current contributes to synaptic depression at a fast central synapse. *Neuron* 20, 797–807.
- Franken, T.P., Roberts, M.T., Wei, L., Golding, N.L., and Joris, P.X. (2015). In vivo coincidence detection in mammalian sound localization generates phase delays. *Nat. Neurosci.* 18, 444–452.
- Franken, T.P., Smith, P., and Joris, P. (2016). *In vivo* whole-cell recordings of the lateral and medial superior olive to interaural time differences of transients. *Assoc. Res. Otolaryngol. Abs.* 39, 451.
- Franken, T.P., Joris, P.X., and Smith, P.H. (2018). Principal cells of the brainstem's interaural sound level detector are temporal differentiators rather than integrators. *eLife* 7, <https://doi.org/10.7554/eLife.33854>.
- Gaik, W. (1993). Combined evaluation of interaural time and intensity differences: psychoacoustic results and computer modeling. *J. Acoust. Soc. Am.* 94, 98–110.
- Geiger, J.R., Melcher, T., Koh, D.S., Sakmann, B., Seeburg, P.H., Jonas, P., and Monyer, H. (1995). Relative abundance of subunit mRNAs determines gating and Ca<sup>2+</sup> permeability of AMPA receptors in principal neurons and interneurons in rat CNS. *Neuron* 15, 193–204.
- Gerkema, M.P., Davies, W.I.L., Foster, R.G., Menaker, M., and Hut, R.A. (2013). The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proc. Biol. Sci.* 280, 20130508.
- Godfrey, D.A., Kiang, N.Y.S., and Norris, B.E. (1975). Single unit activity in the posteroventral cochlear nucleus of the cat. *J. Comp. Neurol.* 162, 247–268.
- Goerlitz, H.R., and Siemers, B.M. (2007). Sensory ecology of prey rustling sounds: acoustical features and their classification by wild Grey Mouse Lemurs. *Funct. Ecol.* 21, 143–153.
- Goerlitz, H.R., Greif, S., and Siemers, B.M. (2008). Cues for acoustic detection of prey: insect rustling sounds and the influence of walking substrate. *J. Exp. Biol.* 211, 2799–2806.
- Golding, N.L., and Oertel, D. (2012). Synaptic integration in dendrites: exceptional need for speed. *J. Physiol.* 590, 5563–5569.
- Goldwyn, J.H., Mc Laughlin, M., Verschooten, E., Joris, P.X., and Rinzel, J. (2014). A model of the medial superior olive explains spatiotemporal features of local field potentials. *J. Neurosci.* 34, 11705–11722.
- Gómez-Álvarez, M., and Saldaña, E. (2016). Different tonotopic regions of the lateral superior olive receive a similar combination of afferent inputs. *J. Comp. Neurol.* 524, 2230–2250.
- Grothe, B., and Sanes, D.H. (1993). Bilateral inhibition by glycinergic afferents in the medial superior olive. *J. Neurophysiol.* 69, 1192–1196.
- Grothe, B., and Sanes, D.H. (1994). Synaptic inhibition influences the temporal coding properties of medial superior olivary neurons: an in vitro study. *J. Neurosci.* 14, 1701–1709.
- Guillery, R.W. (2005). Observations of synaptic structures: origins of the neuron doctrine and its current status. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1281–1307.
- Guinan, J.J., Jr., and Li, R.Y.S. (1990). Signal processing in brainstem auditory neurons which receive giant endings (calyces of Held) in the medial nucleus of the trapezoid body of the cat. *Hear. Res.* 49, 321–334.
- Guinan, J.J., Norris, B.E., and Guinan, S.S. (1972). Single auditory units in the superior olivary complex. II: Locations of unit categories and tonotopic organization. *Int. J. Neurosci.* 4, 147–166.
- Gygi, B., Kidd, G.R., and Watson, C.S. (2007). Similarity and categorization of environmental sounds. *Percept. Psychophys.* 69, 839–855.
- Hartmann, W.M., and Constan, Z.A. (2002). Interaural level differences and the level-meter model. *J. Acoust. Soc. Am.* 112, 1037–1045.
- Hartmann, W.M., and Macaulay, E.J. (2014). Anatomical limits on interaural time differences: an ecological perspective. *Front. Neurosci.* 8, 34.
- Heffner, H.E., and Heffner, R.S. (2008). High-frequency hearing. In *Audition*, P. Dallos and D. Oertel, eds. (Elsevier), pp. 55–60.
- Helfert, R.H., and Schwartz, I.R. (1986). Morphological evidence for the existence of multiple neuronal classes in the cat lateral superior olivary nucleus. *J. Comp. Neurol.* 244, 533–549.
- Henning, G.B. (1974). Detectability of interaural delay in high-frequency complex waveforms. *J. Acoust. Soc. Am.* 55, 84–90.
- Hermann, J., Pecka, M., von Gersdorff, H., Grothe, B., and Klug, A. (2007). Synaptic transmission at the calyx of Held under in vivo like activity levels. *J. Neurophysiol.* 98, 807–820.
- Hruskova, B., Trojanova, J., Kulik, A., Kralikova, M., Pysanenko, K., Bures, Z., Syka, J., Trussell, L.O., and Turecek, R. (2012). Differential distribution of glycine receptor subtypes at the rat calyx of Held synapse. *J. Neurosci.* 32, 17012–17024.
- Huang, H., and Trussell, L.O. (2008). Control of presynaptic function by a persistent Na<sup>+</sup> current. *Neuron* 60, 975–979.
- Huang, H., and Trussell, L.O. (2011). KCNQ5 channels control resting properties and release probability of a synapse. *Nat. Neurosci.* 14, 840–847.
- Huang, H., and Trussell, L.O. (2014). Presynaptic HCN channels regulate vesicular glutamate transport. *Neuron* 84, 340–346.
- Irvine, D.R., Park, V.N., and McCormick, L. (2001). Mechanisms underlying the sensitivity of neurons in the lateral superior olive to interaural intensity differences. *J. Neurophysiol.* 86, 2647–2666.
- Ishikawa, T., Nakamura, Y., Saitoh, N., Li, W.-B., Iwasaki, S., and Takahashi, T. (2003). Distinct roles of Kv1 and Kv3 potassium channels at the calyx of Held presynaptic terminal. *J. Neurosci.* 23, 10445–10453.
- Jalabi, W., Kopp-Scheinpflug, C., Allen, P.D., Schiavon, E., DiGiacomo, R.R., Forsythe, I.D., and Maricich, S.M. (2013). Sound localization ability and glycinergic innervation of the superior olivary complex persist after genetic deletion of the medial nucleus of the trapezoid body. *J. Neurosci.* 33, 15044–15049.
- Johnston, J., Griffin, S.J., Baker, C., Skrzypiec, A., Chernova, T., and Forsythe, I.D. (2008). Initial segment Kv2.2 channels mediate a slow delayed rectifier and maintain high frequency action potential firing in medial nucleus of the trapezoid body neurons. *J. Physiol.* 586, 3493–3509.
- Joris, P.X. (1996). Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J. Neurophysiol.* 76, 2137–2156.
- Joris, P.X. (2003). Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J. Neurosci.* 23, 6345–6350.
- Joris, P.X., and Verschooten, E. (2013). On the limit of neural phase locking to fine structure in humans. *Adv. Exp. Med. Biol.* 787, 101–108.
- Joris, P.X., and Yin, T.C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J. Acoust. Soc. Am.* 91, 215–232.
- Joris, P.X., and Yin, T.C. (1995). Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J. Neurophysiol.* 73, 1043–1062.

- Joris, P.X., and Yin, T.C. (1998). Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J. Neurophysiol.* 79, 253–269.
- Joris, P.X., Carney, L.H., Smith, P.H., and Yin, T.C. (1994a). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J. Neurophysiol.* 71, 1022–1036.
- Joris, P.X., Smith, P.H., and Yin, T.C. (1994b). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J. Neurophysiol.* 71, 1037–1051.
- Joris, P.X., Schreiner, C.E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577.
- Kadner, A., and Berrebi, A.S. (2008). Encoding of temporal features of auditory stimuli in the medial nucleus of the trapezoid body and superior paraolivary nucleus of the rat. *Neuroscience* 151, 868–887.
- Kavanagh, G.L., and Kelly, J.B. (1992). Midline and lateral field sound localization in the ferret (*Mustela putorius*): contribution of the superior olivary complex. *J. Neurophysiol.* 67, 1643–1658.
- Kermack, K.A., and Mussett, F. (1983). The ear in mammal-like reptiles and early mammals. *Acta Palaeontol. Pol.* 28, 147–158.
- Kim, G., and Kandler, K. (2003). Elimination and strengthening of glycinergic/GABAergic connections during tonotopic map formation. *Nat. Neurosci.* 6, 282–290.
- Kim, J.H., and von Gersdorff, H. (2012). Suppression of spikes during posttetanic hyperpolarization in auditory neurons: the role of temperature, I(h) currents, and the Na(+)-K(+)-ATPase pump. *J. Neurophysiol.* 108, 1924–1932.
- Klug, A., and Trussell, L.O. (2006). Activation and deactivation of voltage-dependent K<sup>+</sup> channels during synaptically driven action potentials in the MNTB. *J. Neurophysiol.* 96, 1547–1555.
- Kopp-Scheinpflug, C., Lippe, W.R., Dörrscheidt, G.J., and Rübsamen, R. (2003). The medial nucleus of the trapezoid body in the gerbil is more than a relay: comparison of pre- and postsynaptic activity. *J. Assoc. Res. Otolaryngol.* 4, 1–23.
- Kopp-Scheinpflug, C., Tozer, A.J.B., Robinson, S.W., Tempel, B.L., Hennig, M.H., and Forsythe, I.D. (2011). The sound of silence: ionic mechanisms encoding sound termination. *Neuron* 71, 911–925.
- Krächan, E.G., Fischer, A.U., Franke, J., and Friauf, E. (2017). Synaptic reliability and temporal precision are achieved via high quantal content and effective replenishment: auditory brainstem versus hippocampus. *J. Physiol.* 595, 839–864.
- Kulesza, R.J., Jr. (2014). Characterization of human auditory brainstem circuits by calcium-binding protein immunohistochemistry. *Neuroscience* 258, 318–331.
- Kulesza, R.J., Jr., and Grothe, B. (2015). Yes, there is a medial nucleus of the trapezoid body in humans. *Front. Neuroanat.* 9, 35.
- Kuwabara, N., and Zook, J.M. (1991). Classification of the principal cells of the medial nucleus of the trapezoid body. *J. Comp. Neurol.* 314, 707–720.
- 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.
- Kuwabara, N., DiCaprio, R.A., and Zook, J.M. (1991). Afferents to the medial nucleus of the trapezoid body and their collateral projections. *J. Comp. Neurol.* 314, 684–706.
- Kuwada, S., and Batra, R. (1999). Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit. *J. Neurosci.* 19, 2273–2287.
- Kuwada, S., Batra, R., Yin, T.C., Oliver, D.L., Haberly, L.B., and Stanford, T.R. (1997). Intracellular recordings in response to monaural and binaural stimulation of neurons in the inferior colliculus of the cat. *J. Neurosci.* 17, 7565–7581.
- Laback, B., Egger, K., and Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hear. Res.* 322, 138–150.
- Laumen, G., Ferber, A.T., Klump, G.M., and Tollin, D.J. (2016). The physiological basis and clinical use of the binaural interaction component of the auditory brainstem response. *Ear Hear.* 37, e276–e290.
- Leão, R.M., Kushmerick, C., Pinaud, R., Renden, R., Li, G.-L., Taschenberger, H., Spirou, G., Levinson, S.R., and von Gersdorff, H. (2005). Presynaptic Na<sup>+</sup> channels: locus, development, and recovery from inactivation at a high-fidelity synapse. *J. Neurosci.* 25, 3724–3738.
- Lesage, V., Barrette, C., Kingsley, M.C.S., and Sjare, B. (1999). The effect of vessel noise on the vocal behavior of belugas in the St. Lawrence River Estuary, Canada. *Mar. Mamm. Sci.* 15, 65–84.
- Levine, R.A., Gardner, J.C., Stufflebeam, S.M., Fullerton, B.C., Carlisle, E.W., Furst, M., Rosen, B.R., and Kiang, N.Y.S. (1993a). Binaural auditory processing in multiple sclerosis subjects. *Hear. Res.* 68, 59–72.
- Levine, R.A., Gardner, J.C., Fullerton, B.C., Stufflebeam, S.M., Carlisle, E.W., Furst, M., Rosen, B.R., and Kiang, N.Y. (1993b). Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. *Hear. Res.* 68, 73–88.
- Lewis, E.R., and Fay, R.R. (2004). Environmental variables and the fundamental nature of hearing. In *Evolution of the Vertebrate Auditory System* (Springer Verlag), pp. 27–54.
- Li, L., and Kelly, J.B. (1992). Binaural responses in rat inferior colliculus following kainic acid lesions of the superior olive: interaural intensity difference functions. *Hear. Res.* 61, 73–85.
- Lorteije, J.A.M., Rusu, S.I., Kushmerick, C., and Borst, J.G.G. (2009). Reliability and precision of the mouse calyx of Held synapse. *J. Neurosci.* 29, 13770–13784.
- Louage, D.H., van der Heijden, M., and Joris, P.X. (2005). Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. *J. Neurosci.* 25, 1560–1570.
- Magrath, R.D., Pitcher, B.J., and Dalziel, A.H. (2007). How to be fed but not eaten: nestling responses to parental food calls and the sound of a predator's footsteps. *Anim. Behav.* 74, 1117–1129.
- Mahrt, E., Agarwal, A., Perkel, D., Portfors, C., and Elemans, C.P.H. (2016). Mice produce ultrasonic vocalizations by intra-laryngeal planar impinging jets. *Curr. Biol.* 26, R880–R881.
- Manley, G.A. (2010). An evolutionary perspective on middle ears. *Hear. Res.* 263, 3–8.
- Masterton, B., Heffner, H., and Ravizza, R. (1969). The evolution of human hearing. *J. Acoust. Soc. Am.* 45, 966–985.
- Mc Laughlin, M., van der Heijden, M., and Joris, P.X. (2008). How secure is in vivo synaptic transmission at the calyx of Held? *J. Neurosci.* 28, 10206–10219.
- Meyer, A.C., Neher, E., and Schneggenburger, R. (2001). Estimation of quantal size and number of functional active zones at the calyx of Held synapse by nonstationary EPSC variance analysis. *J. Neurosci.* 21, 7889–7900.
- Michelet, P., Kovacic, D., and Joris, P.X. (2012). Ongoing temporal coding of a stochastic stimulus as a function of intensity: time-intensity trading. *J. Neurosci.* 32, 9517–9527.
- Michelsen, A., and Rohrseitz, K. (1997). Sound localisation in a habitat: an analytical approach to quantifying the degradation of directional cues. *Bioacoustics* 7, 291–313.
- Mogdans, J., and Knudsen, E.I. (1994). Representation of interaural level difference in the VLVp, the first site of binaural comparison in the barn owl's auditory system. *Hear. Res.* 74, 148–164.
- Moore, J.K. (1987). The human auditory brain stem: a comparative view. *Hear. Res.* 29, 1–32.
- Moore, J.K. (2000). Organization of the human superior olivary complex. *Microsc. Res. Tech.* 51, 403–412.
- Moore, J.K., and Moore, R.Y. (1971). A comparative study of the superior olivary complex in the primate brain. *Folia Primatol. (Basel)* 16, 35–51.

- Myoga, M.H., Lehnert, S., Leibold, C., Felmy, F., and Grothe, B. (2014). Glycnergic inhibition tunes coincidence detection in the auditory brainstem. *Nat. Commun.* 5, 3790.
- Nakamura, Y., and Takahashi, T. (2007). Developmental changes in potassium currents at the rat calyx of Held presynaptic terminal. *J. Physiol.* 581, 1101–1112.
- Nayagam, D.A., Clarey, J.C., and Paolini, A.G. (2005). Powerful, onset inhibition in the ventral nucleus of the lateral lemniscus. *J. Neurophysiol.* 94, 1651–1654.
- Nayagam, D.A., Clarey, J.C., and Paolini, A.G. (2006). Intracellular responses and morphology of rat ventral complex of the lateral lemniscus neurons in vivo. *J. Comp. Neurol.* 498, 295–315.
- Park, T.J., and Pollak, G.D. (1993). GABA shapes sensitivity to interaural intensity disparities in the mustache bat's inferior colliculus: implications for encoding sound location. *J. Neurosci.* 13, 2050–2067.
- Perney, T.M., and Kaczmarek, L.K. (1997). Localization of a high threshold potassium channel in the rat cochlear nucleus. *J. Comp. Neurol.* 386, 178–202.
- Pollak, G.D. (2012). Circuits for processing dynamic interaural intensity disparities in the inferior colliculus. *Hear. Res.* 288, 47–57.
- Recio-Spinoso, A., and Joris, P.X. (2014). Temporal properties of responses to sound in the ventral nucleus of the lateral lemniscus. *J. Neurophysiol.* 111, 817–835.
- Rhode, W.S., and Smith, P.H. (1986). Encoding timing and intensity in the ventral cochlear nucleus of the cat. *J. Neurophysiol.* 56, 261–286.
- Richter, E.A., Norris, B.E., Fullerton, B.C., Levine, R.A., and Kiang, N.Y.S. (1983). Is there a medial nucleus of the trapezoid body in humans? *Am. J. Anat.* 168, 157–166.
- Rietzel, H.J., and Friauf, E. (1998). Neuron types in the rat lateral superior olive and developmental changes in the complexity of their dendritic arbors. *J. Comp. Neurol.* 390, 20–40.
- Roberts, M.T., Seeman, S.C., and Golding, N.L. (2013). A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry. *Neuron* 78, 923–935.
- Roberts, M.T., Seeman, S.C., and Golding, N.L. (2014). The relative contributions of MNTB and LNTB neurons to inhibition in the medial superior olive assessed through single and paired recordings. *Front. Neural Circuits* 8, 49.
- Roche, B.E., Schulte-Hostedde, A.I., and Brooks, R.J. (1999). Route choice by deer mice (*Peromyscus maniculatus*): reducing the risk of auditory detection by predators. *Am. Midl. Nat.* 142, 194–197.
- Römer, H. (1998). The sensory ecology of acoustic communication in insects. In *Comparative Hearing: Insects*, R.R. Hoy, A.N. Popper, and R.R. Fay, eds. (Springer), pp. 63–96.
- Römer, H., and Lewald, J. (1992). High-frequency sound transmission in natural habitats: implications for the evolution of insect acoustic communication. *Behav. Ecol. Sociobiol.* 29.
- Rusznák, Z., Bakondi, G., Pocsai, K., Pór, A., Kosztka, L., Pál, B., Nagy, D., and Szucs, G. (2008). Voltage-gated potassium channel (Kv) subunits expressed in the rat cochlear nucleus. *J. Histochem. Cytochem.* 56, 443–465.
- Sakaba, T. (2018). Kinetics of transmitter release at the calyx of Held synapse. *Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci.* 94, 139–152.
- Sanes, D.H. (1990). An in vitro analysis of sound localization mechanisms in the gerbil lateral superior olive. *J. Neurosci.* 10, 3494–3506.
- Sätzler, K., Söhl, L.F., Bollmann, J.H., Borst, J.G.G., Frotscher, M., Sakmann, B., and Lübke, J.H.R. (2002). Three-dimensional reconstruction of a calyx of Held and its postsynaptic principal neuron in the medial nucleus of the trapezoid body. *J. Neurosci.* 22, 10567–10579.
- Schneggenburger, R., and Forsythe, I.D. (2006). The calyx of Held. *Cell Tissue Res.* 326, 311–337.
- Schneggenburger, R., Sakaba, T., and Neher, E. (2002). Vesicle pools and short-term synaptic depression: lessons from a large synapse. *Trends Neurosci.* 25, 206–212.
- Scott, L.L., Mathews, P.J., and Golding, N.L. (2005). Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive. *J. Neurosci.* 25, 7887–7895.
- Seffer, D., Schwarting, R.K.W., and Wöhr, M. (2014). Pro-social ultrasonic communication in rats: insights from playback studies. *J. Neurosci. Methods* 234, 73–81.
- Shen, J.-X., Feng, A.S., Xu, Z.-M., Yu, Z.-L., Arch, V.S., Yu, X.-J., and Narins, P.M. (2008). Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. *Nature* 453, 914–916.
- Shofner, W.P., and Dye, R.H., Jr. (1989). Statistical and receiver operating characteristic analysis of empirical spike-count distributions: quantifying the ability of cochlear nucleus units to signal intensity changes. *J. Acoust. Soc. Am.* 86, 2172–2184.
- Siemers, B.M., and Güttinger, R. (2006). Prey conspicuousness can explain apparent prey selectivity. *Curr. Biol.* 16, R157–R159.
- Siemers, B.M., Goerlitz, H.R., Robsomanitrandrasana, E., Piep, M., Ramamananjato, J.-B., Rakotondravony, D., Ramilijaona, O., and Ganzhorn, J.U. (2007). Sensory basis of food detection in wild *Microcebus murinus*. *Int. J. Primatol.* 28, 291.
- Sierksma, M.C., and Borst, J.G.G. (2017). Resistance to action potential depression of a rat axon terminal in vivo. *Proc. Natl. Acad. Sci. USA* 114, 4249–4254.
- Slabbekoorn, H., and Peet, M. (2003). Ecology: Birds sing at a higher pitch in urban noise. *Nature* 424, 267.
- Smith, P.H., Joris, P.X., Carney, L.H., and Yin, T.C. (1991). Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *J. Comp. Neurol.* 304, 387–407.
- Smith, P.H., Joris, P.X., and Yin, T.C. (1998). Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. *J. Neurophysiol.* 79, 3127–3142.
- Smith, P.H., Massie, A., and Joris, P.X. (2005). Acoustic stria: anatomy of physiologically characterized cells and their axonal projection patterns. *J. Comp. Neurol.* 482, 349–371.
- Sommer, I., Lingenhöhl, K., and Friauf, E. (1993). Principal cells of the rat medial nucleus of the trapezoid body: an intracellular in vivo study of their physiology and morphology. *Exp. Brain Res.* 95, 223–239.
- Spangler, K.M., Warr, W.B., and Henkel, C.K. (1985). The projections of principal cells of the medial nucleus of the trapezoid body in the cat. *J. Comp. Neurol.* 238, 249–262.
- Spirou, G.A., Brownell, W.E., and Zidanic, M. (1990). Recordings from cat trapezoid body and HRP labeling of globular bushy cell axons. *J. Neurophysiol.* 63, 1169–1190.
- Spirou, G.A., Rager, J., and Manis, P.B. (2005). Convergence of auditory-nerve fiber projections onto globular bushy cells. *Neuroscience* 136, 843–863.
- Stanley, E.F. (2016). The Nanophysiology of Fast Transmitter Release. *Trends Neurosci.* 39, 183–197.
- Stebbins, W.C. (1983). *The Acoustic Sense of Animals* (Harvard University Press).
- Strutt, J.W. (1907). On our perception of sound direction. *Philos. Mag.* 13, 214–232.
- Tollin, D.J. (2008). Encoding of interaural level differences for sound localization. In *Audition*, P. Dallos and D. Oertel, eds. (Academic Press), pp. 631–653.
- Tollin, D.J., and Yin, T.C. (2002). The coding of spatial location by single units in the lateral superior olive of the cat. II. The determinants of spatial receptive fields in azimuth. *J. Neurosci.* 22, 1468–1479.
- Tollin, D.J., and Yin, T.C. (2005). Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *J. Neurosci.* 25, 10648–10657.
- Tollin, D.J., Koka, K., and Tsai, J.J. (2008). Interaural level difference discrimination thresholds for single neurons in the lateral superior olive. *J. Neurosci.* 28, 4848–4860.

- Tsai, J.J., Koka, K., and Tollin, D.J. (2010). Varying overall sound intensity to the two ears impacts interaural level difference discrimination thresholds by single neurons in the lateral superior olive. *J. Neurophysiol.* 103, 875–886.
- Tsuchitani, C. (1977). Functional organization of lateral cell groups of cat superior olivary complex. *J. Neurophysiol.* 40, 296–318.
- Tsuchitani, C. (1994). The brain stem evoked response and medial nucleus of the trapezoid body. *Otolaryngol. Head Neck Surg.* 110, 84–92.
- Tsuchitani, C. (1997). Input from the medial nucleus of trapezoid body to an interaural level detector. *Hear. Res.* 105, 211–224.
- Tsuchitani, C., and Boudreau, J.C. (1969). Stimulus level of dichotically presented tones and cat superior olive S-segment cell discharge. *J. Acoust. Soc. Am.* 46, 979–988.
- van Adel, B.A., and Kelly, J.B. (1998). Kainic acid lesions of the superior olivary complex: effects on sound localization by the albino rat. *Behav. Neurosci.* 112, 432–446.
- van der Heijden, M., Lorteije, J.A.M., Plauška, A., Roberts, M.T., Golding, N.L., and Borst, J.G.G. (2013). Directional hearing by linear summation of binaural inputs at the medial superior olive. *Neuron* 78, 936–948.
- van der Poel, J.C., Jones, S.J., and Miller, D.H. (1988). Sound lateralization, brainstem auditory evoked potentials and magnetic resonance imaging in multiple sclerosis. *Brain* 111, 1453–1474.
- Wang, L.Y., Gan, L., Forsythe, I.D., and Kaczmarek, L.K. (1998). Contribution of the Kv3.1 potassium channel to high-frequency firing in mouse auditory neurones. *J. Physiol.* 509, 183–194.
- Weiss, T.F., and Rose, C. (1988). Stages of degradation of timing information in the cochlea: a comparison of hair-cell and nerve-fiber responses in the alligator lizard. *Hear. Res.* 33, 167–174.
- Wiley, R.H. (2015). *Noise Matters* (Harvard University Press).
- Willadsen, M., Seffer, D., Schwarting, R.K.W., and Wöhr, M. (2014). Rodent ultrasonic communication: Male prosocial 50-kHz ultrasonic vocalizations elicit social approach behavior in female rats (*Rattus norvegicus*). *J. Comp. Psychol.* 128, 56–64.
- Wu, S.H., and Kelly, J.B. (1992). Binaural interaction in the lateral superior olive: time difference sensitivity studied in mouse brain slice. *J. Neurophysiol.* 68, 1151–1159.
- Yang, Y.-M., Aitoubah, J., Lauer, A.M., Nuriya, M., Takamiya, K., Jia, Z., May, B.J., Haganir, R.L., and Wang, L.-Y. (2011). GluA4 is indispensable for driving fast neurotransmission across a high-fidelity central synapse. *J. Physiol.* 589, 4209–4227.
- Yin, T.C.T., and Chan, J.C. (1990). Interaural time sensitivity in medial superior olive of cat. *J. Neurophysiol.* 64, 465–488.
- Yost, W.A., and Dye, R.H., Jr. (1988). Discrimination of interaural differences of level as a function of frequency. *J. Acoust. Soc. Am.* 83, 1846–1851.
- Young, E.D., and Oertel, D. (2004). Cochlear nucleus. In *The Synaptic Organization of the Brain*, G.M. Shepherd, ed. (Oxford University Press), pp. 125–163.
- Young, E.D., Robert, J.M., and Shofner, W.P. (1988). Regularity and latency of units in ventral cochlear nucleus: implications for unit classification and generation of response properties. *J. Neurophysiol.* 60, 1–29.
- Yue, L., and Johnson, D.H. (1997). Optimal binaural processing based on point process models of preprocessed cues. *J. Acoust. Soc. Am.* 101, 982–992.