

Early binaural hearing: the comparison of temporal differences at the two ears

Philip X. Joris

KU Leuven

Philip.Joris@kuleuven.be

Orcid <http://orcid.org/0000-0002-9759-5375>

Marcel van der Heijden

Erasmus MC

m.vanderheyden@erasmusmc.nl

Orcid <http://orcid.org/0000-0002-3876-1257>

Running title:

Physiology of early binaural hearing

Corresponding Author:

Philip Joris, Laboratory of Auditory Neurophysiology, Herestraat 49 bus 1021, B-3000 Leuven,
Belgium. Phone: +32 16 33 07 44

Keywords

Hearing

Stereo

Coincidence detection

Temporal processing

Brainstem

Sound localization

Abstract

Many mammals, including humans, are exquisitely sensitive to tiny time differences between the sounds at the two ears. These interaural time differences are an important source of information to detect sounds; to localize them in space; and for environmental awareness. Two brainstem circuits are involved in the initial temporal comparisons between the ears, centered on the medial and lateral superior olive (MSO and LSO). Cells in these nuclei, as well as their afferents, display a large number of striking physiological and anatomical specializations to enable sub-millisecond sensitivity. As such they provide an important model system to study temporal processing in the CNS. We review the progress that has been made in characterizing these primary binaural circuits, as well as the variety of mechanisms that have been proposed to underlie their function.

1. General introduction

Comparisons between receptors in different positions on the body can inform the brain regarding external space. Prominent examples are paired sense organs, particularly eyes and ears. Study of the neural processing exploiting such pairing – “stereo” – is attractive for several reasons. Stereo systems can be isolated behaviorally or physiologically: random-dot stereograms reveal objects when viewed binocularly but appear featureless monocularly (Julesz 1971), and some signals can be heard binaurally but not monaurally (Hirsh 1948; Licklider 1948). Stereo systems obviously require convergence from paired sensors onto single neurons, which facilitates identification of the relevant circuits. An appeal of auditory stereo is that the main cue consists of tiny time differences between the ears, making binaural hearing a model system to study temporal processing. We review the initial processing of interaural temporal differences at the level of the brainstem.

2. Temporal cues to the CNS

In mammals, binaural hearing relies on CNS mechanisms comparing sound-evoked neural events from the two ears. These events reflect the waveform of the physical stimuli as modified by the cochlea and nervous system. Since the auditory nerve (AN) is the bottleneck feeding the CNS, we first consider the transformation of sound waves to AN action potentials (**Figure**

1).

Figure 1

The first major step is the spectral analysis by the cochlea, which decomposes sound into frequency bands. For narrowband sounds such as tones this filtering has only minor effects (e.g. prolonged ringing after the tone ends, **Figure 1d,e**). Wideband noise, on the other hand, is completely transformed into a waveform with fairly regular zero crossings and a slowly fluctuating magnitude (“envelope”). The same wideband noise (**Figure 1c**) gives rise to entirely different filtered waveforms (**Figure 1f,g**) depending on the characteristic frequency (CF) of the cochlear site.

The second major step is transduction into action potentials. Importantly, nerve responses to low-frequency sounds (**Figure 1h,j**) reflect the fine-structure of the waveform, while responses to high frequencies only reflect slower envelope fluctuations (**Figure 1i,k**). This limitation in temporal coding, expressed in terms of the “phase-locking limit” (typically ~3-5 kHz: Rose et al. 1967; Johnson 1980; Weiss & Rose 1988), has major implications on binaural processing.

Rayleigh’s “duplex theory” stated that low- and high-frequency sounds are localized based on their ITDs and ILDs, respectively (Rayleigh (J.W. Strutt) 1907). This view is no longer held, as high-frequency ITDs do contribute to localization and binaural detection (McFadden & Pasanen 1976). But the use of high-frequency ITDs critically depends on the presence of sharp on- and offsets or pronounced envelope fluctuations. Even if present, these cues are easily spoiled by background noise and reverberation. In comparison, the information carried by the temporal fine-structure of low-frequency sounds is more robust (Devore & Delgutte 2010) and generally more accurate (Bernstein 2001).

3. Usefulness of interaural timing cues

A basic use of ITDs is the localization of a single sound source in a silent, anechoic environment. ITD varies with azimuth and its maximum value is determined by "headwidth": the distance between the ears, which is an important variable between species. In humans, it corresponds to a maximum ITD of $\sim 750 \mu\text{s}$. In daily life, however, idealized-static ITD estimation is rare. The sound waveforms at the two ears are rarely simple delayed copies of each other. Different components of complex sounds acquire different, frequency-specific, ITDs when reaching the ears (**Figure 2a**). In the presence of competing sound sources or multiple acoustic paths produced by reverberation, their interference causes ITDs to fluctuate in time and to transiently exceed the headwidth, even when one of the sources is dominant (**Figure 2b,c**).

Figure 2

These complications are not mere confounders of idealized waveform-ITDs: the CNS processes and exploits this dynamic ITD information. The binaural system is thus much more than a simple ITD-meter. Apart from single-source localization, binaural processing enhances source detection and identification. It also helps assess the acoustic properties of the environment and judge the reliability of available acoustic cues.

4. Behavioral limits

There is a rich history of binaural psychophysics in human listeners, ranging from headphone studies to free-field experiments and virtual reality (reviewed by Bernstein 2001; Durlach &

Colburn 1978; Middlebrooks & Green 1991). We restrict our discussion to a few topics that have a close relation to binaural physiology.

4.1. Sensitivity to ITDs

Headphones allow the delivery of waveforms that are identical except for an interaural delay. In the low-frequency range human listeners are extremely sensitive to such pure ITDs, requiring only a few microseconds for detection (Klumpp & Eady 1956). Free-field experiments are consistent with this finding. For low-frequency tones the sensitivity to changing the angle of the loudspeaker from the midline amounts to only a few degrees, corresponding to an ITD change of $\sim 10 \mu\text{s}$ (Mills 1958). When using a non-zero ITD (or a non-zero azimuth $< 60^\circ$) for the baseline condition, sensitivity drops somewhat, but not dramatically. Thus, humans do not have a pronounced “binaural fovea.” The exquisite sensitivity to ITDs at low frequencies is based on the fine-structure of the ongoing portion of the sound, and does not require marked onsets, offsets or a fluctuating envelope. ITD-sensitivity at low frequencies is much reduced when using artificial stimuli in which the envelope-ITD is varied while keeping the fine-structure fixed (Henning 1980)(**Figure 3b**).

Figure 3

For high frequencies (>1500 Hz) the situation is essentially reversed. The sensitivity to fine-structure cues entirely disappears. A striking demonstration of the contrasting phase-sensitivity at low and high frequencies is provided by binaural beats (**Figure 3d**). Because of the high-frequency “phase deafness”, ITDs of high-frequency sounds can only be perceived through a fluctuating envelope (**Figure 3b**)(Henning 1974) or on- and offsets (gating, **Figure 3c**). The sensitivity to ongoing envelope-ITDs of high-frequency stimuli, however, is 2-10 times poorer than sensitivity to ITDs in the fine-structure at low frequencies (Bernstein 2001), and it

degrades when the modulation depth is reduced (Nuetzel & Hafter 1981) e.g. as a result of reverberation (Houtgast & Steeneken 1985). In reverberant conditions, then, the only useful high-frequency ITD cue remaining is the direct field portion of sharp onsets, which arrives before any echoes do. Indeed, the earliest portion of waveforms is known to dominate the perceived lateral position (Wallach et al. 1949).

4.2. Binaural detection and ITDs

Auditory masking refers to soft signals becoming inaudible in the presence of noise. The masking potency of noise can be partially undone by exploiting binaural information. If signal and masker differ in their interaural attributes, the signal's audibility is often enhanced. In a classical headphone study (Hirsh 1948; Licklider 1948) listeners detect tones presented with opposite polarity (antiphasically) to the two ears in the presence of "diotic" (interaurally identical) wideband noise (**Figure 4**). Presenting low-frequency tones antiphasically (**Figure 4b**) enhances their audibility by ~13 dB compared to presenting them identically (**Figure 4a**). This phase-reversal trick was applied to improve radio communication in noisy airplanes. Binaural unmasking is likely to play a role in everyday situations where signal and noise sources are at different spatial locations.

Adding the soft antiphasic tone to the noise masker introduces subtle disparities between the resultant waveforms at the two ears. After peripheral filtering the disparities include time-varying ITDs of the type illustrated in **Figure 2b**. Neural circuits with exquisite sensitivity to static ITDs also sense dynamically varying ITDs (Joris et al. 2006b; Zuk & Delgutte 2017), making them well-suited to realize binaural unmasking and related binaural tasks such as detecting minute interaural decorrelation (Coffey et al. 2006; Louage et al. 2006; Shackleton et al. 2005).

Figure 4

There is psychophysical evidence that dynamic ITDs are indeed the major cue for detecting low-frequency tones in wideband noise (van der Heijden & Joris 2010). The contribution of ITD-processing is also consistent with the fact that binaural unmasking as well as correlation detection are much reduced for tones above 1500 Hz, the same frequency-limit observed for tonal ITDs and binaural beats (**Figure 3d**). And as illustrated in **Figure 2**, larger-than-headwidth ITDs are expected to play a role. Specifically, the interference between signal and noise leads to occasional interaural phase-opposition, creating short-term ITD of half the signal period, i.e., up to several milliseconds. If the binaural system is to exploit these large excursions, it needs sensitivity to ITDs well beyond the headwidth limit. Indeed, larger-than-headwidth ITDs do lead to continued lateralization (Blodgett et al. 1956) and are also likely to contribute to binaural unmasking (van der Heijden & Trahiotis 1999).

There is another interesting link between binaural detection and ITDs. Performance in binaural detection is relatively insensitive to overall interaural delays of the entire stimulus, i.e., signal and noise together (Bernstein & Trahiotis 2018). This robustness against overall delays resembles the absence of a clear “binaural fovea” for ITD-discrimination mentioned above. In terms of binaural processing it suggests that the system has access to a repertoire of internal delays to compensate external delays in the stimulus (i.e., ITDs). Such a range of internal delay was first proposed by Jeffress (1948) and is a key attribute of most binaural models (Colburn & Durlach 1978; Durlach 1972; Stern et al. 1988). Experimental evidence (van der Heijden & Trahiotis 1999) indicates the use of internal delays up to 2-3 ms, much larger than the $\sim 750\text{-}\mu\text{s}$ headwidth-limit of humans. The “best delays” of binaural neurons (described below) can be viewed as the physiological counterpart to the internal delays of binaural psychophysics.

Summarizing, in a limited view of the binaural system as a “single-source ITD meter,” the occurrence of cells tuned to larger-than-headwidth ITDs is something that needs special explanation (e.g. McAlpine et al. 2001). But when considering everyday situations involving competing sound sources, reverberation, etc., the headwidth ceases to be the natural limit of ITD processing.

4.3. Non-human

The ability to localize and detect prey or predator is one of the major functions of hearing. Sound localization of free-field stimuli has been studied extensively in many mammalian species (review: Brown & May 2005). Dichotic studies are technically challenging and few in number (Scott et al. 2007a; Tolnai et al. 2018; Wakeford & Robinson 1974) but basic capabilities of binaural detection and ITD-sensitivity have been documented. A general comparative theme has been to relate behavioral abilities to various anatomical features such as head size and brainstem anatomy. Importantly, some species (e.g. mice, rats) do not have low-frequency hearing. The main animal models to study ITD-sensitivity are rodents with large middle ear spaces (gerbil, guinea pig, chinchilla) and other species with good low-frequency hearing (cat, rabbit, macaque).

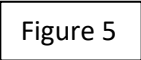
5. Circuit components

Initial binaural interactions occur in two brainstem circuits (**Figure 5**), centered on the medial superior olive (MSO) and lateral superior olive (LSO). These nuclei share several input sources and are embedded in the superior olivary complex (SOC). There is marked variation in the absolute and relative size of SOC nuclei across species, including primates (Moore & Moore

1971), and covariation with other characters (Harrison & Irving 1966; Heffner & Heffner 1992; Masterton et al. 1975).

5.1. Medial Superior Olive (MSO) circuit

The MSO is a striking elongated nucleus: a book-shaped slab only a few cells wide. Early studies (Ramón y Cajal 1909; Stotler 1953) drew attention to the bipolar morphology of its principal neurons, which show two main thick and branching dendrites oriented in opposite directions (Rautenberg et al. 2009; Smith 1995). The bipolar morphology enables dendritic segregation of excitatory inputs from spherical bushy cells (SBCs), which branch to supply lateral MSO-dendrites ipsilaterally and medial dendrites contralaterally (Beckius et al. 1999; Cant & Casseday 1986; Smith et al. 1993). Monaural sounds typically evoke low response rates from MSO-neurons, which are labeled “EE” (excited by contra- and ipsilateral ear). There is a tonotopic gradient along the MSO dorsoventral axis with an underrepresentation of high frequencies (Franken et al. 2015; Guinan et al. 1972; Karino et al. 2011). Whether there is also a rostrocaudal functional gradient has been a topic of considerable interest, discussed below.



MSO neurons also receive inhibitory input, which provides a second path of binaural convergence. Glycinergic terminals cluster on their soma (Clark 1969; Kapfer et al. 2002). Globular bushy cells (GBCs) provide axosomatic terminals to two glycinergic nuclei projecting to the MSO: the posteroventral lateral nucleus of the trapezoid body (pvLNTB) ipsilaterally, and the medial nucleus of the trapezoid body (MNTB) contralaterally (Banks & Smith 1992; Cant & Hyson 1992; Franken et al. 2016a; Kuwabara & Zook 1992; Roberts et al. 2014; Smith et al. 1991, 1998; Spangler et al. 1985; Spirou & Berrebi 1996, 1997; Spirou et al. 1990).

SIDEBAR 1

5.2. Lateral Superior Olive (LSO) circuit

The LSO is an S-shaped nucleus lateral to the MSO. Here, the inputs from the two ears are opposite in sign but not segregated to different dendrites. Ipsilateral inputs are excitatory and dominantly on distal dendrites. They are mostly derived from SBCs (at least partly the same SBCs also projecting to MSO: Shneiderman & Henkel 1985; Smith et al. 1993). Inhibitory inputs derive from the homolateral MNTB and dominate on LSO somata and proximal dendrites (Cant 1984; Cant & Casseday 1986; Spangler et al. 1985). Thus, LSO neurons are “IE”: inhibited by contralateral and excited by ipsilateral sound. This property confers sensitivity to ILDs: the firing rate of LSO neurons increases as sound location changes from contralateral to ipsilateral space.

The LSO is biased towards higher frequencies than the MSO (Gómez-Álvarez & Saldaña 2016; Guinan et al. 1972; Tsuchitani & Boudreau 1966). The combination of a difference in frequency bias (low in MSO, high in LSO) and a difference in binaural sensitivity (to ITDs in MSO, to ILDs in LSO) has led to the classical “duplex” view of these two binaural nuclei stated in most textbooks. It makes a tidy story indeed that these two brainstem circuits are dedicated to the extraction of the two binaural cues. Particularly for the LSO, this is a mischaracterization.

6. Monaural preprocessing

The coding of stimulus fine-structure is a hallmark property of the auditory system, and is taken to an extreme in bushy-type neurons. They feature large conductances with fast kinetics, low input-resistance and short time-constant, properties geared towards temporal coding (reviewed in Young & Oertel 2004). Sound-evoked responses of bushy-type neurons share

general features with AN fibers – hence the label “primary-like.” However, axonal recordings revealed strongly enhanced phase-locking to low-frequency tones compared to the AN (Joris et al. 1994a,b), in two ways. First, spikes occur over a narrower fraction of the stimulus cycle. The resulting “peakedness” of the cycle histograms (**Figure 6a**) is quantified by the “vector strength” (Goldberg & Brown 1969): a number between zero (no preferred phase) and unity (all spikes in a single bin). In the AN, pure-tone vector strength rarely exceeds 0.9, but in “high-sync” bushy cells values near 1 occur for frequencies up to almost 1 kHz. Bushy-type cells are also more consistent in firing a spike on every cycle. Such entrainment results in strikingly unimodal inter-spike-interval histograms. AN-fibers do not entrain: they tend to skip cycles and generate multimodal histograms. These two aspects, accurate temporal coding and superior entrainment, are also obvious in responses non-periodic sounds such as wideband noise (**Figure 6b**), but they can no longer be quantified in terms of vector strength and fraction of skipped cycles because those metrics require periodicity.

Shuffled autocorrelograms (SAC, Joris 2003) provide an analysis not requiring periodic stimuli. The algorithm has a natural resemblance to the cellular process of coincidence detection. The stimulus is repeated n times and all pairs of spike trains are evaluated in terms of coincidences, i.e. spikes occurring at the same instant in post-stimulus time. For example, the red dots in **Figure 6** indicate spikes for which at least one coincident spike occurred in another stimulus repetition. The coincidence count is much higher for the bushy cell than for the AN fiber. Repeating the count with a time shift δ introduced between each pair of spike trains being compared, and varying δ , yields the full correlogram (**Figure 6b**, right column). For the neurons

in **Figure 6** ($CF \sim 500$ Hz), a 1-ms delay produces a very low coincidence count, while a 2-ms delay (close to the characteristic period, CF^{-1}) produces a peak.

Correlograms provide a bridge between monaural temporal properties and binaural sensitivity. Consider a perfect coincidence detector receiving two identical bushy-cell inputs. If a wideband noise with variable ITD is presented to the two ears, the correlogram describes the “ITD function” (variation of firing rate with ITD) of this idealized detector. And if the two inputs reach the detector with a latency difference (e.g. due to differences in path length), its output is a horizontally shifted version of the correlogram, now peaking at a nonzero “best delay.” Alternatively, δ can be seen as the internal delay: the correlogram then describes the activation pattern of a population of coincidence detectors in response to noise with $ITD = 0$, each receiving the same input but with different internal delay.

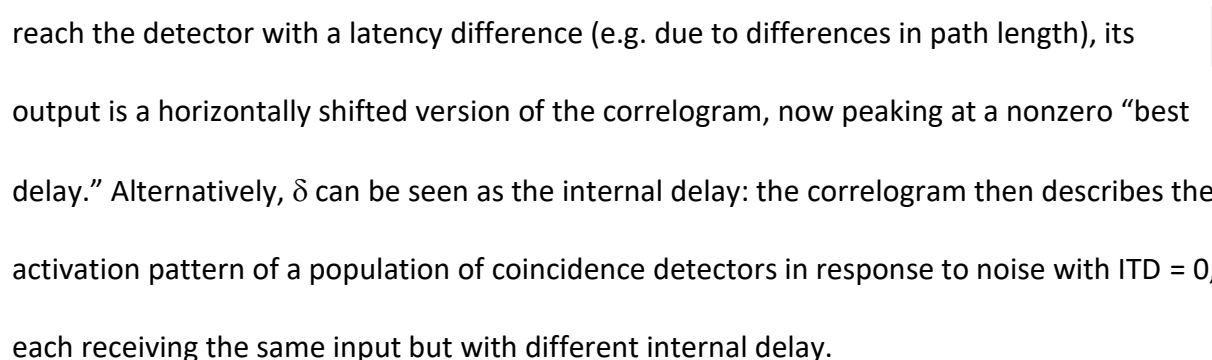


Figure 6

The enhanced temporal coding between AN and bushy-type neurons is interesting in itself. Models of this transformation (reviewed in Joris & Smith 2008) use a monaural form of coincidence detection. Bushy-type neurons code the stimulus more precisely and reliably (**Figure 6**), and this should benefit interaural temporal comparisons. Another, less obvious, benefit is the robustness against intensity variations: if exactly one spike is fired for each cycle, the firing rate is “clamped” at the stimulus frequency and therefore insensitive to suprathreshold changes in stimulus intensity. This likely contributes to the striking invariance in firing rate that binaural neurons can display with changes in stimulus intensity (Yin et al. 1986) and may help solve “Steven’s problem” (van de Par et al. 2001) and “Steve’s problem” (Colburn & Isabelle 2001), which address the difficulty of binaural cross-correlation models to deal with stimulus-level variability.

Several studies have used the correlogram technique to analyze whether enhanced temporal coding by bushy-type neurons improves binaural sensitivity. Simulated ITD-sensitivity to fine-structure and envelope was indeed much higher when based on spike-trains from bushy-type neurons than from AN fibers (van der Heijden et al. 2011). Likewise, neural thresholds for correlation discrimination were better when bushy cells were used for coincidence detection rather than nerve fibers, and could be as low as observed in humans (Louage et al. 2006). Importantly, thresholds were lowest when the population of detectors spanned a central range of internal delays similar to the range observed in binaural neurons. Large internal delays were beneficial for correlation discrimination of narrowband signals, for which the ITD-function shows large secondary peaks. This may explain the psychophysical paradox that decorrelation detection is better for narrow- than for broadband low-frequency noise (Gabriel & Colburn 1981).

7. SOC output

SIDEBAR 2

7.1. Forms of ITD-sensitivity to ongoing sound

Figure 7

The first demonstration of ITD-sensitivity at the single-cell level was in the IC (Rose et al. 1966). **Figure 7a,b** illustrates ITD-tuning in two IC-neurons in response to a broadband noise. ITD is varied over a range much larger than the headwidth. Positive ITDs refer to a stimulus leading in the contralateral ear. The black line in **Figure 7a** is representative of the vast majority of responses recorded in IC and shows the strong dependence of firing rate on ITD in the shape of a damped oscillation. A central peak at the Best Delay (BD) is flanked by deep troughs and

smaller secondary peaks. If the two ears are stimulated by independently generated noise waveforms (tokens “A/B”, cyan), there is no systematic ongoing-ITD but only a varying gating- or onset-ITD (cf. **Figure 3c**), which has no effect on overall spike rate. In the third condition (magenta), the noise token is identical at the two ears, but the polarity of the noise is inverted at one ear (as if reversing the 2 earphone wires). The resulting response is antiphase relative to the response to correlated noise. This is to be expected: waveform inversion gives a π phase shift of all stimulus components.

The response in **Figure 7b** shows another form of ITD-sensitivity. Again, the BD is positive, but the ITD-function shows a single peak surrounded by shallow troughs. Moreover, the ITD-function obtained to anticorrelated noise is virtually identical. This type of sensitivity is observed less frequently (and is little documented at the level of the MSO: Joris 1996; Plauška et al. 2016; Yin & Chan 1990), but is the dominant kind of response observed at higher (>2 kHz) CFs (Devore & Delgutte 2010; Joris 2003). It indicates that the neuron is sensitive to ITDs of the envelope rather than the fine-structure of the waveform. The envelope is symmetric and therefore not affected by changing the polarity in one ear. Decorrelation of the noises to the two ears again results in a low response-rate independent of ITD (**Figure 7b**, cyan). There is also a minority of IC responses (“troughers”) that are better described as being tuned in an inhibitory direction (Batra et al. 1997; McAlpine et al. 2001; Yin et al. 1986).

7.2. Distribution of Best Delays

A consistent finding in IC is that BDs are mostly positive, i.e. neurons are most active when the contra- precedes the ipsilateral stimulus. Studies in cat (Hancock & Delgutte 2004; Yin et al. 1986) found that the range of BDs covers the headwidth of these animals (about 350 to 400 μ s:

Roth et al. 1980). These findings nicely fit the general notion of a tuned system that covers the perceptually relevant cue-range and that is contralateralized like visual and somatosensory systems. This notion was challenged by the discovery that the range of BDs is CF-dependent: the upper bound is roughly hyperbolic at a constant phase value of π , the so-called π -limit. This was first described in guinea pig (McAlpine et al. 1996, 2001) and similar distributions have been found in IC and MSO recordings of various species: cat (Hancock & Delgutte 2004; Joris et al. 2006a) , gerbil (Brand et al. 2002; Day & Semple 2011; Pecka et al. 2008; Plauška et al. 2017), and chinchilla (Bremen & Joris 2013). This finding has implications for the mechanism underlying the BD, but the functional implications are less clear.

According to the “two-channel” model (McAlpine et al. 2001), the observed scattering of BDs is a bug rather than a feature: ITD-tuning is not optimized to have a range of BDs spanning the headwidth at all frequencies, but rather to have the steepest slope of sensitivity centered at zero ITD (Harper & McAlpine 2004). It is proposed that binaural temporal sensitivity relies on a comparison of two channels that show a sigmoidal dependence of firing rate for ITDs within the headwidth; a “push-pull” system in which two populations of MSO neurons, one in each hemisphere, respond in opposite direction to a given stimulus.

While the two-channel model may be adequate for the lateralization of simple stimuli such as single tones, its plausibility beyond such stimuli has been questioned by computational studies (Brette 2010; Day et al. 2012; Fitzpatrick et al. 1997; Goodman et al. 2013; Hancock 2007; van der Heijden et al. 2011). Also, the argument for “optimal slope positioning” ignores that neural sensitivity to decorrelation is largest at the peaks of ITD-sensitivity, not at the slopes (Coffey et al. 2006; Joris et al. 2006b; Louage et al. 2006; Shackleton et al. 2005; Yin et al. 1987). This can

be appreciated from **Figure 7a**: the ITD-functions for the three conditions of correlation have their largest differences near the BD.

8. Mechanisms of sensitivity to ITD

There are vastly more data from the midbrain than from the SOC, from which the IC inherits its ITD-sensitivity. Prominent phase-locked local field potentials (Mc Laughlin et al. 2010), small action potentials (Scott et al. 2007b), and small target size combined with poor accessibility, make single-cell recording in the SOC challenging. Positive identification of physiologically-characterized principal MSO and LSO neurons was recently achieved by intracellular labeling (Franken et al. 2015, 2018), but it is not entirely clear whether MSO and LSO are the only primary sources of ITD-sensitivity, which has often been observed in other SOC locations (Batra et al. 1997; Day & Semple 2011; Franken et al. 2016a; Goldberg & Brown 1969).

Examples of ITD-sensitivity to noise of an MSO and LSO neuron are shown in **Figure 7c,d**. Nearly all MSO-responses published are dominated by fine-structure, whereas most LSO responses are dominated by envelope. Note that the ITD-function of the LSO neuron is of the “trougher” variety, as expected from its IE binaural interaction: when ipsi- and contralateral inputs to the cell are maximally correlated, it is maximally inhibited.

Two concepts that have taken center stage in SOC studies of ITD-sensitivity, are internal delays and coincidence detection. We first discuss these concepts for MSO, and conclude with LSO physiology.

8.1. Internal delay

The mechanism(s) behind the range of BDs surmised psychophysically and observed physiologically sparked much debate. **Figure 8a** illustrates internal delay: it is the difference in delay between the effective input of left vs. right ear on the binaural neuron. If stimulus-ITD is compensated by internal delay, the effective inputs are coincident and the neuron is optimally excited. The finding of an inverse relationship between internal delay (estimated by BD) and CF (McAlpine et al. 1996, 2001) suggests a mechanism that somehow scales with CF^{-1} .

Jeffress (1948) proposed that internal delays are produced by axonal delay-lines. **Figure 8b** shows a rendition based on axonal tracing (Beckius et al. 1999; Smith et al. 1993): branches of the contralateral SBCs towards the rostral pole of the MSO tend to be shorter than towards its caudal pole. Systematic differences in axonal length between ipsi- and contralateral inputs would not only create a range of BDs at each frequency, but also a rostrocaudal map of BDs. Such a gradient was indeed suggested by MSO recordings (Yin & Chan 1990), in a direction consistent with the axonal tracing data. Thus, both anatomical and physiological data supported Jeffress' proposal (Joris et al. 1998). However, this interpretation is challenged by the finding of the inverse BD-CF relationship (McAlpine et al. 2001), which is not predicted by (although not necessarily inconsistent with) axonal delay lines. A re-analysis confirmed the existence of gradients in axonal length, but estimated that the resulting delays were too small, particularly at low CF, and could not account for the BD distribution in cat (Karino et al. 2011). Perhaps gradients in axonal diameter or internodal distance are important variables (Ford et al. 2015; Seidl et al. 2010). However, labeling of physiologically-characterized MSO neurons in gerbil did not support a rostro-caudal map of BDs (Franken et al. 2015). Furthermore, axonal delays are

pure time delays, whereas ITD-functions often express a constant phase-delay between the inputs from the two ears (MSO: Plauška et al. 2016; IC: Yin & Kuwada 1983). This produces asymmetric ITD-functions (**Figure 7a**): the secondary peak on one side of the main peak is smaller than on the other side (unlike the SACs in **Figure 6b**). These asymmetries provide further evidence of other mechanisms than pure time delays. Thus, at present, the central tenet of Jeffress' model, axonal delay-lines producing a map of ITD, no longer appears plausible.

The systematic delays created by the cochlear traveling-wave motivated another proposal for internal delay (Schroeder 1977). If the inputs from the two ears are mismatched in CF, the input from the ear with lower CF incurs a delay (**Figure 8c**). Computational models and a correlation analysis of AN spiketrains tested the plausibility of this scheme (Bonham & Lewis 1999; Joris et al. 2006a; Shamma et al. 1989). Small cochlear disparities are sufficient to generate delays that are significant in binaural terms, and interestingly they scale with CF^{-1} (Joris et al. 2006a) and can account for other features of ITD-sensitivity (Benichoux et al. 2015). A difficulty with this proposal is that, while some "error" in wiring seems plausible, it should be systematic: a bias of contralateral inputs towards lower CF than ipsilateral inputs is needed to generate positive BDs. An examination in gerbil found evidence for a cochlear contribution to BD, but no correlation between mismatches in frequency-tuning and BD (Plauška et al. 2017). Binaural reverse correlation analysis of MSO neurons also revealed tuning mismatches, which moreover were biased towards the contralateral ear (Sayles et al. 2016). Thus, while there is suggestive evidence that cochlear disparities contribute to BD (see also Benichoux et al. 2015; Day & Semple 2011), the weight of that contribution is unclear.

Inhibition was proposed as a source of internal delay (Batra et al. 1997; Brand et al. 2002): contralateral inhibition would delay contralateral excitation (**Figure 8d**), which would generate a within-cycle phase shift rather than a time delay. Modeling (Day & Semple 2011; Zhou et al. 2005) and in-vitro physiology (Myoga et al. 2014; Roberts et al. 2014) showed that the effects are too small to explain the range of BDs observed, and in-vivo juxtacellular and whole-cell recordings do not support an effect of leading inhibition (Franken et al. 2015; van der Heijden et al. 2013). Also, inhibitory terminals are heavily concentrated on MSO somata (Kapfer et al. 2002) so that any delaying or advancing effects would affect both ipsi- and contralateral inputs. There have been several suggestions for sources of delay intrinsic to MSO neurons (**Figure 8e**). Their axon can have a dendritic rather than somatic origin (Smith 1995), which could introduce a differential delay (Zhou et al. 2005). However, this structural asymmetry is only present in a minority of cells (Rautenberg et al. 2009), and the predicted asymmetry in EPSP-AP latency was not observed in-vivo (van der Heijden et al. 2013). An asymmetry in rise-time of contralaterally vs. ipsilaterally evoked EPSPs was proposed to introduce a differential delay (Jercog et al. 2010), but the asymmetry was not replicated in-vitro (Roberts et al. 2014) and in-vivo (Franken et al. 2015; van der Heijden et al. 2013). Franken et al. (2015) found that asymmetries in the temporal pattern of ipsi- and contralateral excitation interacted with membrane properties to generate phase delays. These asymmetries were correlated with CF, but their sources have not been established.

BDs amount to only a fraction of the overall latency between sound stimulus and MSO response (~4 ms). It is interesting that the diverse mechanisms examined give rather underwhelming delays. This easily leads to floccinaucinihilipilification of the respective

mechanisms, but is at once an indication of the optimization in this circuit to tightly control timing. There is the unalluring but real possibility that a combination of factors is at work, perhaps to different degrees in different neurons. Possibly the circuit is tweaked, phylogenetically and/or ontogenetically, to limit BDs to a behaviorally relevant range, without there being one specific and dominant mechanism for internal delay.

8.2. Coincidence detection in MSO

Coincidence detection was an early concept in models implying neural multiplication (referred to as “summation” in Jeffress 1948; Licklider 1951; McCulloch & Pitts 1943). Initial evidence for coincidence detection came from comparison of spike output to monaural and binaural stimulation. Goldberg and Brown (1969) showed that the difference in phase of monaural responses to tones matched the ITD of maximal binaural response. Yin and Chan (1980) extended this analysis to noise and amplitude-modulated tones. However, comparison of the output of the same neuron to monaural and binaural stimulation is not the same as comparing input and output. Such analysis became possible with methods enabling the recording of both subthreshold and spike events (Franken et al. 2015; van der Heijden et al. 2013). **Figure 9** shows an in-vivo whole-cell MSO recording. Ipsi- and contralateral monaural stimulation gives rise to brief, “spike-like”, subthreshold EPSPs phase-locked to the tonal stimulus, sometimes triggering small action potentials. With the stimuli combined into a binaural beat, the in-phase part of the stimulus evokes large and well-phase-locked EPSPs which trigger spikes, and the anti-phase part evokes non-coincident EPSPs which remain subthreshold.

Figure 9

Juxtacellular in-vivo recordings of gerbil MSO showed that simple linear summation of subthreshold events well-predicted spike output to tones (van der Heijden et al. 2013) and broadband stimuli (Plauška et al. 2016). However, whole-cell recordings in the same species revealed discrepancies between the timing of EPSPs and BDs (Franken et al. 2015). These were traced to subtle asymmetries in the pattern of activation to ipsi- and contralateral input which, in interaction with the cell's membrane properties, cause ordering effects that generate a phase-shift in ITD-sensitivity.

SIDEBAR 3

9. The LSO as a complementary ITD-processor

Neuroscience textbooks invariably refer to LSO as the ILD-processor. While MSO-neurons perform a multiplication-type operation, LSO neurons are typically characterized as subtractors (reviewed by Tollin 2003). Because the two ears have opposite synaptic effects, the firing of LSO neurons indeed signals whether sound intensity is larger at the ipsilateral (high firing rate) or contralateral ear (no firing). There are several reasons to suspect that such characterization of LSO is incomplete (reviewed by Joris and Trussell, in press). The assumption that LSO only processes ILDs is difficult to square with many specializations of this circuit (**Figure 5b**) that suggest a role in timing, such as the calyx of Held. As mentioned, behavioral ITD-sensitivity extends to high frequencies based on stimulus envelope (cf. **Figure 7b**), so the processing of envelope-ITDs may seem a plausible *raison d'être* of the temporal specializations. Indeed LSO neurons show ITD-sensitivity to amplitude-modulated high-frequency sounds and even some sensitivity to ITDs of fine-structure at low frequencies (Finlayson & Caspary 1991; Joris & Yin

1995; Tollin & Yin 2002, 2005), but in both cases ITD-sensitivity is rather weak and easily swamped by ILDs, which seems incommensurate with the extraordinary specializations such as the calyx. The only case in which LSO neurons show a steep ITD-dependence over a range of ILDs, is to stimulus transients such as clicks (Irvine et al. 2001; Joris & Yin 1995). Recent data (Franken et al. 2016b) confirm steep ITD-sensitivity to clicks in LSO but not in MSO neurons. Transients may thus be the only temporal stimulus feature for which ITD-sensitivity is superior in LSO compared to MSO. Joris & Trussell (in press) propose that the behavioral relevance of this sensitivity is in the spatial lateralization of transients produced adventitiously during animal locomotion (rustling sounds).

Recent data (Franken et al. 2018) add a surprising twist. The LSO is less homogenous than the MSO: besides principal cells, which constitute the bulk of the nucleus, other morphological classes are distinguished (Helfert & Schwartz 1986). By labeling physiologically-characterized cells, it was found that principal cells do not generate the classical sustained (“chopper”) pattern of response, associated with the LSO since the earliest single-unit studies (Boudreau & Tsuchitani 1968; Guinan et al. 1972). Rather, principal cells are MSO-like, with similar membrane features (short time constants, short PSPs, small action potentials). These cells generate onset-responses to pure tones and appear to be the neurons with acute ITD-sensitivity to transients. Franken et al. (2018) suggest that traditional extracellular recording methods, used in previously studies, biased against recording from principal cells.

10. Summary and outlook

Great strides have been made in the past two decades in the study of the initial temporal binaural interactions in the brainstem. A variety of new techniques, both in-vivo and in-vitro, have yielded much refined knowledge of these reticent nuclei, and there is a healthy variety of hypotheses and models to drive further experimentation. In the near future, technical advances will further refine morphological resolution of the circuit and allow its optical manipulation towards both mechanistic and functional questions. Increasing our knowledge on these nuclei is not just of academic interest: deficient ITD-sensitivity is a key problem in the hearing-impaired, for reasons that are only partly understood (Chung et al. 2016).

Acknowledgements

We thank NC and CD for their award-winning hospitality during the preparation of this review. Our research is currently supported by FWO (Research Foundation-Flanders) G0B2917N, and BOF (Bijzonder Onderzoeksfonds KU Leuven) OT-14-118 (PXJ); and NWO (Netherlands Organization for Scientific Research), ALW 824.15.008 (MvdH). We dedicate this review to our contra- and ipsilateral inputs, Tom Yin and Tino Trahiotis.

Literature Cited

- Agmon-Snir H, Carr CE, Rinzel J. 1998. The role of dendrites in auditory coincidence detection. *Nature*. 393:268–72
- Banks MI, Smith PH. 1992. Intracellular recordings from neurobiotin-labeled cells in brain slices of the rat medial nucleus of the trapezoid body. *J Neurosci*. 12:2819–37

- Batra R, Kuwada S, Fitzpatrick DC. 1997. Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. *J Neurophysiol.* 78:1222–36
- Beckius GE, Batra R, Oliver DL. 1999. Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci.* 19(8):3146–61
- Benichoux V, Fontaine B, Franken TP, Karino S, Joris PX, Brette R. 2015. Neural tuning matches frequency-dependent time differences between the ears. *eLife.* 4:e06072
- Bernstein LR. 2001. Auditory processing of interaural timing information: new insights. *J. Neurosci. Res.* 66(6):1035–46
- Bernstein LR, Trahiotis C. 2018. Effects of interaural delay, center frequency, and no more than “slight” hearing loss on precision of binaural processing: Empirical data and quantitative modeling. *J. Acoust. Soc. Am.* 144(1):292
- Blodgett HC, Wilbanks WA, Jeffress LA. 1956. Effect of Large Interaural Time Differences upon the Judgment of Sidedness. *J. Acoust. Soc. Am.* 28(4):639–43
- Bonham BH, Lewis ER. 1999. Localization by interaural time difference (ITD): effects of interaural frequency mismatch. *J Acoust Soc Am.* 106:281–90
- Boudreau JC, Tsuchitani C. 1968. Binaural interaction in the cat superior olive S Segment. *J Neurophysiol.* 31:442–54
- Brand A, Behrend O, Marquardt T, McAlpine D, Grothe B. 2002. Precise inhibition is essential for microsecond interaural time difference coding. *Nature.* 417:543–47
- Bremen P, Joris PX. 2013. Axonal recordings from medial superior olive neurons obtained from the lateral lemniscus of the chinchilla (*Chinchilla laniger*). *J. Neurosci.* 33(44):17506–18
- Brette R. 2010. On the interpretation of sensitivity analyses of neural responses. *J. Acoust. Soc. Am.* 128(5):2965–72

- Brown C, May BJ. 2005. Comparative Mammalian Sound Localization. In *Sound Source Localization*, eds. A Popper, R Fay, pp. 124–78. Springer
- Cant NB. 1984. The Fine Structure of the Lateral Superior Olivary Nucleus of the Cat. *J Comp Neurol.* 227:63–77
- Cant NB, Casseday JH. 1986. Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei. *J. Comp. Neurol.* 247(4):457–76
- Cant NB, Hyson RL. 1992. Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil. *Hear Res.* 58:26–34
- Chung Y, Hancock KE, Delgutte B. 2016. Neural Coding of Interaural Time Differences with Bilateral Cochlear Implants in Unanesthetized Rabbits. *J. Neurosci.* 36(20):5520–31
- Clark GM. 1969. Vesicle shape versus type of synapse in the nerve endings of the cat medial superior olive. *Brain Res.* 15(2):548–51
- Coffey CS, Ebert CS Jr, Marshall AF, Skaggs JD, Falk SE, et al. 2006. Detection of interaural correlation by neurons in the superior olivary complex, inferior colliculus and auditory cortex of the unanesthetized rabbit. *Hear. Res.* 221(1–2):1–16
- Colburn HS, Durlach NI. 1978. Models of binaural interaction. In *Handbook of Perception*, pp. 467–518. New York: Academic Press
- Colburn HS, Han YA, Cullota CP. 1990. Coincidence model of MSO responses. *Hear Res.* 49:335–46
- Colburn HS, Isabelle SK. 2001. Physiologically based models of binaural detection. In *Physiological and Psychophysical Bases of Auditory Function*, pp. 142–48. Maastricht: Shaker Publishing BV
- Day ML, Koka K, Delgutte B. 2012. Neural encoding of sound source location in the presence of a concurrent, spatially separated source. *J. Neurophysiol.* 108(9):2612–28
- Day ML, Semple MN. 2011. Frequency-dependent interaural delays in the medial superior olive: implications for interaural cochlear delays. *J. Neurophysiol.* 106(4):1985–99

- Devore S, Delgutte B. 2010. Effects of reverberation on the directional sensitivity of auditory neurons across the tonotopic axis: influences of interaural time and level differences. *J Neurosci.* 30(23):7826–37
- Durlach NI. 1972. Binaural signal detection: equalization and cancellation theory. In *Foundations of Modern Auditory Theory*, ed. JV Tobias, pp. 371–462
- Durlach NI, Colburn HS. 1978. Binaural phenomena. In *Handbook of Perception*, pp. 365–465. New York: Academic Press
- Finlayson PG, Caspary DM. 1991. Low-frequency neurons in the lateral superior olive exhibit phase-sensitive binaural inhibition. *J Neurophysiol.* 65:598–605
- Fitzpatrick DC, Batra R, Stanford TR, Kuwada S. 1997. A neuronal population code for sound localization. *Nature.* 388:871–74
- Ford MC, Alexandrova O, Cossell L, Stange-Marten A, Sinclair J, et al. 2015. Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat. Commun.* 6:8073
- Franken TP, Bremen P, Joris PX. 2014. Coincidence detection in the medial superior olive: mechanistic implications of an analysis of input spiking patterns. *Front. Neural Circuits.* 8:42
- Franken TP, Joris PX, Smith PH. 2018. Principal cells of the brainstem’s interaural sound level detector are temporal differentiators rather than integrators. *eLife.* 7:
- Franken TP, Roberts MT, Wei L, Golding NL, Joris PX. 2015. In vivo coincidence detection in mammalian sound localization generates phase delays. *Nat. Neurosci.* 18(3):444–52
- Franken TP, Smith PH, Joris PX. 2016a. In vivo Whole-Cell Recordings Combined with Electron Microscopy Reveal Unexpected Morphological and Physiological Properties in the Lateral Nucleus of the Trapezoid Body in the Auditory Brainstem. *Front. Neural Circuits.* 10:69

- Franken TP, Smith P, Joris P. 2016b. In vivo whole-cell recordings of the lateral and medial superior olive to interaural time differences of transients. *Assoc Res Otolaryngol Abs.* 39:451
- Gabriel KG, Colburn HS. 1981. Interaural correlation discrimination: I. Bandwidth and level dependence. *J Acoust Soc Am.* 69:1394–1401
- Goldberg JM, Brown PB. 1969. Response of Binaural Neurons of Dog Superior Olivary Complex to Dichotic Tonal Stimuli: Some Physiological Mechanisms of Sound Localization. *J Neurophysiol.* 22:613–36
- Golding NL, Oertel D. 2012. Synaptic integration in dendrites: exceptional need for speed. *J. Physiol.* 590(22):5563–69
- Gómez-Álvarez M, Saldaña E. 2016. Different tonotopic regions of the lateral superior olive receive a similar combination of afferent inputs. *J. Comp. Neurol.* 524(11):2230–50
- Goodman DFM, Benichoux V, Brette R. 2013. Decoding neural responses to temporal cues for sound localization. *eLife.* 2:e01312
- Guinan JJ, Norris BE, Guinan SS. 1972. Single auditory units in the superior olivary complex. II: Locations of unit categories and tonotopic organization. *Int. J. Neurosci.* 4:147–66
- Hancock KE. 2007. A physiologically-based population rate code for interaural time differences (ITDs) predicts bandwidth-dependent lateralization. In *Hearing - from Sensory Processing to Perception*, eds. B Kollmeier, G Klump, V Hohmann, U Langemann, M Mauermann, et al., pp. 389–97. Berlin: Springer-Verlag
- Hancock KE, Delgutte B. 2004. A physiologically based model of interaural time difference discrimination. *J Neurosci.* 24:7110–17
- Harper NS, McAlpine D. 2004. Optimal neural population coding of an auditory spatial cue. *Nature.* 430(7000):682–86

- Harrison JM, Irving R. 1966. Visual and nonvisual auditory systems in mammals. Anatomical evidence indicates two kinds of auditory pathways and suggests two kinds of hearing in mammals. *Science*. 154(3750):738–43
- Heffner RS, Heffner HE. 1992. Visual factors in sound localization in mammals. *J. Comp. Neurol.* 317(3):219–32
- Helfert RH, Schwartz IR. 1986. Morphological evidence for the existence of multiple neuronal classes in the cat lateral superior olivary nucleus. *J Comp Neurol.* 244:533–49
- Henning GB. 1974. Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am.* 55:84–90
- Henning GB. 1980. Some observations on the lateralization of complex waveforms. *J Acoust Soc Am.* 68:446–54
- Hirsh IJ. 1948. Binaural summation and interaural inhibition as a function of the level of masking noise. *Am. J. Psychol.* 61(2):205–13
- Houtgast T, Steeneken HJM. 1985. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J Acoust Soc Am.* 77(3):1069–77
- Irvine DR, Park VN, McCormick L. 2001. Mechanisms underlying the sensitivity of neurons in the lateral superior olive to interaural intensity differences. *J Neurophysiol.* 86(6):2647–66
- Jeffress LA. 1948. A Place Theory of Sound Localization. *J Comp Physiol Psychol.* 41:35–39
- Jercog PE, Svirskis G, Kotak VC, Sanes DH, Rinzel J. 2010. Asymmetric excitatory synaptic dynamics underlie interaural time difference processing in the auditory system. *PLoS Biol.* 8(6):e1000406
- Johnson DH. 1980. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am.* 68:1115–22
- Joris P, Trussell LO. The calyx of Held: a hypothesis on the need for reliable timing in an intensity difference encoder. *Neuron*. in press

- Joris PX. 1996. Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J Neurophysiol.* 76(4):2137–56
- Joris PX. 2003. Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci.* 23(15):6345–50
- Joris PX, Carney LH, Smith PH, Yin TC. 1994a. Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol.* 71(3):1022–36
- Joris PX, Smith PH. 2008. The volley theory and the spherical cell puzzle. *Neuroscience.* 154(1):65–76
- Joris PX, Smith PH, Yin TC. 1994b. Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J Neurophysiol.* 71(3):1037–51
- Joris PX, Smith PH, Yin TC. 1998. Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron.* 21(6):1235–38
- Joris PX, Van de Sande B, Louage DH, van der Heijden M. 2006a. Binaural and cochlear disparities. *Proc Natl Acad Sci U A.* 103(34):12917–22
- Joris PX, van de Sande B, Recio-Spinoso A, van der Heijden M. 2006b. Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. *J Neurosci.* 26(1):279–89
- Joris PX, Yin TC. 1995. Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J Neurophysiol.* 73(3):1043–62
- Julesz B. 1971. *Foundations of Cyclopean Perception.* Chicago: University of Chicago Press
- Kapfer C, Seidl AH, Schweizer H, Grothe B. 2002. Experience-dependent refinement of inhibitory inputs to auditory coincidence-detector neurons. *Nat Neurosci.* 5(3):247–53
- Karino S, Smith PH, Yin TCT, Joris PX. 2011. Axonal branching patterns as sources of delay in the mammalian auditory brainstem: a re-examination. *J. Neurosci.* 31(8):3016–31

- Klumpp RG, Eady HR. 1956. Some Measurements of Interaural Time Difference Thresholds. *J. Acoust. Soc. Am.* 28(5):859–60
- Kuwabara N, Zook JM. 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(4):522–38
- Licklider JCR. 1948. The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am.* 20:150–59
- Licklider JCR. 1951. A duplex theory of pitch perception. *Experientia.* 7(4):128–34
- Louage DH, Joris PX, van der Heijden M. 2006. Decorrelation sensitivity of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise. *J Neurosci.* 26(1):96–108
- Masterton B, Thompson GC, Bechtold JK, RoBards MJ. 1975. Neuroanatomical basis of binaural phase-difference analysis for sound localization: a comparative study. *J. Comp. Physiol. Psychol.* 89(5):379–86
- Masterton RB, Imig TJ. 1984. Neural mechanisms for sound localization. *ARP.* 46:275–87
- Mc Laughlin M, Verschooten E, Joris PX. 2010. Oscillatory dipoles as a source of phase shifts in field potentials in the mammalian auditory brainstem. *J Neurosci.* 30(40):13472–87
- McAlpine D, Jiang D, Palmer A. 1996. Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. *Hear Res.* 97:136–52
- McAlpine D, Jiang D, Palmer A. 2001. A neural code for low-frequency sound localization in mammals. *Nat Neurosci.* 4:396–401
- McCulloch WS, Pitts W. 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5(4):115–33
- McFadden D, Pasanen EG. 1976. Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am.* 59:634–39

- Middlebrooks JC, Green DM. 1991. Sound localization by human listeners. *Annu. Rev. Psychol.* 42:135–59
- Mills AW. 1958. On the Minimum Audible Angle. *J. Acoust. Soc. Am.* 30(4):237–46
- Moore JK, Moore RY. 1971. A comparative study of the superior olivary complex in the primate brain. *Folia Primatol. Int. J. Primatol.* 16(1):35–51
- Myoga MH, Lehnert S, Leibold C, Felmy F, Grothe B. 2014. Glycinergic inhibition tunes coincidence detection in the auditory brainstem. *Nat. Commun.* 5:3790
- Nuetzel JM, Hafter ER. 1981. Discrimination of interaural delays in complex waveforms: Spectral effects. *J Acoust Soc Am.* 69:1112–18
- Pecka M, Brand A, Behrend O, Grothe B. 2008. Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *J Neurosci.* 28(27):6914–25
- Plauška A, Borst JGG, van der Heijden M. 2016. Predicting binaural responses from monaural responses in the gerbil medial superior olive. *J. Neurophysiol.* 115(6):2950–63
- Plauška A, van der Heijden M, Borst JGG. 2017. A Test of the Stereausis Hypothesis for Sound Localization in Mammals. *J. Neurosci.* 37(30):7278–89
- Ramón y Cajal S. 1909. *Histologie du système nerveux de l'homme & des vertébrés*. Paris : Maloine
- Rautenberg PL, Grothe B, Felmy F. 2009. Quantification of the three-dimensional morphology of coincidence detector neurons in the medial superior olive of gerbils during late postnatal development. *J. Comp. Neurol.* 517(3):385–96
- Rayleigh (J.W. Strutt). 1907. On our perception of sound direction. *Philos Mag.* 13:214–32
- Roberts MT, Seeman SC, Golding NL. 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

- Rose JE, Brugge JF, Anderson DJ, Hind JE. 1967. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol.* 30:769–93
- Rose JE, Gross NB, Geisler CD, Hind JE. 1966. Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *J Neurophysiol.* 29:288–314
- Roth GL, Kochhar RK, Hind JE. 1980. Interaural Time Differences: Implications Regarding the Neurophysiology of Sound Localization. *J Acoust Soc Am.* 68:1643–51
- Sayles M, Smith PH, Joris PX. 2016. Inter-aural Time Sensitivity of Superior-olivary-complex Neurons is Shaped by Systematic Cochlear Disparities. *Assoc Res Otolaryngol Abs.* 39:273
- Schroeder MR. 1977. New viewpoints in binaural interactions. In *Psychophysics and Physiology of Hearing*, pp. 455–67. New York: Academic Press
- Scott BH, Malone BJ, Semple MN. 2007a. Effect of behavioral context on representation of a spatial cue in core auditory cortex of awake macaques. *J. Neurosci.* 27(24):6489–99
- Scott LL, Hage TA, Golding NL. 2007b. Weak action potential backpropagation is associated with high-frequency axonal firing capability in principal neurons of the gerbil medial superior olive. *J Physiol.* 583(Pt 2):647–61
- Seidl AH, Rubel EW, Harris DM. 2010. Mechanisms for adjusting interaural time differences to achieve binaural coincidence detection. *J Neurosci.* 30(1):70–80
- Shackleton TM, Arnott RH, Palmer AR. 2005. Sensitivity to interaural correlation of single neurons in the inferior colliculus of guinea pigs. *J Assoc Res Otolaryngol.* 6(3):244–59
- Shamma SA, Shen NM, Gopaldaswamy P. 1989. Stereausis: binaural processing without neural delays. *J. Acoust. Soc. Am.* 86(3):989–1006
- Shneiderman A, Henkel CK. 1985. Evidence of collateral axonal projections to the superior olivary complex. *Hear Res.* 19:199–205

- Smith PH. 1995. Structural and functional differences distinguish principal from nonprincipal cells in the guinea pig MSO slice. *J Neurophysiol.* 73:1653–67
- Smith PH, Joris PX, Carney LH, Yin TC. 1991. Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *J Comp Neurol.* 304(3):387–407
- Smith PH, Joris PX, Yin TC. 1993. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J Comp Neurol.* 331(2):245–60
- Smith PH, Joris PX, Yin TC. 1998. Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. *J Neurophysiol.* 79(6):3127–42
- Spangler KM, Warr WB, Henkel CK. 1985. The Projections of Principal Cells of the Medial Nucleus of the Trapezoid Body in the Cat. *J Comp Neurol.* 238:249–62
- Spirou GA, Berrebi AS. 1996. Organization of ventrolateral periolivary cells of the cat superior olive as revealed by PEP-19 immunocytochemistry and Nissl stain. *J. Comp. Neurol.* 368(1):100–120
- Spirou GA, Berrebi AS. 1997. Glycine immunoreactivity in the lateral nucleus of the trapezoid body of the cat. *J. Comp. Neurol.* 383(4):473–88
- Spirou GA, Brownell WE, Zidanic M. 1990. Recordings From Cat Trapezoid Body and HRP Labeling of Globular Bushy Cell Axons. *J Neurophysiol.* 63:1169–90
- Stern RM, Zeiberg AS, Trahiotis C. 1988. Lateralization of complex binaural stimuli: A weighted-image model. *J Acoust Soc Am.* 84:156–65
- Stotler WA. 1953. An experimental study of the cells and connections of the superior olivary complex of the cat. *J Comp Neurol.* 98:401–32
- Tollin DJ. 2003. The Lateral Superior Olive: A Functional Role in Sound Source Localization. *The Neuroscientist.* 9(2):127–43

- Tollin DJ, Yin TC. 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(4):1468–79
- Tollin DJ, Yin TC. 2005. Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *J Neurosci.* 25(46):10648–57
- Tolnai S, Beutelmann R, Klump GM. 2018. Interaction of interaural cues and their contribution to the lateralisation of Mongolian gerbils (*Meriones unguiculatus*). *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* 204(5):435–48
- Tsuchitani C, Boudreau JC. 1966. Single unit analysis of cat superior olive S segment with tonal stimuli. *J. Neurophysiol.* 29(4):684–97
- van de Par S, Trahiotis C, Bernstein LR. 2001. A consideration of the normalization that is typically included in correlation-based models of binaural detection. *J Acoust Soc Am.* 109:830–33
- van der Heijden M, Joris PX. 2010. Interaural correlation fails to account for detection in a classic binaural task: dynamic ITDs dominate NOSpi detection. *J Assoc Res Otolaryngol.* 11(1):113–31
- van der Heijden M, Louage DH, Joris PX. 2011. Responses of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise: implications for the sensitivity to interaural delays. *J Assoc Res Otolaryngol.* 12(4):485–502
- van der Heijden M, Trahiotis C. 1999. Masking with interaurally delayed stimuli: the use of “internal” delays in binaural detection. *J Acoust Soc Am.* 105:388–99
- van der Heijden M, Lorteije JAM, Plauška A, Roberts MT, Golding NL, Borst JGG. 2013. Directional Hearing by Linear Summation of Binaural Inputs at the Medial Superior Olive. *Neuron.* 78(5):936–48
- Wakeford OS, Robinson DE. 1974. Lateralization of tonal stimuli by the cat. *J. Acoust. Soc. Am.* 55(3):649–52

- Wallach H, Newman EB, Rosenzweig MR. 1949. The precedence effect in sound localization. *Am. J. Psychol.* 62(3):315–36
- Weiss TF, Rose C. 1988. A comparison of synchronization filters in different auditory receptor organs. *Hear Res.* 33:175–80
- Yin TC, Chan JC, Carney LH. 1987. Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J. Neurophysiol.* 58(3):562–83
- Yin TCT, Chan JK. 1990. Interaural Time Sensitivity in Medial Superior Olive of Cat. *J Neurophysiol.* 64:465–88
- Yin TCT, Chan JK, Irvine DRF. 1986. Effects of Interaural Time Delays of Noise Stimuli on Low-Frequency Cells in the Cat's Inferior Colliculus. I. Responses to Wideband Noise. *J Neurophysiol.* 55:280–300
- Yin TCT, Kuwada S. 1983. Binaural Interaction in Low-Frequency Neurons in Inferior Colliculus of the Cat. III. Effects of Changing Frequency. *J Neurophysiol.* 50:1020–42
- Young ED, Oertel D. 2004. Cochlear Nucleus. In *The Synaptic Organization of the Brain*, ed. GM Shepherd, pp. 125–63. Oxford: Oxford University Press
- Zhou Y, Carney LH, Colburn HS. 2005. A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J Neurosci.* 25(12):3046–58
- Zuk N, Delgutte B. 2017. Neural coding of time-varying interaural time differences and time-varying amplitude in the inferior colliculus. *J. Neurophysiol.* 118(1):544–63

Terms and Definitions list:

1. **Interaural Time Delay (ITD):** arrival time difference of a sound at the two ears, depending on angle of incidence
2. **Interaural Level Difference (ILD):** sound level difference between the ears caused by acoustic head shadow
3. **Binaural:** based on combining the information from the two ears
4. **Internal delay:** difference in delay at which left and right ear affect binaural neuron
5. **Best delay:** the stimulus ITD that optimally excites a binaural neuron
6. **Characteristic Frequency (CF):** stimulus frequency to which an auditory neuron is most sensitive
7. **Interaural decorrelation:** deviation from perfect statistical correlation ($\rho=1$) of the waveforms entering the two ears
8. **Masking:** one sound becoming inaudible by the presence of another
9. **Superior Olivary Complex (SOC):** amalgam of brainstem nuclei also containing the neurons of primary binaural interaction
10. **Cochlear Nucleus (CN):** first nucleus of auditory CNS, receives all auditory nerve input
11. **Spherical Bushy Cells (SBC):** projection neurons in CN providing main excitatory input to binaural SOC neurons
12. **Globular Bushy Cells (GBC):** projection neurons in CN that excite SOC nuclei that are inhibitory to binaural neurons

13. **Trapezoid Body (TB):** main output pathway of the part of the CN which contains bushy cell axons
14. **Correlogram:** histogram summarizing the temporal relation between events; discrete version of correlation function
15. **Fine-structure:** finest temporal details of a waveform at the scale of individual cycles
16. **Envelope:** hull of a waveform that tracks its intensity fluctuations on a timescale of multiple cycles
17. **Phase locking:** a neuron's ability to code fine-structure
18. **Vector strength:** metric for degree of phase locking

Related Resources list:

Audio demos

- Binaural beats at low (LFbinbeat.wav) and high (HFbinbeat.wav) frequencies, called out as Supplemental Audio1 and Supplemental Audio2 in caption Figure 3.
- Binaural masking (LowFreqBinMask_high-pi.wav and LowFreqBinMask_low_pi.wav), called out as Supplemental Audios 3 and 4 in caption Figure 4.
- Separate captions for link to the supplemental content are provided in separate file: Supplemental_Audio.docx

Sidebars:

SIDEBAR 1:

Giant synapses

The monaural neurons providing excitation and inhibition to MSO and LSO receive limited convergence of inputs via large axosomatic terminals. This is particularly striking for SBC, GBC, and MNTB neurons, which have similar morphology with a compact “bushy” dendritic tree. We group them as “bushy-type” neurons. SBCs and GBCs receive nerve input via the endbulbs and modified endbulbs of Held, respectively. MNTB principal cells are mono-innervated by GBCs via the calyx of Held, one of the largest synapses in the brain.

SIDEBAR 2:

The Acoustic Chiasm

Projections of MSO and LSO form the “acoustic chiasm” (Masterton & Imig 1984). While the MSO gives excitatory projections to the ipsilateral IC, the LSO makes excitatory projections to the contralateral IC and largely inhibitory projections ipsilaterally (**Figure 5**). From the midbrain upward, neurons are driven by sounds in contralateral space and lesions result in localization deficits in contralateral space (Champoux et al. 2007; Jenkins & Masterton 1982; Litovsky et al. 2002).

SIDEBAR 3:

Cellular Specializations

Recent in-vitro studies (reviewed by Golding & Oertel 2012) uncovered the extraordinary cellular specializations underlying coincidence detection at the time scale needed in MSO. Models drew attention to the need for a mechanism to reduce monaural “autocoincidences” (Colburn et al. 1990; Franken et al. 2014): this is thought to be a dendritic function tied to the

segregation of ipsi- and contralateral afferents to opposite dendrites (Agmon-Snir et al. 1998; Golding & Oertel 2012). EPSPs are shaped by a low-voltage-activated K^+ conductance to maintain a short time course while propagating from dendrites to soma. Large resting conductances (g_{KL} and g_h) make the cells leaky and decrease the membrane time constant. The neurons are electrically compartmentalized so that action potentials show poor backpropagation from the axon to the soma – which contributes to the difficulties in extracellular recording of spikes.

Figure Captions

Figure 1: Schematic illustration of the transformation from sound waveforms to neural excitation. *First row*: sound waveforms of (a) a low-frequency tone, (b) high-frequency tone, (c) wideband noise. *Second row*: the corresponding waveforms filtered by the cochlea. (d,e) low- and high-frequency tones at their own characteristic cochlear place; (f,g) the same wideband noise at apical and basal cochlear positions, respectively. *Third row*: timing of action potentials in the auditory nerve (AN), shown as peri-stimulus time histogram (*top*) and corresponding dot raster (*bottom*). The dots represent the timing of action potentials evoked by repeated presentations of the stimuli in the panel above. Responses to low frequencies (h,j) show fine-structure (“phase locking”), while the responses to high frequencies (i,k) only show envelope coding.

Figure 2: ITD is not a straightforward localization cue. (a) ITD depends on azimuth, but also on frequency, causing the components of one complex sound to attain different ITDs. (Modified from Durlach & Colburn 1978). (b) Competing sound sources cause ITD to fluctuate in time. Two sources occur in free-field at either side of the listener’s head (± 90 deg). The “signal” (near the left ear) is 20 dB above the “noise” (near the right ear). A representative snippet of the resulting waveforms at the two ears (middle) is shown, together with its instantaneous time-varying ITD (bottom). Note the excursions of ITD well exceeding the headwidth (grey line). (c) Statistical distribution of short-term ITD as shown in b for three signal-to-noise ratios. (b,c) Signal and noise were 100-Hz-wide noise bands centered at 300 Hz.

Figure 3. Different types of ITD cues in human psychophysics. Aspects of ITD that always occur together in natural conditions can be teased apart using artificial stimuli: (a) ongoing ITDs in the fine structure; (b) ongoing ITDs in the envelope; (c) onset and offset ITDs. Relative sensitivity to different ITD-types is strongly frequency-dependent: fine-structure ITD dominates at low frequencies, but becomes imperceivable above 1500 Hz (see text). ITD types in a-c are illustrated using amplitude-modulated tones, but also apply to other stimuli. (d) Binaural beats provide a striking demonstration of the frequency-dependence of binaural fine-structure cues. Two tones of slightly differing frequency are presented to the two ears, causing a running interaural phase difference. When listening over headphones, they evoke a periodically varying spatial percept with low-frequency tones (e.g., 500/508 Hz, **Supplemental Audio 1**), but not with high-frequency tones (e.g. 2000/2008 Hz, **Supplemental Audio 2**).

Supplemental
Audio 1 and 2

Figure 4: Binaural unmasking of a tone in noise. (a) The listener is unable to detect a soft tone (green wave, scaled up 7-fold) in the presence of a noise masker (black wave) when noise and tone waveforms are presented to the two ears with identical polarity. (b) When the tone, but not the noise masker, is presented antiphasically (polarity-reversed in one ear), it becomes audible (**Supplemental Audios 3 and 4**).

Supplemental
Audio 3 and 4

Figure 5. The two primary circuits of binaural interaction. TB and LL are fiber tracts. Red indicates inhibitory projections; black excitatory projections. The projection of LSO to ipsilateral IC is predominantly inhibitory (red dashed line).

Figure 6. Enhancement of temporal coding from the AN to bushy cells, both tuned near 500 Hz. (a) Response to a 460-Hz tone, repeated 50 times. Dot rasters represent the temporal firing

pattern; red dots represent spikes occurring at the same instant (within 50 μ s) in at least one other stimulus repetition. Phase locking in the AN (upper panels) is manifest in the vertical dot alignment and the unimodal cycle histograms. The limited temporal precision of AN firing is reflected by the imperfect dot alignment, the large fraction of “unmatched spikes” (black dots), the shallow-peaked histogram and the vector strength of 0.76 (see text). The enhanced temporal coding by bushy cells (lower panels) is evident from the near-perfect dot alignment, complete dominance of “matched spikes” (red dots), sharply peaked histogram and vector strength of 0.98. (b) Responses to wideband noise show the same enhancement of temporal precision, again reflected by improved alignment and dominance of matched spikes. Cycle histograms and vector strength cannot be defined for a noise stimulus, but normalized shuffled autocorrelograms (SACs, see text) provide a quantitative analysis of temporal precision. The SAC of the bushy-cell response (bottom) has a much narrower and higher central peak than the SAC of the AN response (top). Modified from Joris et al. (2006).

Figure 7. Sensitivity to ITDs of a noise stimulus, in IC (*a,b*), MSO (*c*), and LSO (*d*). Left panels show low-frequency neurons (*a*: 405 Hz, *c*: 780 Hz); right panels show high-frequency neurons (*c*: 5220 Hz, *d*: 5040 Hz). Black lines are responses to ITDs of identical (correlated) noise in the two ears; magenta line is for anticorrelated noise; cyan line for uncorrelated noise. The yellow rectangle shows the approximate headwidth. The waveforms are cartoons, for a positive and negative ITD, of vibration at the cochlear place from which the neuron ultimately derives its input. Neurons were recorded in cat (*a,b,d*) and gerbil (*c*). Modified from Joris (2003)(*a,b*), Plauska et al. (2017)(*c*), and Joris and Yin (1995)(*d*).

Figure 8. Cartoon of hypothesized sources of internal delay. The trapezoidal shapes represent the basilar membranes of ipsi- and contralateral ear. Circles are MSO neurons with simplified inputs. (a) The observed distribution of internal delays shows a longer effective delay contra- than ipsilaterally, symbolized by a loop, and a larger difference at low (larger loop) than at higher frequencies (shorter loop). (b) Jeffress' (1948) model proposes axonal delay lines creating a range of internal delays at all frequencies. The delay lines would need to be longer at the lowest frequencies. (c) Schroeder (1977) proposed an asymmetry in cochlear innervation causing lower CF (more apical location) of contra- than ipsilateral inputs. (d) Brand et al. (2002) propose delay of the contralateral input by preceding inhibition (red). (e) Some sources of delay are proposed to be intrinsic to the nucleus (see text).

Figure 9. Example traces of a gerbil MSO neuron in response to tones with a 1-Hz binaural beat. Each response shows a 45-ms snippet from a 5-s long recording to tones presented monaurally (a,b) or binaurally (c,d). With monaural stimulation, small action potentials, variable in amplitude, are present at low rate. When presented together, the phase of the tones drifts alternately in-phase (c) and anti-phase (d), and activity varies between high rates (c) and no spiking (d). Data provided by TP Franken.

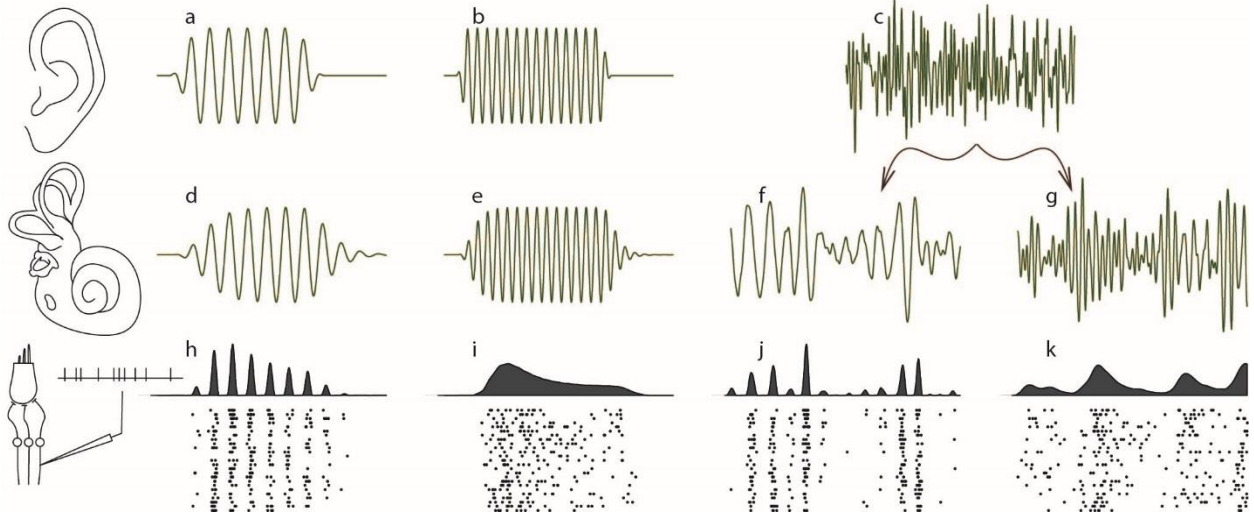


Figure 1

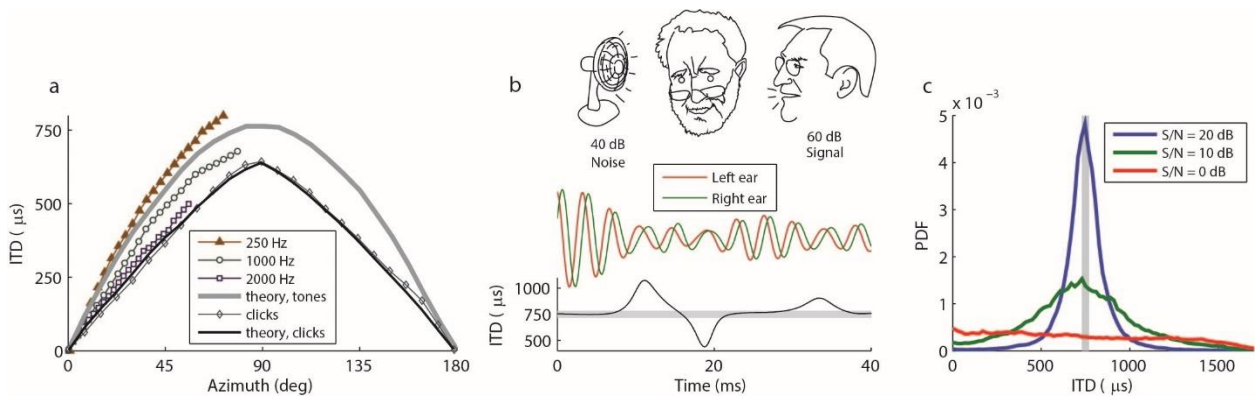


Figure 2

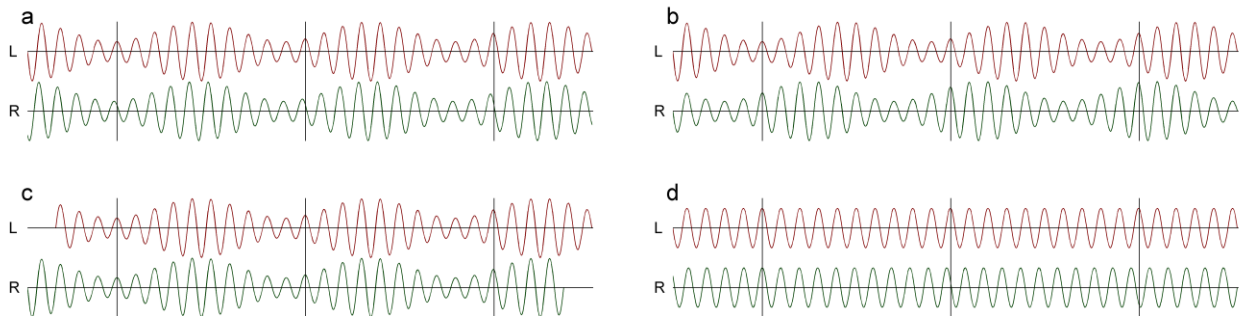


Figure 3

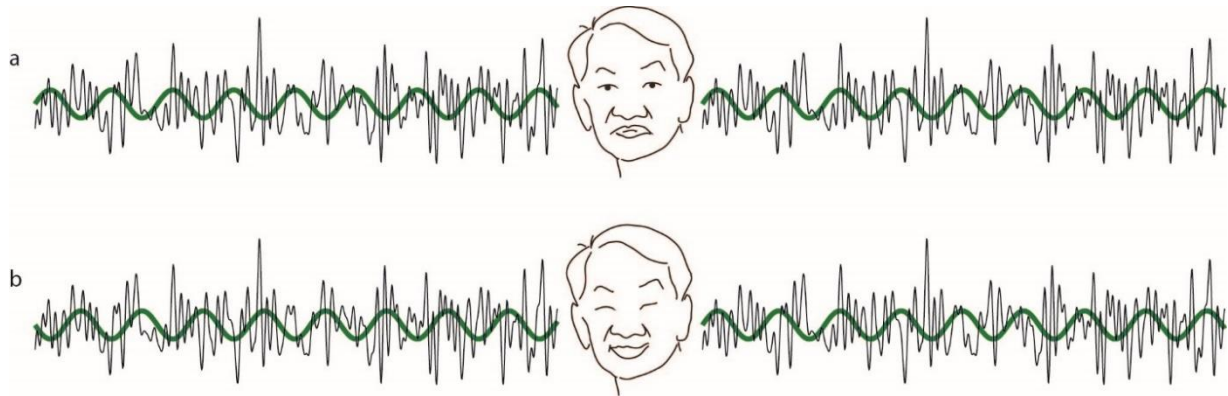


Figure 4

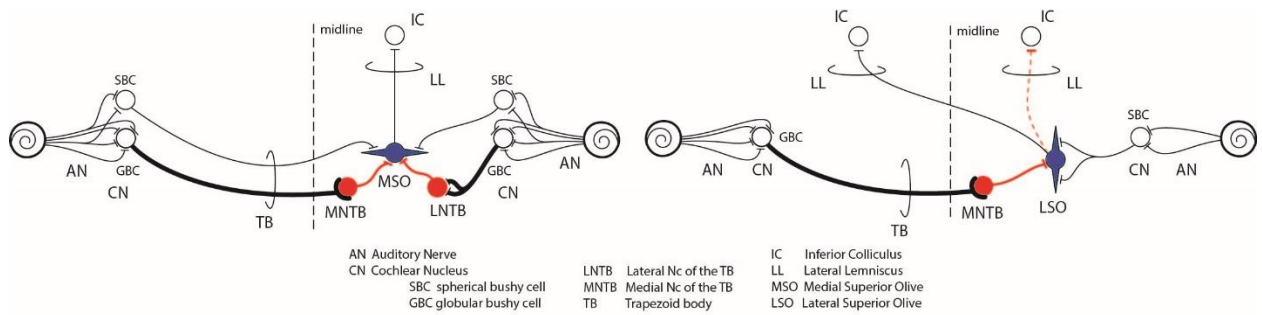


Figure 5

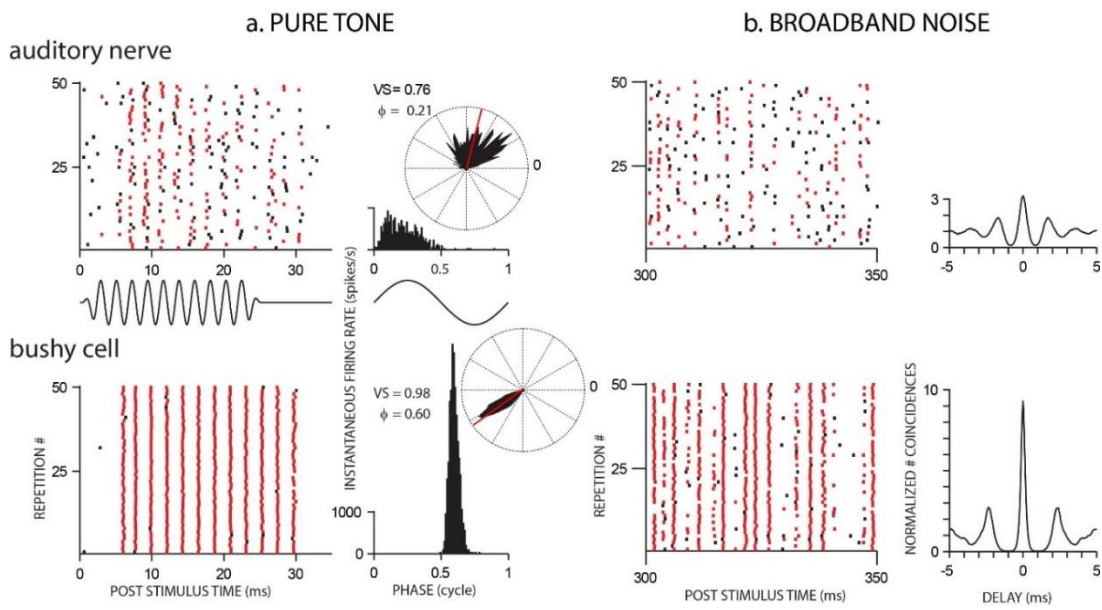


Figure 6

Inferior Colliculus

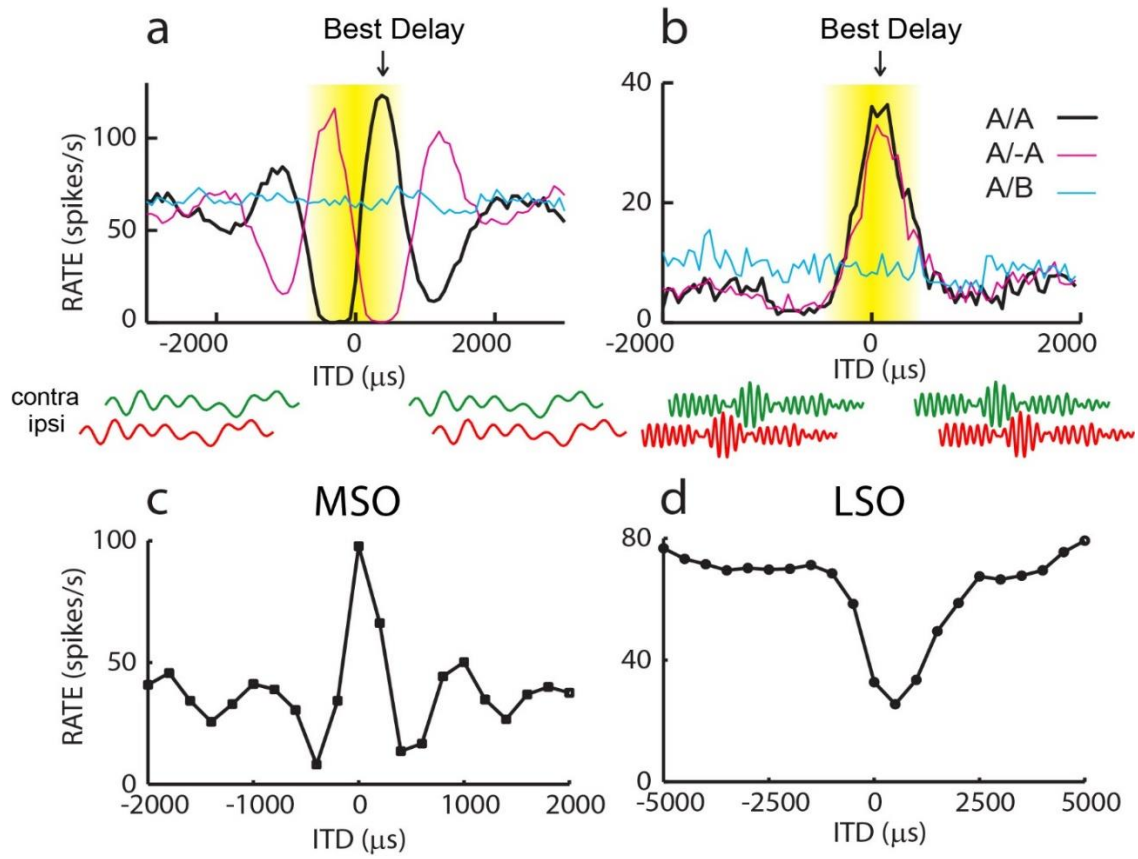


Figure 7

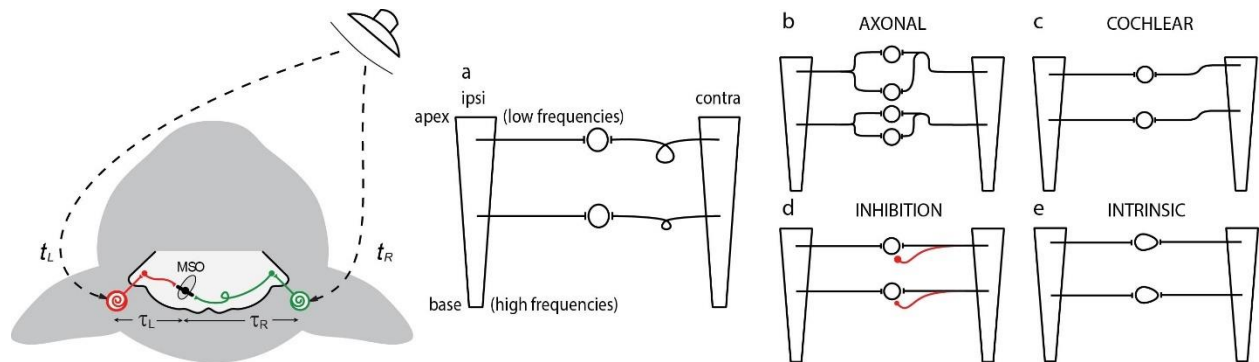


Figure 8

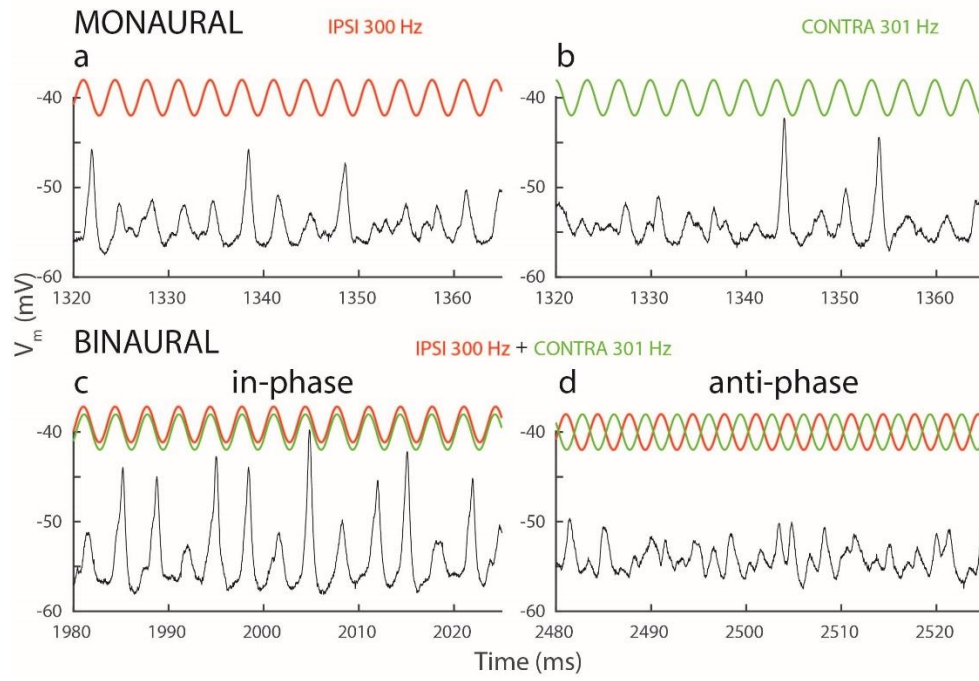


Figure 9