

Sound Source Localization Mechanisms in Gerbil Medial Superior Olive

Andrius Plauška

ISBN: 978-94-6416-421-3

The research described in this thesis was performed at the Department of Neuroscience, Erasmus Medical Center Rotterdam.

Cover art: Inverted image of Golgi-stained medial superior olive neurons in Mongolian Gerbil.

Printing: Ridderprint BV, The Netherlands, www.ridderprint.nl, info@ridderprint.nl

Copyright © 2021 Andrius Plauška

All rights reserved. No parts of this publication may be reproduced, stored in retrieval system or transmitted in any form by any means, electronical, mechanical, photocopying, recording or otherwise without permission of the author or, when appropriate, the scientific journal in which parts of this thesis have been published.

Sound Source Localization Mechanisms in Gerbil
Medial Superior Olive

Hoe de bovenste olijkern in gerbils bijdraagt aan geluidslokalisatie

Thesis

to obtain the degree of Doctor from the

Erasmus University Rotterdam

by command of the

rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Tuesday 9 February 2021 at 10.30 hrs

by

Andrius Plauška

born in Šilutė, Lithuania

Doctoral Committee:

Promotor: Prof.dr. J.G.G. Borst

Other members: Dr. M. Schonewille
Prof.dr. A.J. van Opstal
Prof.dr. A.G. Kohlrausch

Copromotor: Dr. M. van der Heijden

Contents

Chapter 1	7
<i>INTRODUCTION</i>	7
Chapter 2	33
<i>Directional hearing by linear summation of binaural inputs at the medial superior olive</i>	33
Chapter 3	71
<i>Predicting binaural responses from monaural responses in the gerbil medial superior olive</i>	71
Chapter 4	101
<i>A test of the stereausis hypothesis for sound localization in mammals</i>	101
Chapter 5	129
<i>DISCUSSION</i>	129
References	144
Appendix	157
<i>Summary</i>	158
<i>Samenvatting</i>	161
<i>Curriculum Vitae</i>	165
<i>PhD-Portfolio</i>	167
<i>Acknowledgements</i>	168

Chapter 1

INTRODUCTION

For many species, predators and prey alike, sound localization is important for survival. In contrast to a visual stimulus, localizing a sound is often a difficult task. The auditory system uses subtle spectral cues to determine the vertical location of a sound. To localize a sound in the horizontal plane (azimuth), it can use small differences in the arrival times or the intensity of sounds across the two ears. The main focus of this thesis is the circuitry involved in low-frequency sound localization in azimuth. We begin this introduction with a review of the anatomy of the cochlea and the neural circuitry involved in sound localization. These sections will be followed by a more detailed description of sound localization mechanisms and research questions addressed in this thesis.

Sound localization: periphery

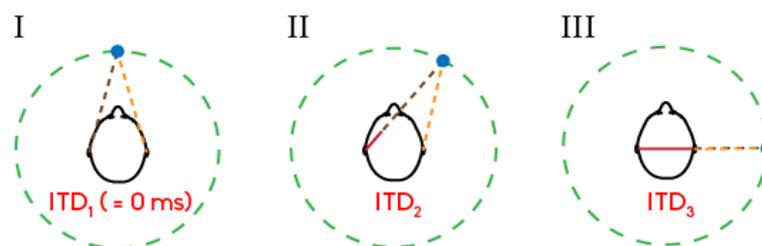


Figure 1. Sound source azimuthal location creates interaural time differences.

A schematic depiction of the relation between sound source and ITDs. If the sound source (blue circle) is in front of the listener (I), the sound will reach left (brown dashed line) and right (orange dashed line) ears at the same time. If the sound source is off-center with respect to the listener (II), the sound will reach one ear before the other (in this case the sound will reach the right ear earlier). The extra time it takes for the sound to reach the other ear (red solid line) is called the interaural time difference (ITD) and is dependent on the sound source location and the head size. In the extreme case when the sound comes from one side (III), the ITD is the largest possible for the given head size; $ITD_3 > ITD_2 > ITD_1 = 0$ (ms). For a human head the distance between both ears is about 24 cm. The speed of sound in air is about 340 ms, yielding a maximal ITD of 0.7 ms. In contrast, in a small rodent such as the gerbil, the head size yields a maximal ITD of not more than 0.13 ms.

The mammalian hearing system can be subdivided into a peripheral and a central part. The peripheral auditory system consists of the outer, middle and inner ear, while the neural pathways that conduct and process sound-related information represent the central auditory system.

The perception of sound begins with the outer ear collecting incoming sound waves. The pinna plays a large role in determining the sound source elevation (Middlebrooks and Green 1991). Sounds coming from different elevations undergo different changes in sound spectrum due to reflections in the pinna and ear canal. These spectral changes are cues that play a major role in monaural sound localization in the vertical plane.

Sound localization in azimuth depends on the sound source position relative to both ears and the dimension of the listener's head. For high-frequency (> 3 kHz) sound waves, the head creates an acoustic shadow for the ear that is further from the sound source. As a result, the sound arriving at the ear that is closer to the source has a larger amplitude than the ear that is further from the source – the origin of the interaural level difference (**ILD**) cue. This cue, however, is less useful for low-frequency sound waves because the attenuation from the head shadow is smaller for low frequencies.

Azimuthal sound localization for low frequencies (< 3 kHz) makes use of differences in travel time of sound to both ears (**Figure 1**). When a sound source in the horizontal plane is directly in front of (or directly behind) the listener, sound waves reach both ears at the same time. In any other case the sound will reach one ear later than the other. This time delay is called an interaural time difference (**ITD**).

Sound waves are guided through the ear canal towards the tympanic membrane, or the eardrum, where the transformation from the pressure wave to mechanical vibrations occurs. These mechanical vibrations are conducted via three ossicles – malleus, incus and stapes, of which the latter is attached to the oval window membrane. This is where the inner ear begins.

The mechanical vibrations arriving at the oval window initiate travelling waves in the fluid-filled cochlea. The basilar membrane (**BM**) is the main structure in the cochlea involved in the frequency analysis of incoming sounds. Different sound frequencies lead to maximal BM displacements at different locations along the cochlea. Higher frequencies preferentially displace the BM closer to the cochlear base, whereas lower frequencies mainly cause

displacements in the apical part. The organ of Corti, which rests on the BM, contains the inner hair cells, which convert the vibrations of the BM into electrical signals. The inner hair cells are innervated by spiral ganglion neurons, which form the auditory nerve (**AN**) that conducts the auditory information in the form of action potentials (**AP**) to the cochlear nucleus in the brainstem.

Duplex theory of azimuthal sound localization

Thompson (1882) and Rayleigh (1907) were the first to recognize that two dominant cues – interaural level differences and interaural time differences can be used to localize sound in azimuth. A lot of experimental support for this so-called duplex theory of sound localization of pure tones has been obtained, but for complex sounds, ITD cues are also used for the localization of high-frequency sound. It is now known that both cues are processed in parallel neural pathways, which involve to some extent different auditory nuclei in the auditory brainstem. Which of these two cues can be used depends on the frequency composition of incoming sounds. Higher frequency tones tend to be obstructed by the head shape more than low-frequency tones; the head creates an acoustic shadow for these short wavelength waves. This results in an ILD cue, a difference in sound intensity between the two ears. On the other hand, the difference in the arrival time of low frequency tones creates a meaningful phase difference (ITD cue). It is worth noting that due to phase locking (see below) ITD cue is not present for high-frequency tones, but is relevant for high frequency complex sounds. At the extremes of the frequency spectrum (either very low or very high frequencies), one of the two cues dominates the sound localization mechanism, while at the mid-frequency range neither of the cues is strong. Thus, the boundary between the utilization of either ITDs or ILDs is not a steep jump between the two, but rather a gradual transition and depends on the dimension of the head. For humans, the frequency demarcation separating the low and high frequencies is between 2 and 3 kHz.

For pure tones, duplex theory has been supported by a large set of psychophysical experiments, which showed *a*) reduced acuity in azimuthal sound localization for mid-frequency sounds (Casseday and Neff 1975; Mills 1958; Stevens and Newman 1936), and *b*) that in humans, ITD detection in the ongoing fine structure of high frequency pure tones is

obscured (Licklider et al. 1950; Mills 1960). One important thing to note is that broadband high-frequency sounds can be localized based on their low-frequency envelopes (Klumpp and Eady 1956; Leakey et al. 1958). For a recent review of binaural processing of temporal information, see Joris and van der Heijden (2019).

Sound localization: neural circuitry

Circuitry in mammals

Throughout this thesis the Mongolian gerbil (*Meriones unguiculatus*) will be used as a model species for low-frequency sound localization in mammals. Gerbils are well known for their low-frequency sensitivity and sound localization capabilities (Ryan 1976) and are widely used in auditory research. **Figure 2** shows a schematic view of the gerbil brainstem to illustrate the auditory nuclei involved in low-frequency sound localization. Auditory nerve fibers terminate in the anteroventral cochlear nucleus (**AVCN**; **CN** – cochlear nucleus in **Figure 2**). Spherical bushy cells (**SBC**) in AVCN provide excitatory glutamatergic inputs to the ipsilateral and contralateral medial superior olive (**MSO**) (Clark 1969a; b; Kil et al. 1995) and the ipsilateral lateral nucleus of the trapezoid body (**LNTB**) (Spirou and Berrebi 1997). Globular bushy cells (**GBC**) in AVCN form excitatory glutamatergic connections with the contralateral medial nucleus of the trapezoid body (**MNTB**) (Kuwabara et al. 1991). Additionally, the MSO receives two glycinergic inhibitory inputs – one from the ipsilateral LNTB (Cant and Hyson 1992; Kuwabara and Zook 1992; Roberts et al. 2014) and one from the ipsilateral MNTB (Clark 1969a; b; Kuwabara and Zook 1992).

To summarize - a typical MSO neuron receives two excitatory inputs – one from the ipsilateral and one from the contralateral AVCN – onto its dendrites and two somatic inhibitory inputs, one from the ipsilateral LNTB and one from the ipsilateral MNTB (Clark 1969a; b). This arrangement of two dominant excitatory inputs converging from both ears is also referred to as an EE type. The primary projections of MSO are the ipsilateral dorsal nucleus of the lateral lemniscus (**DNLL**) and the ipsilateral inferior colliculus (**IC**) (Adams 1979; Roth et al. 1978).

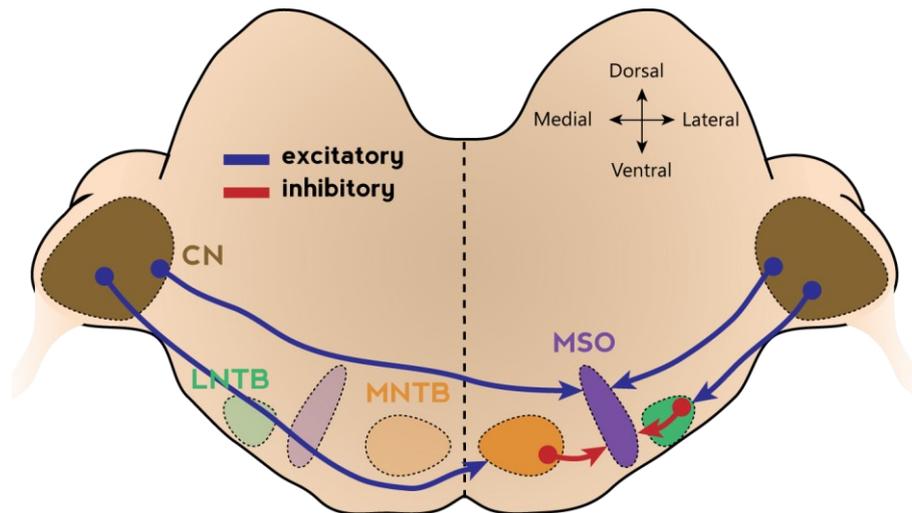


Figure 2. MSO input scheme.

A schematic view of a cross section of the gerbil brainstem showing the principal nuclei involved in low-frequency sound localization in azimuth in mammals. The MSO receives two excitatory inputs (blue arrows) – one from ipsilateral and one from the contralateral cochlear nucleus. Additionally, the MSO receives an inhibitory input from the ipsilateral lateral nucleus of the trapezoid body (LNTB) and an inhibitory input from the ipsilateral medial nucleus of the trapezoid body (MNTB). Dashed line indicates midline.

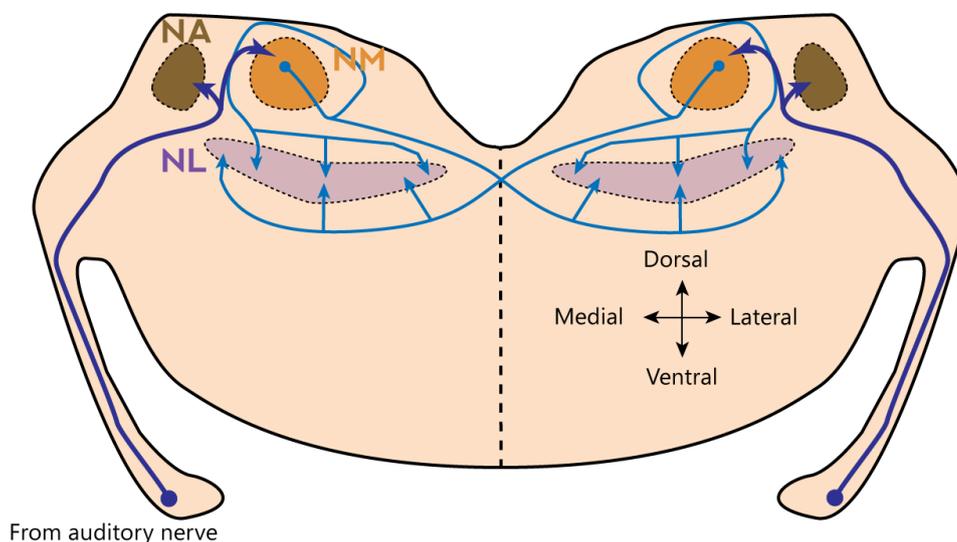


Figure 3. NL input scheme.

A schematic view of a transverse section of the barn owl brainstem showing the principal nuclei involved in low-frequency sound localization in azimuth in avians. Nucleus angularis (NA) and nucleus magnocellularis (NM) both receive direct excitatory inputs from the auditory nerve. NM sends bilateral excitatory projections to nucleus laminaris (analog of MSO in mammals). Dashed line indicates midline.

Circuitry in avians

Avians are well known for their precise sound localization capabilities; an excellent example is the barn owl (*Tyto alba*). **Figure 3** shows a schematic view of a transverse brainstem section from the barn owl, depicting three nuclei that play a key role in sound localization. In birds, the auditory nerve fibers form endbulb terminals on the neurons in the nucleus magnocellularis (**NM**) and bouton-like terminals on the neurons of the nucleus angularis (**NA**). These projections to NA and NM have a tonotopical organization (Rubel and Parks 1975). In the barn owl, nucleus magnocellularis preserves information about the timing of sound source and nucleus angularis (in a parallel pathway to NM) encodes sound level-related information. NA inherits its tonotopic organization from the auditory nerve; high best frequencies are mapped dorsolaterally, low best frequencies ventromedially. Nucleus laminaris receives bilateral inputs from NM and is responsible for detecting interaural time differences (Carr and Konishi 1990; Overholt et al. 1992). Inputs from NM to NL are tonotopically arranged; ipsilateral NM projects onto dorsal dendrites of NL, contralateral NM projects onto ventral dendrites of NL. All the aforementioned connections in avians employ glutamate as the neurotransmitter.

Phase locking

The basilar membrane of the cochlea exhibits tonotopy: waves of different frequencies preferentially produce vibrations at different locations. These vibrations result in the movements of the stereocilia of the inner hair cells (**IHC**). If the hair bundle moves towards the longest stereocilium the membrane potential of the inner hair cell will depolarize, whereas the bundle motion in the opposite direction produces a hyperpolarization. For a low-frequency wave, this translates to an increased probability of generating an action potential in the auditory nerve when the sinusoidal wave is moving in one direction and decreased probability in the opposite direction. This is the origin of **phase-locking** in auditory nerve fibers, which is the property that these fibers preferentially fire during a specific phase of a tone (Rose et al. 1967; Tasaki 1954). An AN fiber typically innervates only a single IHC, meaning that the phase-locking observed in auditory nerve is related to the membrane potential changes in single inner hair cells. The ability to show phase-locking is inherited by many nuclei within the auditory system. It is limited to lower frequencies owing to the limited

temporal acuity of hair cells and auditory neurons. In mammals, the ability to phase-lock starts to decline at 1 kHz and is no longer present at 3-4 kHz, i.e. at higher frequencies there is no correlation between action potentials in AN and stimulus phase.

Phase-locking is quantified by calculating the “vector strength” (**VS**), a metric that was first employed by Goldberg and Brown (1969). Its magnitude ranges from 0 (not synchronized) to 1 (perfectly synchronized). Note that phase-locking does not imply that an AP is fired at each cycle of the stimulus. In fact, the degree of phase locking is independent of the fraction of “skipped” stimulus cycles.

AVCN projections

Auditory nerve fibers innervate neurons in the cochlear nucleus. The anterior aspect of the anterior ventral cochlear nucleus contains spherical bushy cells, which receive input from the AN in the form of 1-2 large endbulbs of Held (Brawer and Morest 1975; Cant and Morest 1984). Globular bushy cells in the posterior aspect of the AVCN receive up to 20 smaller somatic terminals from AN – called modified endbulbs of Held (Liberman 1991). Phase-locking improves in the bushy cells compared to AN (Joris et al. 1994a; Smith et al. 1991; Smith et al. 1993). This enhancement in synchronization is seen for both low and high characteristic frequency cells when stimulated with low-frequency tones (Joris et al. 1994a; Joris et al. 1994b). However, as the stimulation frequency goes above 2 kHz, phase locking drops rapidly and becomes better in AN than in bushy cells; above 4 kHz it disappears.

The primary target of globular bushy cells is the contralateral MNTB, where the GBC axons form a synaptic connection in the form of a calyx of Held. In most cases this calyceal connection is between one GBC and one MNTB neuron (Guinan and Li 1990; Kuwabara et al. 1991; Spirou et al. 1990). GBC axons cross the midline through the ventral aspect of the trapezoid body.

Spherical bushy cells primarily target the ipsilateral lateral superior olive (LSO) and bilaterally project to the MSO. Smith et al. (1993) identified several smaller SBC projections to the contralateral ventral nucleus of the trapezoid body (**VNTB**), contralateral lateral nucleus of

the trapezoid body, contralateral ventral nucleus of lateral lemniscus (**VNLL**) and ipsilateral LNTB.

MSO specialization for sound localization

MSO and NL nuclei are the first binaural nuclei in mammals and birds, respectively. Their physiological and anatomical properties make them exceptionally suitable for the comparison of timed inputs originating from ipsilateral and contralateral ears.

Anatomical arrangement

In mammals, the MSO is a part of the superior olivary complex, which is located in the ventral brainstem. Guinan et al. (1972) showed that the MSO has a tonotopic organization; low-frequency sensitive neurons are located dorsally and high-frequencies are represented in the ventral portion of the nucleus. An anatomical study by Kiss and Majorossy (1983) found three types of neurons in the MSO: apart from the most frequently observed principal cells the MSO also contains multipolar and marginal cells. Principal cells have a striking bipolar shape, with two major dendrites extending medially and laterally, which receive inputs from the ipsi- and contralateral AVCN, respectively (**Figure 4**). It has been shown in gerbil that MSO receives additional, inhibitory inputs from LNTB and MNTB (Cant and Hyson 1992).

Principal MSO cells are densely packed in a narrow (<100 μm thickness) sheet (Rautenberg et al. 2009). The combination of good phase locking and the tight MSO arrangement and alignment of dendrites to both sides gives rise to very large field potentials (Galambos et al. 1959; Goldberg and Brown 1968; Mc Laughlin et al. 2010). Since inputs to MSO are segregated from both ears, monaural stimulation generates an open field response which changes as the electrode traverses MSO somatic layer. For a single MSO neuron monaural stimulation generates synaptic current into one dendrite forming a current sink while soma and the other dendrite compensate this sink with a current efflux creating a current source. A local population of MSO neurons responsive to the same stimulus will respond in unison creating local field potentials of opposing polarities across the somata layer forming a dipole-like field. A recent study showed that the experimentally observed field potential features can be

modelled by coupling the intracellular and extracellular domains of MSO effectively simplifying three-dimensional volume conductor model into a one-dimensional problem (Goldwyn et al. 2014). Further modelling and data showed that the typical dipole-like responses around MSO somata are restricted to frequencies above 1 kHz; frequencies below 1 kHz evoked monopole-like neurophonic responses (Goldwyn et al. 2017). These monopole-like fields seem to be caused by low-frequency driven somatic inhibition interacting with dendritic excitation. For additional details on field potential in MSO see studies by Biedenbach and Freeman (1964) and Clark and Dunlop (1968).

Changes in monaurally-evoked field potentials can conveniently be used for MSO somatic layer identification due to dipole-like field features. Advancing electrode dorsally and presenting monaural pure tone stimulus (~1 kHz) one can observe the reversal of local field potential as electrode crosses MSO somatic layer. The depth at which this reversal occurs indicates that the electrode is in the somatic layer of MSO.

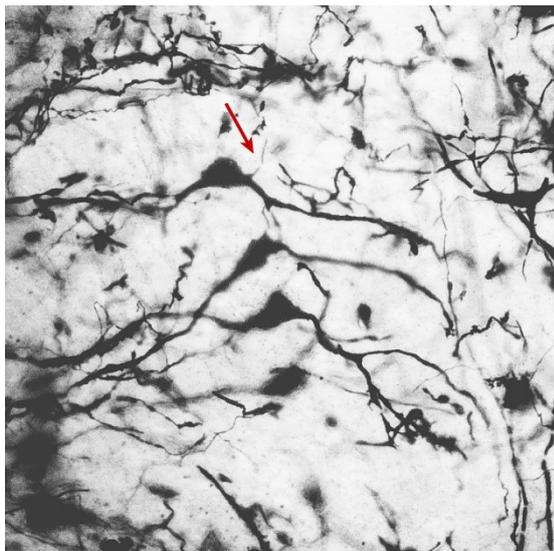


Figure 4. MSO principal cells.

Four principal MSO neurons, stained with Golgi method, show a striking bipolar organization. Each principal MSO neuron is spindle shaped and has two major dendrites at the opposite ends of the soma – one dendrite receives inputs from the ipsilateral AVCN and the other from the contralateral AVCN. A putative axon emerging from soma and dendrite junction can be seen in the upper, most dorsal neuron (red arrow).

Physiological properties

Principal MSO neurons exhibit unique intrinsic features that allow them to act as coincidence detectors with submillisecond precision. An *in vitro* study of gerbil MSO by Scott et al. (2005) showed that MSO principal neurons undergo large electrophysiological changes in the period from hearing onset (P14) to adult phase (P30-P36). MSO neurons in mature gerbils have a very low input resistance (~7 M Ω), short membrane time constant (~0.3 ms) and narrow EPSP width (<0.5 ms). The low input resistance and fast membrane time constant appear to be

mediated by low-voltage-activated potassium channels (K_{LVA}) with Kv1.1 subunits (Mathews et al. 2010; Scott et al. 2005) in combination with the hyperpolarization-activated cation current I_h , which is carried by HCN channels (Baumann et al. 2013; Khurana et al. 2012).

The importance of K_{LVA} for coincidence detection in MSO was first proposed by Svirskis et al. (2002) and detailed by subsequent studies. I_h and K_{LVA} currents show a strong developmental increase following hearing onset. During the first week of hearing both currents increase up to 13-fold and fourfold, respectively, in magnitude, and obtain progressively faster activation kinetics (Khurana et al. 2012; Scott et al. 2005). In adult gerbils a considerable fraction of both channels will be open at rest. Khurana et al. (2011) showed that I_h and K_{LVA} together help to maintain uniform EPSP amplitudes during long sound stimulation. Since the activation of the HCN channels in the MSO is at relatively positive membrane potentials, its resting conductance is very large at resting potential (around -58 mV), reducing membrane time constant and coincidence detection window by $\sim 300\%$. I_h effectively results in a more positive resting membrane potential, which increases the recruitment of K_{LVA} channels. Together these channels sharpen the coincidence detection window. Synaptic depolarization will open additional K_{LVA} channels, causing rapid repolarization, truncating the duration of the response and imparting high temporal precision for converging inputs to reach the response threshold and elicit an action potential. Mathews et al. (2010) showed that K_{LVA} channel expression is biased towards soma and proximal dendritic regions, thus compensating for dendritic filtering that would broaden excitatory post-synaptic potentials (**EPSPs**). This way MSO neurons can preserve submillisecond time resolution of EPSPs, which is essential for high temporal fidelity in the summation of EPSPs from either ear. This is the primary mechanism of MSO specialization for high temporal precision detection.

K_{LVA} are not the only potassium channels present in the MSO. Recent work by Nabel et al. (2019) shows the presence of functional high-voltage-activated potassium channels (K_{HVA}) as well. Furthermore, both types of channels have a distinct distribution pattern. K_{LVA} are more biased towards the soma and co-localize with glycinergic inputs, whereas K_{HVA} are also found in distal dendrites and co-localize with HCN1 channels.

Voltage-gated sodium channels are restricted mainly to perisomatic and axonal compartments. Interestingly, $\sim 92\%$ of VGSCs in MSO soma are inactivated at the resting potential, but the remaining non-inactivated VGSCs can amplify subthreshold EPSPs near AP

threshold, counterbalancing K_{LVA} current. This was shown by Scott et al. (2010) by comparing subthreshold EPSP amplitude dependence on injected current size (EPSC) in control and with application of the sodium channel blocker tetrodotoxin (TTX). In the presence of TTX the current needed to reach the AP threshold was almost twice as large as that for control. Thus, due to their fast activation and inactivation kinetics, VGSCs are thought to perform fast amplification of depolarizing synaptic inputs. The same study provides evidence that voltage-gated sodium channels can counterbalance inhibitory synaptic potentials (see Zhou et al. (2005)).

APs in the MSO are initiated in or near the axon initial segment and get highly attenuated in the soma and dendrites during backpropagation. In mature gerbil MSO neurons, somatic action potentials are only 5-10 mV. MSO neurons have a submillisecond absolute refractory period for AP initiation and propagation in the axon (Scott et al. 2007). Unlike in avians (Kuba et al. 2006; Kuba et al. 2005; Parameshwaran et al. 2001), action potential characteristics don't seem to depend on tonotopic positioning of MSO neurons. However, a recent study by Baumann et al. (2013) shows that in mature gerbil, I_h properties differ significantly between ventral (I_h largest) and dorsal (I_h lowest) part of the MSO.

Responses to auditory stimuli

Monaural responses

MSO neurons play a key role in sound localization by integrating the inputs from both ears. MSO principal neurons receive excitatory inputs exclusively from spherical bushy cells. As a result, their auditory response properties are comparable to that of AVCN. When stimulated monaurally, MSO neurons exhibit selectivity to a specific frequency range. The frequency to which a neuron responds the most is called best frequency (**BF**). Monaural BFs can change with changing stimulus intensity (sound pressure level or **SPL**). Best frequency at the lowest SPL to which neuron still responds is called characteristic frequency (**CF**). It has been shown that a single MSO neuron can have different CFs for ipsilateral and contralateral ears (Day and Semple (2011); this thesis). Similarly to spherical bushy cells (Kuenzel et al. 2011), MSO

frequency selectivity broadens as SPL increases with BFs becoming lower. As SPL increases, the MSO firing rate saturates. Goldberg and Brown (1969) were amongst the first to show that MSO neurons show phase locking, i.e. they preferentially fire during a particular phase of the monaurally presented pure tone stimulus (at BF) (also see Moushegian et al. (1964a)). They also showed that the preferred phases for the same BF or CF can be different for contralateral and ipsilateral inputs.

Binaural responses

Early studies have shown that these neurons typically respond poorly to monaural stimulation, but their firing rate surpasses the sum of the monaural responses during binaural stimulation at a specific time delay between monaural inputs (Goldberg and Brown 1969; Yin and Chan 1990). The interpretation for this finding has been that MSO neurons act as coincidence detectors: at the favorable time delay, the EPSPs evoked by the inputs from both ears are thought to sum at the soma and trigger an action potential, thus making MSO very sensitive to interaural time disparities (Batra et al. 1997a; Crow et al. 1978; Moushegian et al. 1964a; b; Moushegian et al. 1975; Spitzer and Semple 1995; Yin and Chan 1990).

The response of MSO neurons to a varied delay between monaural stimulation of both ears yields a rate-ITD function (**rITDf**) – during favorable delays MSO firing rate is the highest; it falls off as delays start deviating from the ‘best’ ITD and during the ‘worst’ ITD MSO response may be even lower than its response to monaural stimulation (**Figure 5C**). The time delay between the stimuli, presented to both ears, to which MSO responds the best is referred to as the best ITD (**BITD**). Similarly, the worst ITD is the time delay to which MSO responds the least. BITD and worst ITD are represented as a peak and a trough in rITDf, respectively. Many experimental and modelling observations led to the conclusion that MSO cells act as coincidence detectors of their monaural inputs.

The difference in preferred monaural phases at BF was shown to correspond to the preferred ITD of that MSO neuron (**Figure 5**) (Goldberg and Brown 1969). This idea was fleshed out in a later study by Yin and Chan (1990), where best ITD could be well predicted from monaural phase mismatches and further supported by our findings (this thesis).

Internal delays and coincidence detection

Seventy-three years ago, Jeffress (1948) presented a model on how binaural EE type neurons could operate as ITD detectors. This visionary theory set the theoretical foundation for most subsequent coincidence detection studies in the avian and mammalian auditory system. Jeffress' model of binaural cell operation has three major properties. First, both inputs to the binaural cell should carry accurate timing information about the stimulus. Second, binaural cells act as coincidence detectors – well-timed converging inputs result in a maximal response with high sensitivity to input arrival time disparities. Finally, the afferents projecting onto binaural cells form opposing delay lines, an arrangement which results in a spatial map of best ITDs. The different best ITDs between MSO neurons originate from differences in internal delay. The internal delay compensates for the difference in the arrival time at both ears at the best ITD. In the Jeffress' model, the internal delay was an axonal delay. The source of this internal delay is still a matter of debate. Several prominent theories for the origin of internal delays, next to Jeffress' axonal delay model, will be presented in later sections.

MSO neurons often have positive BITDs, i.e. ITDs that are biased towards contralateral ear leading stimulation (Batra et al. 1997a; Spitzer and Semple 1995; Yin and Chan 1990). Moreover, BITDs are often larger than physiologically-relevant ('ecological') time delays (which depend on head size of the species), especially for neurons with low CF (Brand et al. 2002; Galambos et al. 1959; Moushegian et al. 1964a; Moushegian et al. 1975; Pecka et al. 2008).

Binaural responses to different auditory stimuli

Rose et al. (1966) were the first to test the relation between BITD and sound frequency. They showed that in the inferior colliculus (IC) rate ITD functions for different frequencies either had a common peak or a common trough at a fixed ITD. Yin and Kuwada (1983) showed that those cells that had a common peak at one ITD for all relevant frequencies should have a strictly linear relationship (i.e. proportionality) between interaural phase at those frequencies. This strict linearity of the phase-frequency plot is described by its slope, which equals the common best delay (BITD) of the tones of different frequencies. Yin and Kuwada, however, also reported IC neurons whose phase-frequency curves deviated from this simple

proportionality, approximating straight lines having a nonzero intercept. This represents a combination of a constant (frequency-independent) time delay and a constant phase shift. The time delay corresponds to the slope of the phase-frequency curve and is traditionally named the characteristic delay (**CD**). The intercept is called the characteristic phase (**CP**).

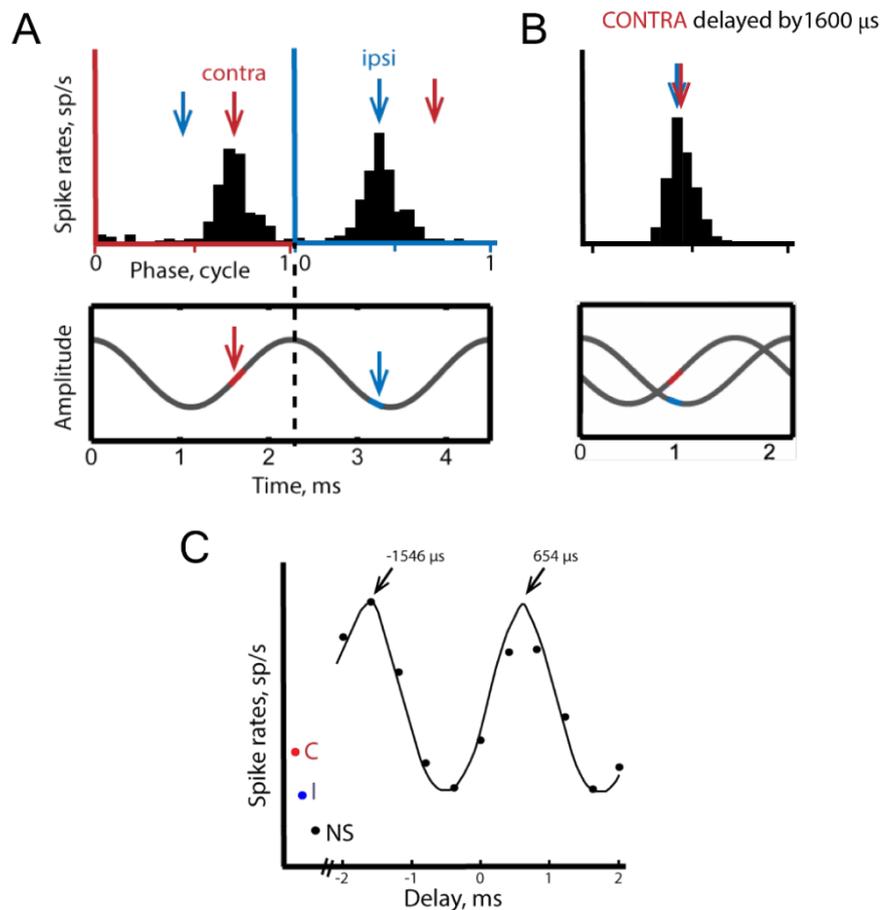


Figure 5. Phase locking in MSO.

(A) Upper: period histograms of responses to monaural ipsi- or contralateral pure tone stimulation at 444 Hz of an MSO neuron. Red arrow indicates the mean preferred phase for contralateral ear stimulation, blue arrow indicates the mean preferred phase for ipsilateral ear stimulation. Lower: corresponding mean preferred phases are shown on the stimulus waveform. (B) Upper: period histogram of MSO response to binaural stimulation when contralateral stimulus was delayed by 1600 µs with respect to ipsilateral stimulus. Lower: at this ITD, the stimulus arrives at both ears at preferred phases (same as on the left but contra waveform is shifted). (C) Rate ITD function of the same MSO neuron as in A and B. The neuron fires maximally when stimulus to the contralateral ear is delayed by 1546 µs. During the worst ITD, MSO response can be below monaural responses (C – contra, I – ipsi). NS – spontaneous activity. Data adapted from Goldberg and Brown (1969).

The name “characteristic delay” for the slope was motivated by Yin and Kuwada’s incorrect statement that the CD is representative of the range of best ITDs of the frequencies that best excite the neuron. In reality, the numerical simulation in their own Fig. 1C shows that is not the case: it has a CD of 200 us, whereas the best ITD ranges from 300-400 us. The correct relation between CD, CP and best ITD is

$$\text{BITD} = \text{CD} + \text{CP} * T,$$

where T is the period of the tone. For fig. 1C of Yin and Kuwada (1983) the 300-400 us range of BITD immediately follows from the $\text{CP} = 0.2$ cycle and the frequency range of 1-2 kHz.

Characteristic phase represents a constant phase difference between phase-locked inputs from both ears. If two monaural inputs with the same phase locking but different internal delays are presented, MSO would respond the most when an appropriate interaural delay is introduced to compensate for internal delays. If these delays are independent of frequency, then ITD will be the same as CD (the slope of phase vs frequency). In this case, the inputs to the MSO at the interaural delay corresponding to the CD are exactly in-phase, resulting in $\text{CP} = 0$. If one input is inhibitory then the two inputs are out of phase, yielding $\text{CP} = 0.5$ (Joris and Yin 1995). In this case, it is the troughs of the rITDf (“worst ITDs”) rather than their peaks that are aligned across frequency. Any CP value between 0 and 0.5 indicates that a constant phase difference is somewhere between peak and trough of rITDf. Several studies have shown that the CP for MSO clusters around 0 (Batra et al. 1997a; Spitzer and Semple 1995; Yin and Chan 1990), while other studies have reported a bias toward small positive values (Bremen and Joris 2013; Pecka et al. 2008). Independent of the population distribution, all of these studies report individual MSO neurons having clear nonzero CP values, that is, neurons whose binaural sensitivity cannot be described by a constant, frequency-independent interaural delay.

Natural sounds commonly consist of many frequency components, which raises the question – how does MSO localize sounds other than pure tones? Yin and Chan (1990) compared MSO responses to wideband stimuli with a composite of rITDfs recorded at individual frequency components and found that the two rITDfs are very alike. This finding suggests that MSO cells sum the spectral components of wideband stimulus approximately linearly. The same study also showed that when both ears are presented with uncorrelated wideband stimuli, the MSO

shows no ITD sensitivity. This suggests that MSO cells perform computation comparable to cross-correlation of the inputs. The mechanisms behind MSO as a coincidence detector is another topic of this thesis and these questions will be looked at in more detail in further chapters.

Role of LSO in sound localization

High-frequency (>3 kHz) sound localization in azimuth is governed by another major nucleus of the superior olivary complex – the lateral superior olive (**LSO**). LSO neurons are Inhibited by stimulation of the contralateral ear and Excited by stimulation of ipsilateral ear; this arrangement is denoted as IE type. The LSO exhibits a distinct S-shaped structure with the tonotopic axis running along the curved axis of the S; high frequencies are represented medially and low frequencies laterally (Guinan et al. 1972; Tsuchitani 1977). Two dominating inputs to the LSO come from the AVCN and MNTB respectively (Cant and Casseday 1986; Glendenning et al. 1985; Glendenning et al. 1991; Spangler et al. 1985). The ipsilateral, excitatory AVCN inputs originate primarily from SBC axons (Glendenning et al. 1985). The inhibitory MNTB inputs are relayed from the contralateral GBCs. Efferents of the LSO project bilaterally to the IC and the DNLL (Glendenning and Masterton 1983; Roth et al. 1978).

The LSO is considered to be responsible for the initial stage of ILD encoding. Monaural stimulation of the ipsilateral ear typically shows a monotonic increase and eventual saturation in LSO firing rate with increasing sound intensity. An ILD function of an LSO neuron is typically obtained by measuring the spike rate dependence when the ipsilateral stimulus is presented at a fixed chosen SPL and the contralateral stimulus SPL is varied. A typical ILD function of an LSO neuron exhibits a sigmoidal shape – increasing sound level in the contralateral ear progressively inhibits the ipsilaterally driven responses until eventually the LSO neuron stops firing altogether.

Although ILD appears to be the most important cue for LSO neurons, temporal information about amplitude modulation (**AM**) is available to LSO from the three afferents that can phase lock to AM stimuli: SBCs, GBCs and MNTB (Joris and Yin 1998). Phase-locking to AM stimuli makes the LSO neurons sensitive to ITDs. In the case of no interaural phase difference the inhibitory signal will reach LSO at approximately the same time as the excitation in many cells,

yielding a minimal LSO response. On the other hand, when both signals reach LSO out of phase (IPD = 0.5), the excitatory response receives no suppression, which results in the strongest LSO response close to that of monaural ipsilateral stimulus. This ITD-based sound localization in LSO is limited by the inability to follow modulation frequencies higher than 300 Hz (Joris 1996), which may be caused by the relatively long duration of the IPSPs evoked by the MNTB afferents. Low-frequency LSO neurons also show ITD sensitivity.

Juxtacellular recordings

In vivo recordings from the MSO have proven to be notoriously difficult due to their location in the ventral aspect of the brainstem, the thin somatic layer of the MSO, and the small size of somatic action potentials. For this reason, most MSO data are based on extracellular recordings, however these usually suffer from high local field potential ('neurophonic') contamination. Whole-cell MSO recordings are technically difficult to establish owing to the high density of principal neurons and spindle-shaped somata. In my thesis, I have, therefore, mostly employed the juxtacellular (also known as loose-patch) configuration.

Juxtacellular recordings are established by advancing a patch clamp pipette until it makes contact. The electrode is further pushed down for a few micrometers into the cell. During the approach typically minimal positive pressure is maintained. When a contact between the electrode and the cell is made, electrode resistance increases, and the positive pressure is released. Recordings are typically performed in current clamp mode.

By performing simultaneous whole-cell recordings from MNTB cells, Lorteije et al. (2009) showed that the waveforms recorded in loose-patch configuration closely resembled sum of scaled versions of the membrane potential and its first time derivative. This suggests that juxtacellular potentials correspond to the local membrane currents, which are the sum of resistive and capacitive currents. In Chapter 2 of this thesis the relationship between whole-cell and juxtacellular recordings in MSO will be investigated.

Models of MSO operation

The previously introduced Jeffress' ITD detection model relies on two assumptions for EE type binaural neurons: they should operate as coincidence detectors and the internal delay source is axonal delay lines. Chapter 3 of this thesis will show that a linear cross-correlation of the inputs forms an adequate description of the action of the MSO neurons. The origin of the internal delay in the mammalian auditory system is still a matter of great debate. It has been shown that some avian species have clear axonal delay lines, as predicted by Jeffress. However, anatomical studies in mammals have not found evidence for the necessary afferent axonal arrangement to support this proposal (Karino et al. 2011). Together with the unknown role of somatic inhibition, these findings gave rise to several other theories trying to explain the source of internal delays in mammalian MSO. Two prominent alternatives are **a)** well-timed inhibition onto MSO modifies ITD tuning and **b)** cochlear delay (stereausis) model, where times to reach CF-sensitive zones in the cochlea differ for both ears. Several other models have been proposed as well and they will be briefly discussed here too.

Jeffress' axonal delay model

The model proposed by Jeffress (1948) explains internal delay to originate from the difference in the length of excitatory paths from the ipsilateral and contralateral sides. These excitatory paths are composed of axons from bushy cells that terminate on the dendrites of MSO cells. For that reason, this model is also referred to as 'axonal delay' model. This model is represented by an axonal line layout with opposing input length gradients from contralateral and ipsilateral sides constituting a place code (**Figure 6A**).

Studies in avians showed that such an arrangement of axonal delay lines indeed exists and that it is the most likely source of the internal delay. In avians, the nucleus magnocellularis projects bilaterally to the ipsilateral dorsal aspect of the nucleus laminaris (MSO homologue) and to the ventral aspect of the contralateral NL. It has been shown that in the barn owl the ITD representation comes from a place code layout (Wagner et al. 2007). Seidl et al. (2010) showed that in the chick the difference in axonal lengths between contralateral and ipsilateral inputs alone cannot be responsible for the internal delay, since the contralateral axon from NM to NL is on average > 1600 μm longer than the ipsilateral axon. However, this axonal

length offset is compensated by systematic differences in axon diameter and internode distance, resulting in a gradual change in conduction velocity. The same was found in the barn owl (Carr et al. 2015).

Unlike in avians, studies in mammalian MSO circuitry met with difficulties in trying to confirm the presence of a place code. Anatomical evidence for delay lines has been investigated by Smith et al. (1993), Beckius et al. (1999) and Karino et al. (2011).

While Smith et al. (1993) did not find evidence for systematic opposing gradient delay lines between the two sides, which were postulated by Jeffress (1948), they did find evidence for a rostral to caudal neural delay of the contralateral but not the ipsilateral innervation of the MSO by the SBCs. In contrast to the first study, Beckius et al. (1999) showed that the contralateral and ipsilateral MSO innervations can exhibit delay lines of opposing gradients. On the ipsilateral side, rostrocaudal gradient slopes for both experimental subjects were similar, but those on contralateral side were very different. Additionally, there was a discrepancy between lengths of axons terminating at a similar rostrocaudal location. The two latter observations cast doubts on whether axonal delay lines can be the dominant mechanism responsible for the ITD sensitivity.

Karino et al. (2011) performed a more detailed re-examination of the data by (Smith et al. 1993), but found that ipsilateral projections can also form a caudally directed delay line pattern. Moreover, the distribution of estimated axonal delays did not match the distribution of best delays obtained from physiological measurements.

In sum, there is evidence of Jeffress' proposed axonal configurations contralaterally, but this is overall less clear for ipsilateral projections in the cat. Three major points arguing against pure axonal delay lines were raised: a) there was no evidence for a relationship between ITD and CF; b) collaterals did not span across the full rostrocaudal extent of the MSO, complicating the formation of a systematic gradient along this axis; c) combined contralateral and ipsilateral delay values, while relatively small, could account for the observed best delays, but to span the full range of them an additional delay mechanism is needed. Neither of the three studies in the cat found significant differences in axonal diameters of afferents from both sides. These findings point to the absence of an anatomical axonal delay line arrangement

from both ears to MSO, but do not reject the possibility of systematic conduction velocity changes across rostro-caudal dimension as seen in avians.

A very important finding in avian sound localization circuitry was that differences in conduction velocities due to differences in internode lengths and axon diameters can compensate for differences in axon length (Fischer and Seidl 2014; Seidl et al. 2014; Seidl et al. 2010). This has recently been investigated in Mongolian gerbil as well (Seidl and Rubel 2016). In gerbils with mature hearing capability (P20), internode length was 1.85 times longer in the contralateral axon than in the ipsilateral. Contralateral axons also featured larger diameter than their counterparts. This implies a differential velocity regulation in collaterals from contralateral and ipsilateral AVCN. Fiber trajectory measurements gave estimated average difference between contralateral and ipsilateral AVCN-MSO pathway length of $2138 \pm 102 \mu\text{m}$. A uniform conduction velocity would result in $267 - 1069 \mu\text{s}$ difference between binaural inputs to MSO. It is very likely that increased diameter and internode length on the contralateral fibers compensate for longer travel distance (Brill et al. 1977). In avians a 1.95-fold difference in internode length resulted in a 2.39-fold increase in conduction velocity (Seidl et al. 2014). A similar mechanism might minimize differences in conduction time between ipsi- and contralateral inputs to the mammalian MSO.

Well-timed inhibition model

The presence of two glycinergic inhibitory inputs, one from the ipsilateral ear (through ipsilateral LNTB) and one from the contralateral ear (through ipsilateral MNTB), raises the question what the physiological function of inhibition is in the operation of the MSO. Recordings from the MSO in brain slices showed the presence of IPSPs when stimulating afferent pathways to the MSO (Grothe 2000; Grothe and Sanes 1993; 1994; Roberts et al. 2013; Smith 1995). Furthermore, inhibitory inputs from the contralateral ear show tight phase-locking (Smith et al. 1998; Tollin and Yin 2005).

Pharmacological blocking of these glycinergic inhibitory inputs by iontophoretic application of strychnine led to a shift of the BITD towards 0 ms (Brand et al. 2002; Pecka et al. 2008). A model was proposed to explain this shift, which featured a brief contralaterally-driven IPSP preceding the EPSP and causing a delay in reaching action potential threshold (**Figure 6B**). The

possible role of inhibition in the operation of the MSO will be addressed in more detail throughout further chapters of this thesis.

Cochlear delay model

Basilar membrane maps different frequencies to different locations along its axis, with high frequencies activating the BM at its basal end and low frequencies at its apical end. The travelling wave for two different CF sounds will travel different distances, resulting in different times needed to reach their respective activation zones in BM. Asymmetric innervation from two cochleae as a source of internal delays is the foundation of the cochlear delay model, which is also referred to as the stereausis model due to its similarity to stereopsis in visual processing (Bonham and Lewis 1999; Schroeder 1977; Shamma et al. 1989) (**Figure 6C**). It has been shown that small interaural differences in CF could account for differences in internal delay (Bonham and Lewis 1999; Joris et al. 2006). An obvious requirement for this model to work in MSO is an asymmetric CF for the inputs from both ears. Some evidence has been presented in favor of this model (Day and Semple 2011). Chapter 4 of this thesis will address stereausis model in more detail.

Other internal delay models

The three models discussed above outline possible internal delay sources stemming from different MSO circuitry elements. Some models, however, try explaining internal delays through differences in inputs that are converging onto the MSO.

The first model is based on differences in the rising slopes between ipsi- and contralateral EPSPs (Jercog et al. 2010). This model is based on experimental evidence obtained in slice experiments showing that the rise times of ipsilateral EPSPs are on average 500 μ s faster than those of contralateral EPSPs. Faster rise times of ipsilateral potentials can 'outrun' low-threshold potassium channels more easily, causing interaural disparity.

The second model, proposed by Zhou et al. (2005), relies on interaural asymmetry in the delay between bilateral EPSPs and the triggering of action potentials. Under the assumption that the axon originates from the dendrite that receives ipsilateral inputs, contralateral excitation

will be more sensitive to synaptic inhibition and to the action of somatic voltage-dependent ion conductances than the ipsilateral excitation, providing a physiologically relevant internal delay. These two models suggest that part of the internal delay may be created at the level of the MSO neuron. Both models will be briefly addressed in the following chapter.

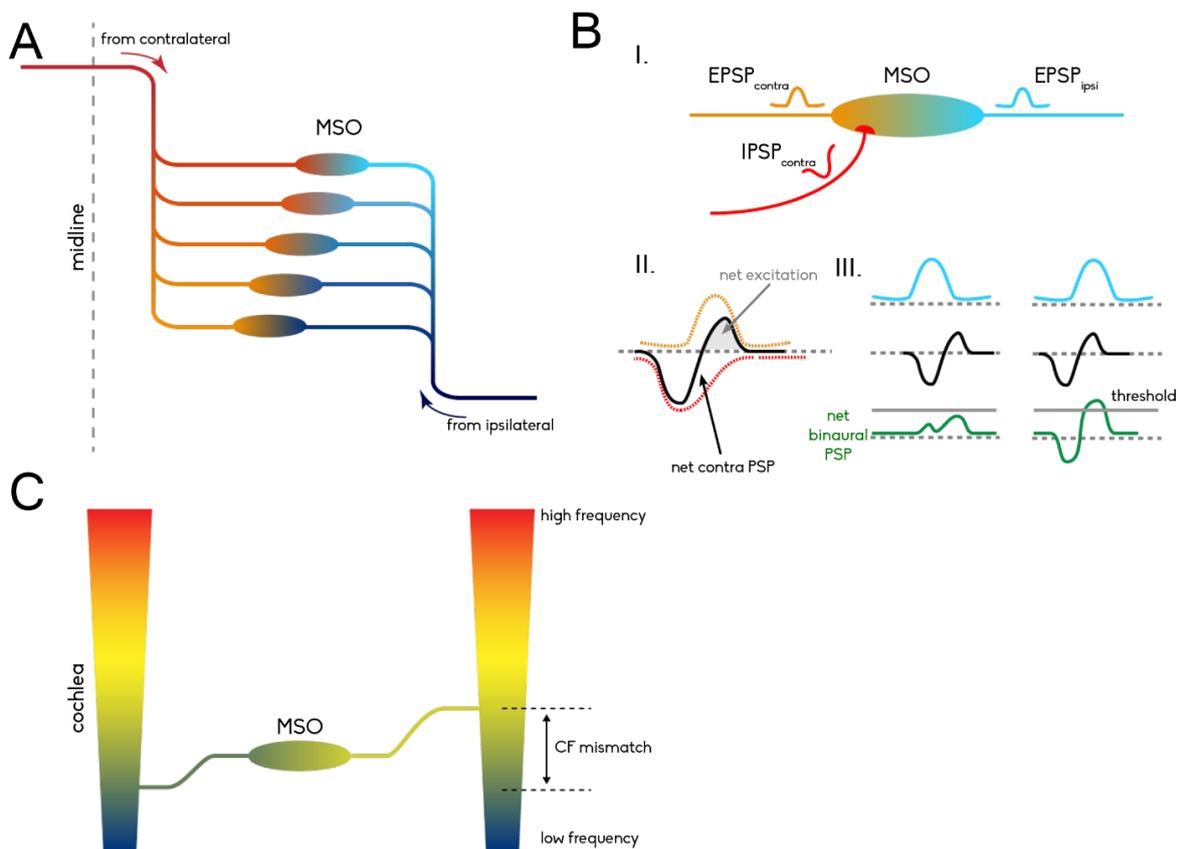


Figure 6. Three internal delay models for MSO.

(A) Jeffress' axonal delay model. Excitatory inputs from ipsilateral and contralateral ears take different times to converge onto different MSO neurons due to variation in the lengths of afferents. **(B)** Well-timed inhibition model. *(I)* Schematic depiction of two EPSPs from contralateral and ipsilateral ears (yellow and blue, respectively) and one contralateral IPSP (red) converging onto an MSO neuron. *(II)* Net contralateral PSP (black) is the sum of the contralateral EPSP (yellow) and the contralateral IPSP (red). Gray area represents the effective excitation part from the contralateral side. *(III)* Binaural interaction of ipsilateral EPSP (blue) and contralateral PSP (black) shown as a linear sum (green) at two different ITDs. Grey line indicates action potential threshold. **(C)** Cochlear mismatch (stereausis) model. MSO neurons receive inputs from two ears at places on the basilar membrane, corresponding to different CFs. As a result, the travelling wave in the cochlea where the CF is lower will take more time to activate the input towards the MSO than in the other ear, effectively creating an internal delay.

Models of input integration

While it is commonly agreed that MSO acts as a coincidence detector, the mechanisms of input integration are still debated. Agmon-Snir et al. (1998) proposed a nonlinear input integration mechanism in which the non-stimulated dendrite acts as a current sink, reducing the ability of inputs coming from the stimulated dendrite to trigger action potentials. This would explain why binaural stimulation triggers action potentials more efficiently than monaural stimulation. Alternatively, Colburn et al. (1990) proposed that the two bilateral inputs sum linearly, but that the relationship between the number of simultaneous inputs and the threshold for AP initiation is non-linear. A simple linear input integration can be achieved by linear cross-correlation, a mechanism shown to be applicable in avians (Fischer et al. 2008; Fischer et al. 2011), but thought to be less appropriate in mammals (Batra et al. 1997b; Batra and Yin 2004; Franken et al. 2015). Chapter 3 of this thesis will address the suitability of cross-correlation to describe binaural integration in the mammalian MSO.

Scope of this thesis

Due to its ventral position in the brainstem, the anatomical approach towards MSO is very hard. Another major hurdle is its strong local field potentials, which contaminate extracellular recordings in this nucleus. Furthermore, the spindle-shaped MSO soma hinders the whole-cell patch clamp approach. These obstacles are the underlying cause for the scarcity of *in vivo* MSO data in the field. The majority of studies have had to rely either on an *in vitro* approach or on modelling of existing experimental data.

A ventral surgical approach combined with the use of the neurophonic for localization and juxtacellular recordings allowed us to gather a large dataset of MSO responses to various auditory stimuli. This dataset enabled us to tackle several controversial questions about MSO operation: what mechanism is the most plausible source of internal delays and how are the ipsi- and contralateral inputs integrated?

Chapter two of this thesis focuses on subthreshold responses recorded in the gerbil MSO. We establish the validity of juxtacellular recordings by comparing them to their whole-cell counterparts; we compare ipsi- and contralateral subthreshold events and address the action

potential generation mechanism. The observed linear interaction between the inputs and nonlinear input-output relationship allowed us to put forward a relatively simple model for MSO operation and reject less plausible ones.

The third chapter tackles to what extent input summation can be described by cross-correlation. Here we compare recorded binaural responses and predictions from monaural responses for wideband auditory stimuli. Using a fitting function for recorded data and predictions based on cross-correlation, we find a striking similarity between the two, confirming the MSO acting as a linear cross-correlator.

Finally, chapter four aims at the stereausis theory as a potential internal delay source. Juxtacellular recordings were accompanied by electrical round window stimulation, allowing us to compare BITDs with and without cochlear delay being introduced. Together with the absence of correlation between CF mismatches and BITDs, these findings showed no evidence supporting cochlear-delay model. Findings in these chapters lead us to conclude that from the three prominent internal delay theories, Jeffress' axonal delay model seems to be the most plausible, similar to the avian MSO homologue operation.

Chapter 2

Directional hearing by linear summation of binaural inputs at the medial superior olive

**Marcel van der Heijden, Jeannette A. M. Lorteije, Andrius Plauška,
Michael T. Roberts, Nace L. Golding, J. Gerard G. Borst**

Neuron. 2013; 78:936-948

Abstract

Neurons in the medial superior olive (MSO) enable sound localization by their remarkable sensitivity to submillisecond interaural time differences (ITDs). Each MSO neuron has its own “best ITD” to which it responds optimally. A difference in physical path length of the excitatory inputs from both ears cannot fully account for the ITD tuning of MSO neurons. As a result, it is still debated how these inputs interact and whether the segregation of inputs to opposite dendrites, well-timed synaptic inhibition, or asymmetries in synaptic potentials or cellular morphology further optimize coincidence detection or ITD tuning. Using *in vivo* whole-cell and juxtacellular recordings, we show here that ITD tuning of MSO neurons is determined by the timing of their excitatory inputs. The inputs from both ears sum linearly, whereas spike probability depends nonlinearly on the size of synaptic inputs. This simple coincidence detection scheme thus makes accurate sound localization possible.

Introduction

Sixty-five years ago, Jeffress proposed a cellular model to explain how ITDs are used to localize sounds (Jeffress 1948). He postulated neurons that fired when inputs from both ears arrived at the same time. He further postulated delay lines introducing different travel times of inputs from either ear which would allow these coincidence detectors to be specifically tuned to certain ITDs. Experimental work showed that principal neurons of the MSO fulfil many of the predictions of his model, including tuning for certain ITDs (Goldberg and Brown 1969; Spitzer and Semple 1995; Yin and Chan 1990). Because these cells are such good coincidence detectors, they have even been compared to logical AND gates (Herz et al. 2006).

It has been very difficult to record the synaptic inputs of MSO neurons *in vivo* because of their location in the ventral brainstem, the large field responses (Biedenbach and Freeman 1964; Galambos et al. 1959; Mc Laughlin et al. 2010), unusually low input resistance, fast time course of synaptic potentials (Mathews et al. 2010), and the small size of the somatic action potentials (Scott et al. 2007; Scott et al. 2005), which altogether make it harder to distinguish between synaptic potentials and action potentials during *in vivo* extracellular recordings from the somatic region. Consequently, two aspects of Jeffress' theory are still disputed (reviewed in (Ashida and Carr 2011; Grothe et al. 2010)). The first involves the anatomical arrangement of the inputs from both ears, which are segregated to opposite dendrites (Grothe et al. 2010). It has been proposed that this arrangement favours binaural inputs over monaural inputs, since it would be difficult for monaural inputs to reach threshold owing to the current sink of the non-stimulated dendrite (Agmon-Snir et al. 1998). This would explain how MSO neurons can be such efficient coincidence detectors, being driven much more effectively by optimal binaural stimuli than by monaural sounds (Goldberg and Brown 1969; Langford 1984; Spitzer and Semple 1995; Yin and Chan 1990). In an alternative model, inputs from both ears sum linearly, but the efficient coincidence detection results from a non-linear relation between the number of simultaneous inputs and spike probability (Colburn et al. 1990). The other area of debate involves the mechanisms causing most MSO neurons to be preferentially activated by contralaterally leading sounds. Difficulties in matching the observed path lengths with the

distribution of “best delays” (Beckius et al. 1999; Karino et al. 2011; Seidl et al. 2010), have inspired alternative models to the anatomical delay lines of Jeffress’ theory. A subject for debate is whether the arrival of the excitatory inputs determines ITD tuning, as Jeffress (Jeffress 1948) originally proposed. In addition to the excitatory inputs originating from the spherical bushy cells of ipsi- and contralateral cochlear nuclei, the MSO neurons also receive prominent glycinergic inhibitory inputs on soma and proximal dendrites arising mainly from the medial nucleus of the trapezoid body (MNTB; contralateral ear), but also from the lateral nucleus of the trapezoid body (LNTB; ipsilateral ear; reviewed in (Grothe et al. 2010). Pharmacologically blocking the inhibitory inputs to the MSO neurons can shift the best ITD from contralaterally leading towards 0 μ s (Brand et al. 2002; Pecka et al. 2008). To explain this observation, a model has been proposed in which brief IPSPs activated by contralateral sounds immediately precede the EPSPs, thus delaying the triggering of the action potential (Brand et al. 2002; Pecka et al. 2008). This well-timed inhibition model predicts a significant phase-dependent interaction between the postsynaptic potentials of both ears for in vivo recordings. A second model which also proposes a central role for the MSO neurons in shaping the internal delays is based on an interaural disparity in EPSP slopes, the contralateral inputs being less effective in triggering spikes because their slower risetime leads to larger activation of low-threshold potassium channels. The interaural disparity in risetimes would then favor instances in which the more effective ipsilateral inputs arrive first (Jercog et al. 2010). This model predicts a difference in slope between postsynaptic potentials of both ears for in vivo recordings. A third model assumes an interaural asymmetry in the delay between ipsi- and contralateral EPSPs and generation of action potentials (Zhou et al. 2005). This model predicts during in vivo recordings a difference in the delay between ipsi- and contralateral EPSPs and the respective APs they trigger. A test of these different models therefore requires direct recording of the inputs of MSO neurons in vivo. To investigate how signals from both ears interact in MSO neurons, we made juxtacellular (loose-patch) and whole-cell recordings from principal neurons of the low-frequency area of the MSO in gerbils, which, like humans, use ITDs for sound localization (Heffner and Heffner 1988; Maier and Klump 2006).

Results

Juxtacellular recordings can resolve inputs to MSO neurons

We used a ventral approach to make juxtacellular (loose-patch) recordings from principal neurons of the low-frequency area of the somatic layer of the gerbil MSO (Figures 1 and S1). We studied binaural interactions using “binaural beat” stimuli (Yin and Chan 1990), for which the tone frequencies always differed by 4 Hz between the ears. The 4-Hz beat causes the interaural phase difference (IPD) to change continuously over the 250-ms beat period. In all MSO cells, binaural beats triggered complex responses (Figure 1A, B). Remarkably, rapid, positive fluctuations were also observed in the absence of sound stimulation (Figure 1D). These spontaneous fluctuations were smaller than the tone-evoked fluctuations. They depended critically on pipette position, since they disappeared upon withdrawal of the pipette. The estimated half width of these spontaneous events was $415 \pm 73 \mu\text{s}$ (mean \pm standard deviation; $n = 19$ cells), similar to EPSPs measured in slice recordings (Scott et al. 2005). We therefore interpret these randomly timed events as the postsynaptic response to the spontaneous activity of spherical bushy cells (SBCs), the main excitatory inputs to MSO. The extracellularly recorded EPSPs (eEPSPs) could not be well delineated owing to their high rate. Lower bound estimates of spontaneous input rates were obtained by peak counting. In most (14/19) cells, peak rate exceeded 500/s.

During tone stimulation, the size of the events increased (Figure 1B). Half width of tone-evoked events was $438 \pm 73 \mu\text{s}$. The largest events triggered extracellularly recorded action potentials (eAPs). These events had an amplitude of $1.0 \pm 0.5 \text{ mV}$ and a maximum rate of rise of $6.4 \pm 3.1 \text{ V/s}$. eAPs were generally small, sometimes even smaller than the eEPSPs that triggered them, in agreement with the small size of somatic APs in whole-cell slice recordings (Scott et al. 2005), which is caused by restricted invasion of the somatodendritic compartment by the backpropagating axonal AP (Scott et al. 2007). Nevertheless, eAPs could be readily identified by their steep downward slope immediately following the peak (Figure 1C, E). The latency between eEPSPs and eAPs was inversely related to eEPSP size (Figure 1F, G); on average it was $168 \pm 20 \mu\text{s}$ ($n = 19$ cells), with an average coefficient of variation of 0.24. Spontaneous rates ranged from 0 sp/s (5/19 cells) to 12.5 sp/s, (median value 0.4 sp/s),

comparable to estimates from extracellular recordings (Goldberg and Brown 1969; Yin and Chan 1990).

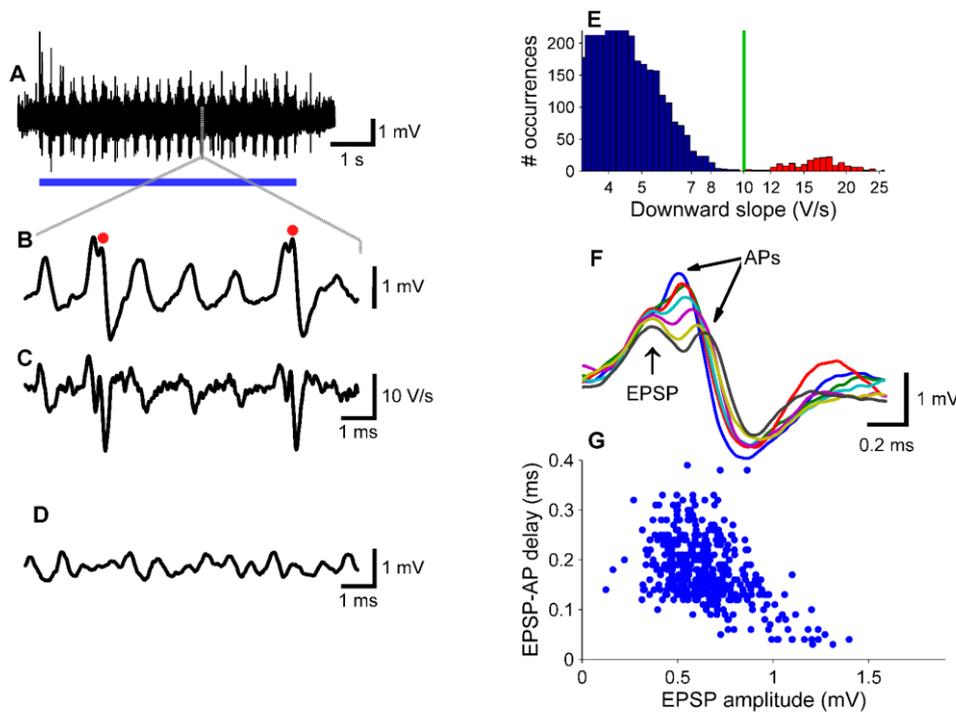


Figure 1. Juxtacellular recordings in MSO.

(A) Juxtacellular recording from a neuron in the somatic layer of the MSO, which was identified based on field potentials (Figure S1), showing the response to a 4-Hz binaural beat (700/704 Hz tone; 50 dB SPL). Stimulus presentation is marked by the blue bar. (B) Short segment of the recording of (A). Two action potentials are marked with red dots. (C) Time derivative of segment shown in (B) illustrating that action potentials can be identified based on their steep downward slopes. (D) Segment of spontaneous activity of the same cell. (E) Bimodal distribution of downward slopes, enabling the distinction of subthreshold events (blue) and action potentials (red). Green line indicates threshold criterion. (F) Action potentials time-aligned on the preceding EPSPs. Smaller EPSPs result in larger EPSP-AP latencies. (G) Scatter plot of EPSP-AP latency versus EPSP magnitude. Characteristic frequency (CF): 680 Hz.

Relation between juxtacellular and whole-cell recordings

The highly unusual properties of the principal neurons were also observed in whole-cell recordings in vivo. A total of 3 neurons were recorded for a sufficiently long period to allow binaural beat stimulation (Figure 2A-C). Membrane potential was -60 ± 3 mV ($n = 3$). Spontaneous fluctuations were observed with half widths that were somewhat larger than

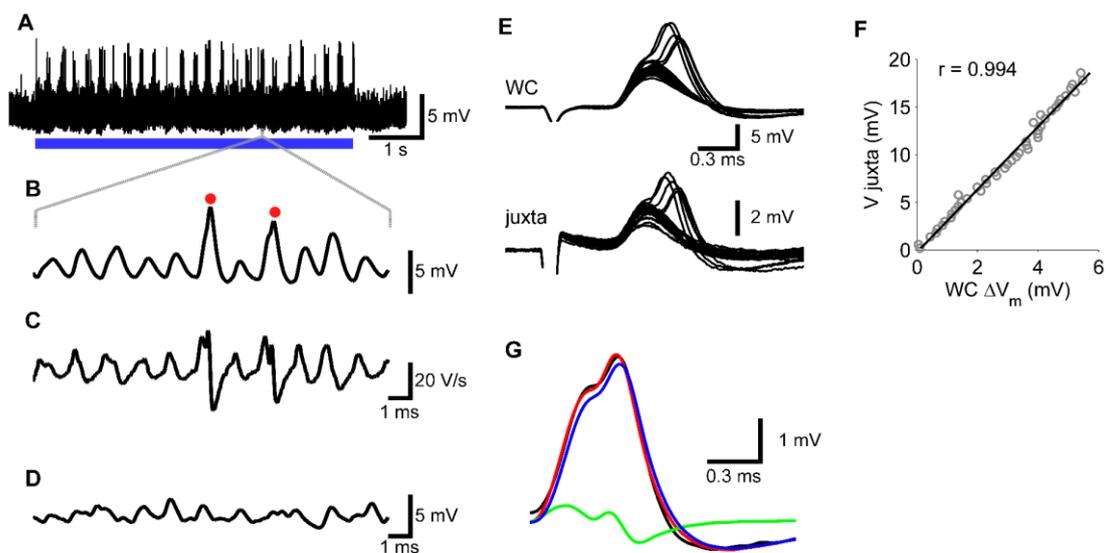


Figure 2. Whole-cell recordings in MSO.

(A-D) shows an in-vivo whole cell recording; (E-G) illustrates the relation between a juxtacellular and a whole-cell recording obtained from a paired recording in brainstem slices. (A) Response to a 700/704-Hz, 40-dB-SPL binaural beat from an MSO neuron with a CF of 790 Hz. Resting membrane potential was -60 mV. (B) Short segment of the trace shown in (A). Two action potentials are marked by red dots. (C) Time derivative of the trace shown in (B), illustrating the faster repolarization phase of action potentials. (D) Segment of spontaneous activity of the same cell. (E) Simultaneous whole-cell and juxtacellular recordings of principal neuron in MSO slice showing EPSPs evoked from ipsilateral afferent stimulation, which in some cases triggered APs. (F) Relation between juxtacellular and intracellular peak EPSP amplitudes. Solid line shows line fit ($r = 0.994$). (G) The juxtacellularly recorded EPSP (black trace) can be well approximated by the sum (red trace) of a scaled version of the membrane potential (blue trace; resistive coupling constant 298 mV/V) and a scaled version of the time derivative of the membrane potential (green trace; capacitive coupling constant 8.2 $\mu\text{V/V/s}$).

juxtacellularly recorded spontaneous fluctuations (Figure 2D). The smallest events could not be identified unambiguously, but using a minimum amplitude criterion of 0.5 mV, we estimated average rates of about 900 events / s. These events had half widths of $608 \pm 142 \mu\text{s}$. During binaural beat stimulation, the size of the EPSPs increased and they showed good phase locking (Figure 2A, B). Tone-evoked EPSPs had a half width of $601 \pm 122 \mu\text{s}$. The largest EPSPs evoked APs. APs had an average amplitude of only $8.5 \pm 1.3 \text{ mV}$ ($n = 3$), but could be reliably identified based on their faster rate of repolarization (Figure 2C). Suprathreshold

EPSPs had an estimated average amplitude of 4.6 ± 1 mV and a maximum rate of rise of 20.2 ± 3.7 V/s. The estimated delay between EPSPs and APs was 216 ± 34 μ s. Juxtacellular recordings provide a measure for the local membrane currents, which consists of a resistive component, which is proportional to the intracellular membrane potential and a capacitive component, which is proportional to the first derivative of the membrane potential (Freygang and Frank 1959; Lorteije et al. 2009). A comparison of juxtacellular and whole-cell recordings indeed suggests that the shape of EPSPs and APs in juxtacellular recordings (Figure 1B) was intermediate between membrane potentials (Figure 2B) and their first derivative (Figure 2C).

To test whether juxtacellular potentials can be used in a quantitative manner, we made simultaneous juxtacellular and whole-cell current-clamp recordings from MSO principal neurons in electrophysiologically mature gerbil slices (Scott et al. 2007). Spontaneous inputs as shown in Figure 1C and 2D were not observed, in agreement with previous slice recordings. Comparison of the shape of EPSPs evoked by afferent stimulation in juxtacellular (eEPSP) and whole-cell recordings (iEPSP) showed that the juxtacellular recordings could be approximated by a mixture of a scaled-down version of the intracellular membrane potential and its time derivative. The relative contribution of both components varied between cells. An example with a relatively large resistive component is shown in Figure 2E. In 9 cells in which EPSPs were afferently evoked, the resistive coupling constant was 127 ± 96 mV/V and the capacitive coupling constant was 5.6 ± 5.1 μ V/V/s. The relation between the amplitude of iEPSPs and eEPSPs was linear (Figure 2F); average correlation was $r = 0.945 \pm 0.036$ ($n = 9$). Linearity was also excellent for IPSPs, which were evoked by conductance clamp ($r = 0.991 \pm 0.015$; $n = 5$; Figure S2 A,B). To further evaluate the linearity of the relation between intracellular and extracellular amplitudes, we injected intracellular depolarizing and hyperpolarizing currents, which showed that peak amplitudes were linearly related in the voltage range between -50 and -70 mV ($r = 0.989 \pm 0.010$; $n = 6$), but that outside this range, the relation changed, probably because of a voltage-dependent change in the resistive component of the juxtacellular membrane currents (Figure S2C, D). Because of the limited voltage range over which the membrane potentials operated in vivo (Figure 2A, B), we conclude that in vivo juxtacellular recordings can be used to quantify subthreshold activity in the MSO.

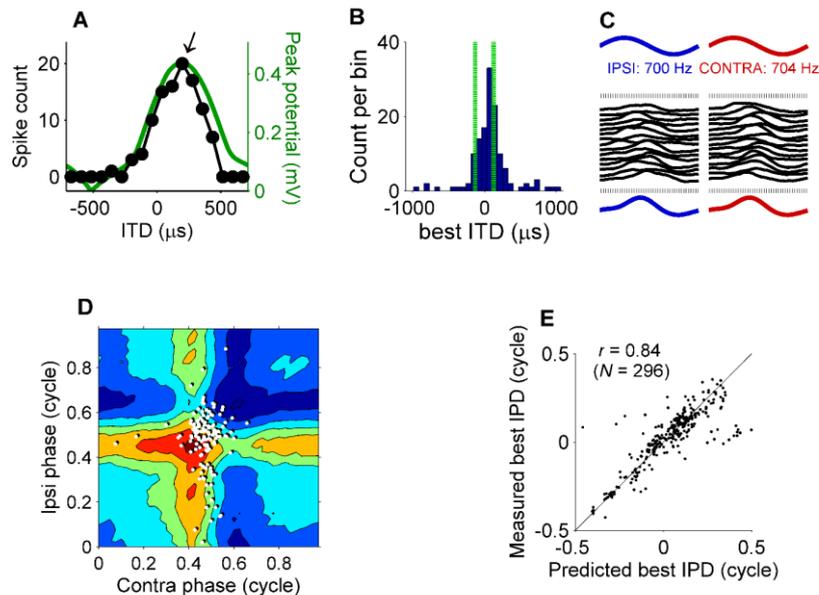


Figure 3. ITD tuning of MSO cells and their subthreshold input.

(A) Relation between number of triggered spikes and ITD for the recording shown in Figure 1A. Arrow indicates the best ITD at +200 μs . *Green line*, peak of subthreshold potential against ITD (see text). (B) Histogram of best ITD values. The histogram was compiled from the 285 binaural-beat recordings (from 19 cells, CFs ranging from 300 Hz to 930 Hz, median 560 Hz) that showed significant (Rayleigh test, $p < 0.001$) phase locking to the 4-Hz beat frequency. Vertical green lines mark the $\pm 130 \mu\text{s}$ physiological range of ITDs. (C) Phase-locked averaging of recordings. After removal of the action potentials, the same binaural beat response was divided in snippets having either the 700-Hz period of the ipsilateral stimulus (left) or the 704-Hz contralateral period (right), yielding the ipsi- and contralateral cycle-averages shown in the bottom traces. (D) Two-dimensional representation of subthreshold input (*colored contours*; 0.2-mV spacing) as a function of both monaural phases, obtained by averaging over repeated instants during the stimulation with the same combination of ipsi- and contralateral phase. *White dots*: eAPs from the same recording. (E) Scatter plot of measured best IPDs against predictions derived from subthreshold input.

ITD tuning of MSO neurons can be predicted from their inputs

In Figure 3A (black circles), the number of triggered spikes of the recording of Figure 1A is plotted against ITD, showing a “best ITD” of 200 μs , a “worst ITD” of about $-500 \mu\text{s}$, and a vector strength (a measure for phase locking to the binaural beat) of 0.78. The best ITD of single MSO cells was not constant, but often varied considerably with frequency (Figure S4), providing evidence against the explanation of best ITDs solely by delay lines (Day and Semple 2011). Population data of best ITD showed a bias for contralateral lead ($91 \pm 282 \mu\text{s}$; $n = 285$;

Figure 3B), and 43% of the best ITDs were outside the physiologically relevant ITD range of the gerbil of $\sim 130 \mu\text{s}$ (Brand et al. 2002; Day and Semple 2011; Pecka et al. 2008; Spitzer and Semple 1995).

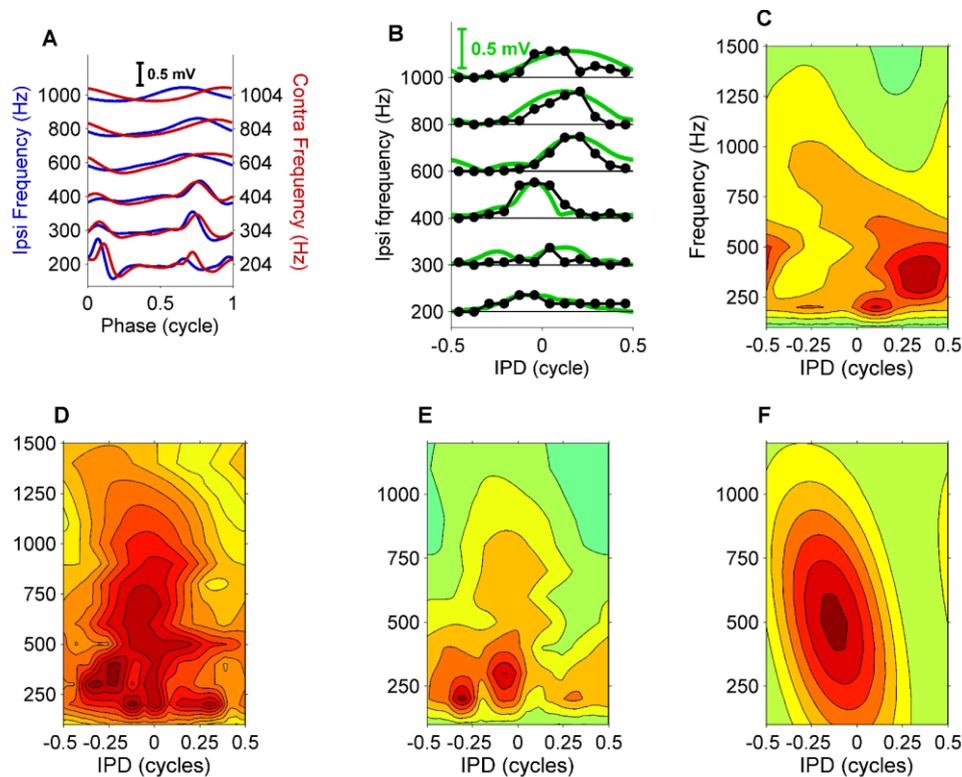


Figure 4. MSO neurons show complex ITD tuning.

(A) Cycle-averaged subthreshold input (as in Figure 3C), multiple frequencies tested. Lowest frequencies show multiple, interaurally matched, preferred latencies in the inputs from both ears. (B) Binaural tuning and its prediction from subthreshold inputs (as in Figure 3A), multiple frequencies tested. Spike count curves (*black symbols*) were normalized to the peak subthreshold potential (*green lines*) to facilitate comparing of their binaural tuning. (C-E) Binaural receptive fields for three MSO neurons (CF = 420, 680, 790 Hz). The peak subthreshold input is shown as a function of both IPD and stimulus frequency, thus combining IPD tuning and frequency tuning of the subthreshold input. Spacing of contours is 0.05 mV. (F) Simple ITD tuning of a hypothetical MSO cell tuned at 500 Hz having a constant, frequency-independent best ITD of $250 \mu\text{s}$. The constant best ITD corresponds to a best IPD that is proportional to the stimulus frequency.

Such tuning beyond the physiological range is consistent with the idea that ITDs follow a “slope” code (Grothe et al. 2010). To resolve whether ITD tuning can be predicted from the inputs (Jeffress 1948), we determined the cycle-averaged subthreshold response for both ears. We removed the eAPs and separately averaged the recording across the cycles of the

respective frequencies presented to each ear (Figure 3C). The latency between the peaks of the two averages thus obtained was 190 μs , close to the observed best ITD of 200 μs .

During its 250-ms cycle, the 4-Hz binaural beat stimulus traverses all possible combinations of ipsi- and contralateral phase, allowing a two-dimensional representation of the subthreshold input as a function of both monaural phases (Figure 3D). The horizontal and vertical ridges in this graph reveal the phase locking of the binaural subthreshold response to the ipsi- and contralateral tone, respectively. The crossing point of these ridges combines the favored phases of both ears, and the peak created by this combination of monaural phases is where one expects the eAPs. The actual timing of eAPs (white dots in Figure 3D) was slightly offset relative to the peak. The direction and magnitude of this offset represents an average latency of 158 μs between peak subthreshold input and APs, consistent with the average EPSP-AP latency of this recording of 173 μs . Thus, Figure 3D shows that subthreshold responses predicted ITD tuning well.

ITD tuning of MSO neurons is complex

Binaural tuning of the subthreshold input was further analyzed by determining, for each value of IPD, the peak potential of the portions of the recording corresponding to that IPD, (i.e., the maximum across diagonal sections of Figure 3D). The IPD-dependence of this peak potential is shown in Figure 3A (*green line*) along with the cycle histogram of eAPs. Again, the binaural tuning of the spikes matches the binaural tuning of the subthreshold input quite well. Figure 3E compares measured best ITDs with predictions from the subthreshold input (as exemplified by the peak of the green curve in Figure 3A) for all our recordings having significant (Rayleigh test, $p < 0.001$; 22 cells, including 3 cells recorded in whole-cell mode) binaural tuning. The correlation $r = 0.84$ confirms the predictability of binaural tuning from the monaural inputs.

The shape of the cycle-averaged subthreshold inputs varied with stimulus frequency (Figures 4A and S5), higher frequencies yielding sinusoidal shapes similar to the intracellularly recorded subthreshold waveforms in nucleus laminaris cells of the barn owl (Funabiki et al. 2011). Responses to low-frequency (<500 Hz) stimuli often showed multiple peaks per tone cycle (e.g., Figure 4A, 200/204-Hz responses). Analysis of SBC recordings previously recorded

in our lab suggested that multiple peaks could already be present in individual inputs to the MSO neurons (Figure S6). Interestingly, the multiple peaks were often matched between the inputs of both ears (Figures 4A and S3). We also expanded the analysis of binaural tuning of the subthreshold input (green curve in Figure 3A) to multiple frequencies (Figure 4B). When

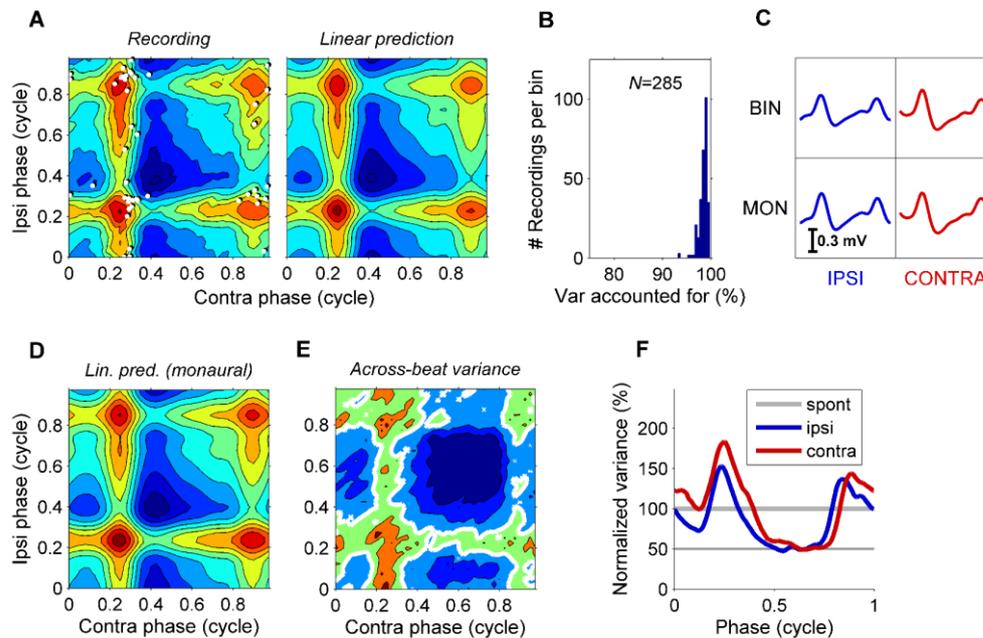


Figure 5. Inputs from both ears sum linearly.

(A) 2D representation of the subthreshold inputs (*left*), with eAPs indicated as *white dots*. *Right* panel is the prediction of a purely linear interaction obtained by adding the cycle averages of the two ears as shown in Figure 3D. Stimulus: 300/304-Hz, 70 dB-SPL binaural beat. Contour spacing 0.1 mV. CF: 680 Hz. (B) Histogram of variance explained by the linear prediction. Population data from 19 MSO cells. (C) Cycle-averaged input waveforms obtained with binaural stimulation (*top row*) and consecutive monaural presentation of the same tones (*bottom row*). (D) Prediction of the subthreshold input of panel A obtained by simply summing the waveforms obtained under monaural stimulation shown in panel C, bottom row. (E) Across-beat-cycle variance corresponding to the across-beat-cycle mean shown in panel A. Contour spacing 0.018 mV². The *thick white contour line* demarcates the variance of spontaneous activity. Most of the time (64 %) during binaural-beat stimulation, the variance is below the spontaneous variance. (F) Across-tone-cycle variance obtained from monaural responses to 70-dB-SPL, 300/304 Hz, normalized to the spontaneous value. For both ears, the variance is periodically reduced to ~50% of the spontaneous value.

displayed as contour plots (Figure 4C-E), these data yield a *binaural receptive field*, in which the effects of stimulus frequency and interaural phase are combined. If a constant, frequency-independent time difference between the inputs existed, the binaural receptive field would

show a single, elongated ridge having a skewed orientation (Figure 4F), because a fixed delay causes a phase shift that is proportional to frequency. The actual binaural receptive fields (Figure 4C-E) do not have this simple form, revealing the complex, frequency-dependent, binaural tuning of the subthreshold inputs.

Linear summation of inputs from both ears

The ability to measure the inputs to the MSO neurons *in vivo* allowed us to test how inputs from both ears sum. To this end, we compared the measured averaged response during the beat cycle with the prediction from a purely linear interaction of the monaural contributions obtained by averaging across the respective ipsi- and contralateral tones (Figure 5A, Movie S1). The observed responses closely followed the linear prediction, which accounted for 97.9 % of the variance. The success of the linear prediction was a general finding, and was observed for both juxtacellular and whole-cell recordings (Figures 5B, S8). Careful inspection of the raw traces did not reveal fast, downward going events that specifically preceded the positive events, both in whole-cell and in juxtacellular recordings (Figures 2B, S3). Simultaneous juxta- and whole-cell slice recordings indicated that the resolution of the juxtacellular recordings allows detecting IPSPs with an amplitude <1 mV (Figure S2). We therefore did not find evidence for well-timed inhibition, nor for a substantial effect of the current sink presented by the non-stimulated dendrite. To further test this linearity, we compared the binaural beat response with the responses to monaural stimulation using the same tones as in the binaural beat stimuli (Figure 5C). Summing the monaural responses provided an excellent prediction of the binaural responses (Figure 5D), accounting for 95.5% of the variance. The small deviations are analyzed in Figure S7.

Lack of excitatory inputs contributes to low firing rates at “worst ITD”

Previous, extracellular recordings from MSO have shown that firing rate at the “worst ITD” is generally lower than the rates obtained by monaural stimulation of either ear, and can even drop below the spontaneous rate (Goldberg and Brown 1969; Spitzer and Semple 1995; Yin and Chan 1990). We observed that subthreshold responses were highly stereotyped, repeating themselves each beat cycle. We therefore determined not only the mean

subthreshold potential (Figure 5A), but also the variance across beat cycles (Figure 5E). The across-cycle variance varied systematically during the cycle. It was clearly larger when responses were large, but in between, during 64% of the beat cycle, it systematically dropped below the spontaneous level. At its absolute minimum, it amounted to only 5% of the spontaneous level. Especially since the inhibitory inputs are large and few (Couchman et al. 2010), the deep trough of the across-beat-cycle variance appears to signify an absence of excitatory inputs rather than the presence of well-timed inhibition. More examples are shown in Figure 6. The periodic reduction of the variance below the spontaneous value was observed in all 22 ITD-sensitive cells. Considering the excellent phase locking of SBCs (Joris and Smith 2008), the most likely interpretation of this phase-locked variance trough is the periodic absence of SBC inputs at those instants where the silent intervals from both ears coincide. This hypothesis is supported by the observation that the periodic reduction of the variance (Figures 5E and 6E-H) became less pronounced for higher stimulus frequencies (Figure S9A), as expected from a decline of SBC phase locking. We also determined the variance across stimulus cycles during monaural stimulation. For both ipsi- and contralateral stimulation, the minimum variance during the cycle was ~50% of the spontaneous level (Figure 5F), consistent with the periodic absence of synaptic inputs from the stimulated ear. Apparently, in this cell the input from each ear contributed ~50% of the total variance of the spontaneous activity. The periodic reduction of variance below spontaneous levels upon monaural stimulation of either ear was a general finding (546/559 recordings; all 18 cells monaurally tested, including two cells recorded in whole-cell mode). Again, the reduction of activity during the unfavorable part of the stimulus cycle became less pronounced with increasing frequency (Figure S9B). We conclude that, most likely, the low firing rate at worst ITD is primarily due to the absence of spontaneous excitatory inputs, whose random timing leads to “accidental coincidences” under monaural stimulation (Colburn et al. 1990).

Inputs from both ears have similar risetimes and EPSP-AP delays

We next tested the predictions of two other models suggesting that ITD tuning is not primarily determined by the timing of the excitatory inputs. Firstly, we did not find evidence for an asymmetry in the risetimes of ipsi- and contralateral responses (Figure 7A; a similar lack of asymmetry was observed for the whole-cell data), in contrast to a slice study, which found

that the slopes of EPSPs evoked by ipsi- or contralateral stimulation differed substantially (Jercog et al. 2010). Secondly, we did not find evidence for an interaural asymmetry in the delay between EPSPs and action potentials (Figure 7B), which could shift ITD tuning (Zhou et al. 2005).

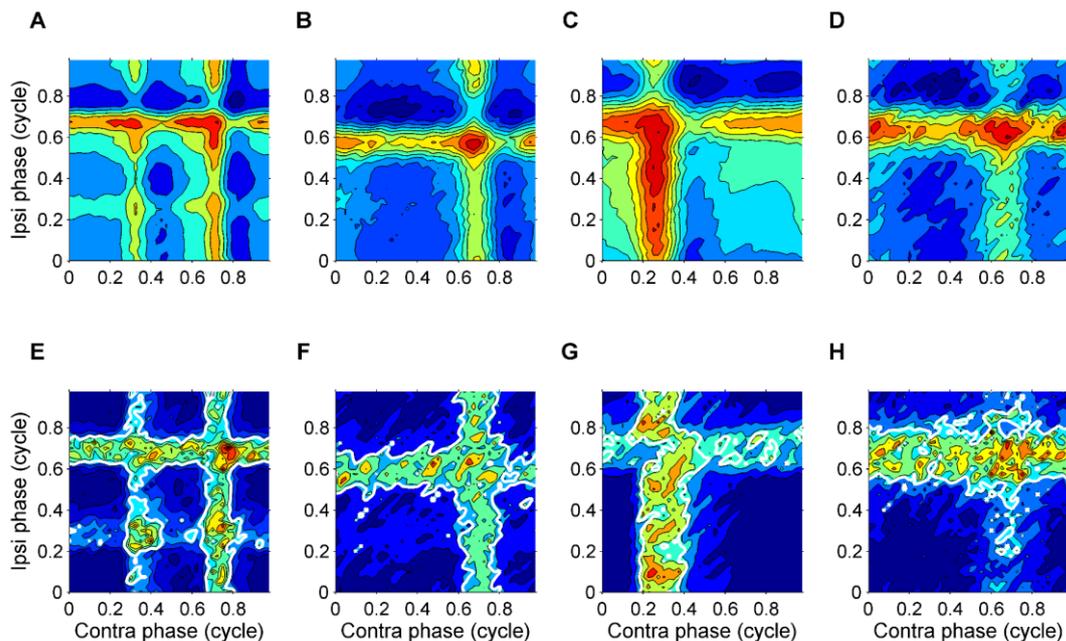


Figure 6. Variance in the response to binaural beats during the beat cycle.

In each column the upper and lower graphs show the mean subthreshold input and its variance, respectively. **(A-D)** Mean subthreshold input as a function of both monaural phases of the binaural beat stimulus in 4 different MSO neurons (cf. Figure 3D). CFs: 680, 430, 300, 480 Hz. **(E-H)** Associated variance (cf. Figure 5E). **(A, E)** 200/204-Hz, 80 dB SPL. **(B, F)** 200/204-Hz, 50 dB SPL. **(C, G)** 300/304-Hz, 60 dB SPL. **(D, H)** 500/504-Hz, 50 dB SPL. The fraction of time that the variance shown in panels **E-H** was below the spontaneous value was 71%, 71%, 78%, and 72%, respectively. The minimum values of the variance over the beat cycle were 5%, 16%, 6%, and 6% of their respective spontaneous values.

Nonlinear input-output relation helps in coincidence detection

The remarkably linear interaction between the inputs from both ears raises the question how the output of these cells can have such good sensitivity to ITD. Figure 8A illustrates how subthreshold monaural inputs can interact to trigger a spike. Binaural stimulation at best ITD evoked on average more than 3 times as many spikes as the sum of monaurally evoked spike

counts (Figure 8B; (Goldberg and Brown 1969; Spitzer and Semple 1995; Yin and Chan 1990). The subthreshold responses in our binaural recordings allowed us to study the relation between the averaged subthreshold potential and the instantaneous firing rate. This relation followed a power relation (Figure 8C), indicating that the nonlinear spike triggering mechanism helps the MSO neurons to be coincidence detectors.

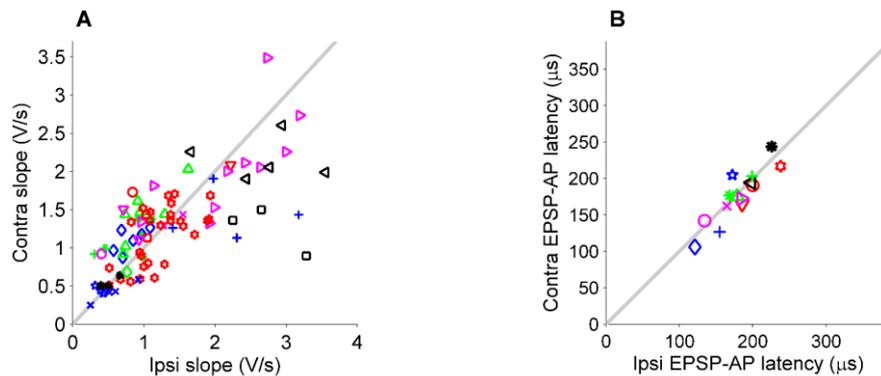


Figure 7. Interaural symmetry of recorded waveforms.

(A) Cycle-averaged ipsilaterally and contralaterally evoked EPSPs have similar risetimes. For each binaural beat response, the steepest slopes of the cycle-averaged subthreshold input (cf. Figures 3B, C; 5C, F) were determined, using either the ipsilateral or the contralateral stimulus frequency. Peak-to-peak values of both cycle-averaged subthreshold responses had to exceed twice the RMS of the spontaneous activity in order to be included, yielding $n = 89$ recordings, 19 cells (juxtacellular recordings). The ipsi/contra pairs of steepest slopes are shown as a scatter plot, each cell indicated by a different symbol. The mean pair-wise difference (ipsi minus contra) was 0.05 ± 0.45 V/s ($p > 0.43$, Student's T-test). (B) EPSP-AP latencies in monaural responses. For the 14 cells for which monaural responses to both ears were available (1 whole-cell, 13 juxtacellular recordings), we compared the latency between EPSPs and APs (cf. Figure 1E, F) and compared them between ipsilateral and contralateral tones. Each symbol represents the average of all monaurally evoked APs of one cell. The mean difference (ipsi minus contra) across cells was 5 ± 16 μ s. A Student's T-test revealed no significant difference between ipsi- and contralaterally evoked EPSP-AP latencies ($p > 0.47$; $n = 14$ cells).

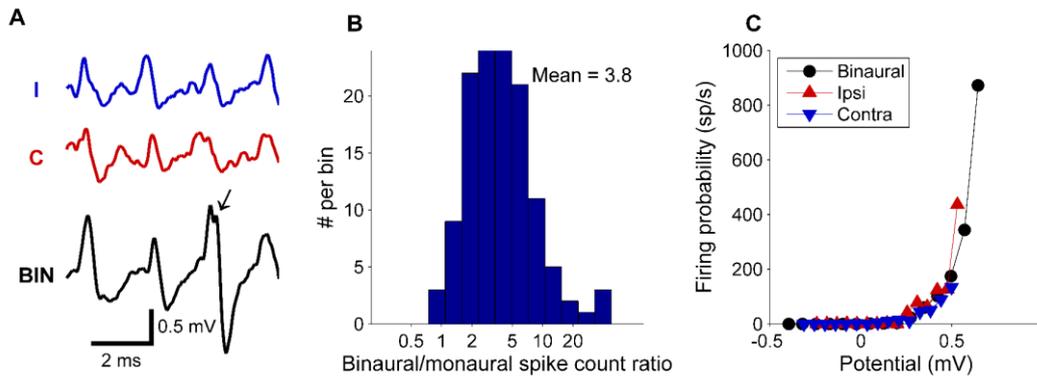


Figure 8. Coincidence detection is realized by a nonlinear input-output relation.

(A) Comparison of monaural responses (upper two traces) and binaural responses (lower trace). Frequencies were 500 Hz (ipsilateral) and 504 Hz (contralateral). All traces show a periodic (~ 2 -ms) depolarization. The larger binaural responses are closer to the firing threshold and trigger an AP (arrow). CF: 680 Hz. (B) Histogram of the ratio of number of spikes evoked by binaural stimulation at best ITD to the sum of monaurally evoked spikes ($n = 18$ cells). The mean value of 3.8 indicates a sizeable binaural facilitation. (C) Instantaneous firing rate as a function of the averaged subthreshold potential derived separately for monaural and binaural stimulation, showing an expansive (“power-law”) relation.

Discussion

The ability to measure the synaptic inputs to the MSO neurons allowed us to study how these neurons integrate information from both ears. We show here that ITD tuning of these neurons is determined by the timing of their excitatory inputs, that these fast excitatory inputs from both ears sum linearly, and that spike probability depends nonlinearly on the size of synaptic inputs.

Subthreshold events in the MSO

We used a juxtacellular approach to record from MSO neurons in vivo. In contrast to earlier studies in gerbil (Brand et al. 2002; Day and Semple 2011; Pecka et al. 2008; Spitzer and Semple 1995), we used a ventral approach, which made it easier to map where the MSO cell layer was located. The use of field potentials (Galambos et al. 1959; Mc Laughlin et al. 2010)

was critical for determining the cell layer. Within the somatic layer, all cells were excited by both ears, whereas several previous studies found that many cells were inhibited by one ear (Barrett 1976; Caird and Klinke 1983; Goldberg and Brown 1968; 1969; Hall 1965; Moushegian et al. 1964a). Even though our sample size was limited, and there may be species differences, this suggests that some of the reported heterogeneities in the properties of MSO neurons are caused by differences in response properties between MSO neurons within and outside of the somatic layer (Guinan et al. 1972; Langford 1984; Tsuchitani 1977).

The recordings from the MSO neurons were characterized by the presence of clear subthreshold responses, even in the absence of sounds, and by the presence of low-amplitude spikes. The observation that the spontaneous events could be picked up even in the juxtacellular recordings is partly due to their low membrane resistance, which is caused by the presence of I_h and low-threshold K^+ channels already open at rest (Khurana et al. 2012; Khurana et al. 2011; Mathews et al. 2010; Scott et al. 2005). In agreement with this, the resistive coupling measured in simultaneous juxtacellular and whole-cell recordings was much larger than in principal neurons of the MNTB, whereas the capacitive coupling was similar (Lorteije et al. 2009). The small size of the somatic action potential is in agreement with slice recordings (Scott et al. 2005), and is caused by the restricted backpropagation of the axonal action potential to the soma (Scott et al. 2007). The high spontaneous event rates of at least 500 events/s were in agreement with average spontaneous firing rates of SBCs of ~ 56 sp/s (Kuenzel et al. 2011) and the estimate of minimally 4-8 SBCs innervating each gerbil MSO neuron (Couchman et al. 2010). The EPSP kinetics largely matched results obtained with slice recordings. Half widths of EPSPs in juxtacellular recordings were somewhat smaller than in adult slice recordings (~ 0.55 ms; (Scott et al. 2005), to which the capacitive component in the juxtacellular recordings may contribute, whereas the intracellularly recorded EPSPs had a half width that was somewhat larger than of EPSPs in slice recordings, to which both dispersion in sound-evoked events and the relatively large series resistances may have contributed.

Variability of EPSP-AP delays

The EPSP-AP delay was remarkably variable and was on average about 200 μs , which is larger than the physiological ITD range of the gerbil. Similar delays have been observed in a slice study (Scott et al. 2007). This delay consists of the travel time of EPSP to initial segment, spike initiation and the backpropagation of the AP to the soma, which is physiologically less relevant. The EPSP-AP delay depended systematically on EPSP amplitude (Scott et al. 2007); larger EPSPs resulted in smaller EPSP-AP delays, in agreement with the idea that the EPSP-AP delay of EPSPs that are barely suprathreshold contribute considerably to jitter, as was also found in the SBCs, which form the excitatory inputs to the MSO neurons (Kuenzel et al. 2011).

Linear summation of inputs from both ears

The ability to measure the inputs to the MSO neurons *in vivo* allowed us to test how inputs from both ears sum. We found that the interaction between the inputs from both ears was remarkably linear. The ipsilateral EPSP did not depend on the phase of the contralateral EPSP (and vice versa).

Our data are in good agreement with experiments in neocortical and hippocampal slices, in which a general finding was that distant inputs sum linearly, whereas inputs on the same dendritic branch interact nonlinearly (Cash and Yuste 1999; Gasparini and Magee 2006; Polsky et al. 2004; Tamás et al. 2002). Linear summation was also observed in an *in vivo* study in visual cortex (Jagadeesh et al. 1993). Apparently, in our *in vivo* experiments the somatic depolarization by the inputs of either ear was not large enough to create a substantial loss of driving force for the inputs from the other ear. The exact cellular mechanisms underlying the remarkable linear behavior of the MSO neurons remain to be investigated, but slice studies have suggested that the interplay of the different voltage-dependent ion channels in the MSO neurons can actively linearize the interaction between binaural inputs (Khurana et al. 2011; Scott et al. 2010).

Implications of linear summation

In a simulation study (Agmon-Snir et al. 1998), it has been proposed that the segregation of the inputs from both ears to opposite dendrites favors binaural inputs over monaural inputs by two different mechanisms. Firstly, inputs from the same ear would tend to sum nonlinearly, because the local depolarization will reduce driving force. Secondly, it would be more difficult for monaural inputs to reach threshold owing to the current sink of the non-stimulated dendrite. The activation of potassium channels might contribute to this non-linear interaction as well (Grau-Serrat et al. 2003; Mathews et al. 2010). The linearity of the summation argues against a prominent role of these mechanisms. Our results do not allow us to infer to what extent inputs sum sublinearly at a single dendrite. However, our results do suggest that the current sink imposed by the non-stimulated dendrite is not very large, since the size of the EPSP from one ear did not depend measurably on the phase of the stimulation to the other ear, and thus on the membrane potential of the other dendrite.

A similar argument can be put forward against the theory that well-timed (phase locked), contralateral, inhibition originating from the MNTB delays the time point at which the action potential threshold is reached (Brand et al. 2002; Pecka et al. 2008). This theory provides an elegant explanation for the observation that best ITDs typically show a bias for contralateral lead, which we also observed in the present study. This theory also predicts a significant phase-dependent influence of the sound from one ear on the response to the sound presented to the other ear, since well-timed inhibition should interact with excitation even if it is entirely of the shunting type. In contrast to these predictions, we found that the timing of the input from either ear is unaffected by the phase of the input from the other ear. Our results therefore suggest that the timing of the inhibitory input from either ear is not sufficiently precise to allow it to shift the ITD tuning (Joris and Yin 2007; Zhou et al. 2005). This argument still holds true in the presence of inhibition from both ears. We cannot entirely exclude that the use of anesthetics may have influenced the timing precision of the inhibition. Effects of ketamine/xylazine on subcortical auditory processing are typically mild (Smith and Mills 1989; Ter-Mikaelian et al. 2007), and both bushy cells (Kuenzel et al., 2011) and primary neurons of the MNTB (Hermann et al. 2007) in gerbil show considerable spontaneous activity even under ketamine/xylazine anesthesia. Decreased inhibition has been reported in the dorsal cochlear nucleus (Navawongse and Voigt 2009). However, the original evidence

favoring well-timed inhibition was also obtained under ketamine/xylazine anesthesia (Brand et al. 2002; Pecka et al. 2008). Another possible confounder is that most of the inhibition is somatic, and may have been disrupted when we made recordings. However, somatic inhibitory responses in the MSO are not disrupted by positive pressures at least ten fold higher than what we used during approach of cells for juxtacellular recordings (Couchman et al. 2012).

Possible role of inhibition in MSO

The presence in the MSO of strong glycinergic inhibitory inputs originating from both the ipsi- (LNTB) and contralateral ear (MNTB) is well established, but its function has been debated (reviewed in (Grothe et al. 2010)). Because of the linearity of the interaction between both ears, a role of well-timed inhibition in shifting the best ITD (Brand et al. 2002; Pecka et al. 2008) seems unlikely. The low variance at the worst ITD suggests that it is the periodic absence of excitatory input rather than phase-locked inhibition that sets the firing rate during the worst ITD. A possible role for inhibition is that it may improve the dynamic range of the MSO neurons, similar to its proposed role in the nucleus laminaris (Yamada et al. 2013), the avian equivalent of the MSO, and in the SBCs (Kuenzel et al. 2011). This role is in agreement with the strong increase in spontaneous activity, clear broadening of ITD tuning and strongly reduced effect of ITD on spike rate observed upon application of the glycine receptor antagonist strychnine (Brand et al. 2002; Pecka et al. 2008) and the relatively slow kinetics of glycinergic synaptic potentials compared to the glutamatergic synaptic potentials (Magnusson et al. 2005).

Lack of contribution of MSO neurons to internal delays

Apart from the lack of evidence for a role of well-timed inhibition, we also did not find support for the two other models that propose that MSO neurons contribute to the creation of internal delays. The suggestions that interaural asymmetries in synaptic potentials (Jercog et al. 2010) or cellular morphology (Zhou et al. 2005) may contribute to ITD tuning of MSO cells are contradicted by our observation that the slopes of subthreshold inputs were similar for

both ears (Figure 7A), in agreement with a recent slice study (Fischl et al. 2012), and we obtained a similar result for the EPSP-AP latencies (Figure 7B). The interaural symmetry of EPSP-AP latencies agrees with the observation that in the gerbil MSO axons typically emerge directly from the soma (Scott et al. 2005). Our data therefore indicate that ITD tuning depends critically on the exact timing of the excitatory inputs to the MSO neurons, and that the MSO neuron itself does not make a large contribution to the internal delay.

ITD tuning

ITD tuning was complex. Two features were remarkable. Firstly, at low sound frequencies we observed multiple preferred latencies in the responses for both ears. Most likely, this is inherited from the SBCs. Spike timing dependent plasticity has been suggested as a possible mechanism for the coincidence of these inputs (Gerstner et al. 1996), and our results suggest that, if it is, it can work for multiple preferred latencies, indicating a hitherto unknown complexity to the tuning of the MSO neurons. It should be noted that these multiple latencies were typically obtained at low frequencies and high intensities, so their contribution to natural stimuli remains to be established. Behaviorally, localization is poorer for pure tones than for more “natural”, wideband sounds. Future work using wideband stimulation is required to test how our findings generalize to a wider range of stimuli.

A second property that added to the complexity of the tuning was that a comparison of the inputs from both ears indicated that ITD tuning was frequency-dependent. This observation by itself argues against the original Jeffress model (Jeffress 1948), in which a delay line was the only source for ITD tuning. Since we did not observe any evidence for a contribution of the MSO neurons themselves to the delay line, this is compatible with the idea that cochlear tuning disparities contribute to the creation of internal delays (Day and Semple 2011; Joris et al. 2006).

Coincidence detection

The backbone of the Jeffress hypothesis is the presence of coincidence detectors, neurons that fire when inputs from both ears arrive at the same time (Jeffress 1948). Experimental

evidence for this hypothesis was obtained from recordings in which the ITD was systematically varied (Goldberg and Brown, 1969). A key finding was that the best ITD could be predicted from the preferred latencies of the monaural responses. Our data extend these findings in three ways. Firstly, we show that the best ITD can be well predicted from the timing of the monaural subthreshold responses. Secondly, we provide a simple explanation for the low firing rate during the worst ITD. The observation that during worst ITD the firing rates become lower than during the response to monaural stimulation in many cells was basically unexplained. Three possibilities have been put forward: a role for well-timed inhibition (Yin and Chan 1990), a role for low-threshold potassium conductance which is activated during depolarizations (Grau-Serrat et al. 2003; Mathews et al. 2010) or the absence of active excitatory inputs because of good phase locking (Colburn et al. 1990). A variance analysis provided evidence favoring the latter possibility, although a specific role of inhibition, low-threshold potassium channels or a combination of the two in the very low firing rates during the worst ITD cannot be excluded. Thirdly, to function as good coincidence detectors, MSO neurons must have a clearly higher spike rate at the best ITD for binaural stimulation than the sum of the spike rates during monaural stimulation of the left and the right ear. We observed a supralinear relation between firing rate and the averaged subthreshold potential (Figure 8C), which is in agreement with the power-law relation between spike probability and membrane potential in other neurons (Silver 2010). This non-linear relation has the effect to greatly increase the probability that a spike is triggered when EPSPs from both ears arrive at the same time. Together, our results indicate that binaural facilitation in MSO neurons results from the nonlinear increase in spiking probability brought about by the linear sum of the inputs from the two ears.

Experimental Procedures

Animal procedures

All experiments were conducted in accordance with the European Communities Council Directive (86/609/EEC) and approved by the institutional animal ethics committee. After brief exposure to isoflurane, a total of 11 young-adult Mongolian gerbils (84 ± 7 days postnatal; 50-70 g) were injected intraperitoneally with a ketamine-xylazine mixture (65/10 mg/kg). Anaesthesia was monitored with the hind limb withdrawal reflex and additional ketamine-xylazine was given to maintain anaesthesia. Rectal temperature was maintained between 36.5 and 37.5 °C with a homeothermic blanket system (Stoelting Co., Wood Dale, Ill.). Both pinnae were surgically removed. We used a ventral approach to reach the MSO. Animals were supine-positioned, with their heads immobilized by a metal pedestal glued to the dorsal skull. Skin and soft tissue overlaying the trachea were removed and the trachea was intubated. Animals continued breathing independently. The right bulla was opened fully using a forceps; a hole was made in the left bulla to prevent pressure buildup in the left middle ear. Based on cranial landmarks, a ~1 mm diameter craniotomy was created by carefully scraping the bone between the bulla and the brainstem with a small handheld drill, exposing the brain surface slightly laterally from the MSO. Dura, arachnoids and pia mater were removed locally.

Histology

In some experiments, recording locations were marked with biocytin (0.5%) which was added to the pipette solution, or with post-recording injection of saturated Alcian Blue at the recording position (Figure S1). In these experiments, animals were sacrificed with a lethal dose of Nembutal and subsequently perfused intracardially with saline, followed by a 4% paraformaldehyde solution. Brains were further processed as described in (Horikawa and Armstrong 1988) with minor modifications. Histology confirmed MSO as the recording location in 6 of 6 animals.

In vivo electrophysiology

Thick-walled borosilicate glass micropipettes with filament had a resistance of 3.5–6 M Ω when filled with recording solution. Pipettes were filled with Ringer solution for juxtacellular recordings, which contained NaCl 135, KCl 5.4, MgCl₂ 1, CaCl₂ 1.8, HEPES 5 mM; for whole-cell recordings the pipette contained (in mM): 138 K-gluconate, 8 KCl, 0.5 EGTA, 10 HEPES, 10 Na₂Phosphocreatine, 4 MgATP, 0.3 NaGTP (pH 7.2 with KOH). Electrodes were typically inserted laterally (and ventrally) from the cell layer and advanced in dorsomedial direction at an angle of 20-30 degrees with the vertical. The thin somatic layer (Rautenberg et al. 2009) was identified based on the polarity reversal of the local field potential response ('neurophonics') during alternating monaural click stimuli to the left and right ear (Figure S1; (Biedenbach and Freeman 1964; Clark and Dunlop 1968; Galambos et al. 1959).

Pipettes had a high positive pressure (>300 mbar) when crossing the brain surface, which was lowered to 10-30 mbar when approaching the cell layer (located at 400-1000 μ m from the surface). Juxtacellular (loose-patch) or whole-cell recordings were made by slowly advancing the pipette while monitoring both its resistance and the presence of EPSP or spike activity. For juxtacellular recordings, pressure was released if a neuron was approached, and slight negative pressure was briefly applied while moving the electrode another 2 to 10 μ m towards the cell until pipette resistance increased to a value of typically 30 M Ω . Because physical contact with a cell is essential for the large size of the juxtacellular potentials (Lorteije et al. 2009), we consider it very unlikely that another, nearby cell contributed significantly to the measured potentials. A further argument supporting good unit isolation was that the shortest eAP interval that we observed in any of the juxtacellular recordings was >1 ms, as expected from recordings from single neurons obeying the refractory period. Details of the whole-cell in vivo recordings are described in *Supplemental Experimental Procedures*.

Data were acquired with a MultiClamp 700B patch-clamp amplifier and pCLAMP 8 software (Axon Instruments). Further details are described in *Supplemental Experimental Procedures*.

Slice electrophysiology

Dual somatic whole cell and juxtacellular recordings were made at 37 °C from MSO neurons in 200 µm horizontal slices prepared from P29-46 gerbils as described previously (Scott et al. 2005). Slices were bathed in ACSF containing (in mM): 125 NaCl, 25 glucose, 25 NaHCO₃, 2.5 KCl, 1.25 NaH₂PO₄, 1.5 CaCl₂, 1.5 MgSO₄. Whole cell recording electrodes were filled with (in mM): 115 K-gluconate, 4.42 KCl, 0.5 EGTA, 10 HEPES, 10 Na₂Phosphocreatine, 4 MgATP, 0.3 NaGTP. Juxtacellular recording electrodes were filled with the same solution used for in vivo juxtacellular recordings. Juxtacellular seal resistance averaged 24 ± 7 MΩ. EPSPs were evoked by local stimulation of excitatory afferents in the presence of 1 µm strychnine. IPSPs were generated via conductance clamp (Toro-8 digital signal processing board, Cambridge Conductance software) simulation of an inhibitory conductance with a double exponential waveform (time constants = 0.28 ms rise, 1.85 ms decay) and reversal potential of -85 mV. Current steps were delivered through the whole cell electrode. Data were acquired using a MultiClamp 700B amplifier and custom algorithms in IGOR Pro. EPSP data were analyzed by binning both whole cell and juxtacellular responses according to the peak EPSP amplitude measured in the whole cell recording (0.2-0.6 mV bins), then averaging the responses in each bin. Similarly, IPSP data were averaged according to the simulated conductance, and current step data were averaged according to the amplitude of the current step. Comparisons between whole cell and juxtacellular recordings were made using these average responses. Capacitive and resistive coupling constants were estimated as described previously (Lortelje et al. 2009).

Auditory stimulation

Auditory stimuli were generated using custom MATLAB software. Stimuli were generated using a TDT2 system (PD1, Tucker Davis Technologies) and presented in a close-field configuration to the animal with Shure speakers (frequency range 22 Hz to 17.5 kHz) attached to the ear canal via a small tube. The correct stimulus levels and phases were attained by calibrating the drivers in situ at the level of the tympanic membrane using the microphone housed in the probe. The transfer characteristics of the probe were taken into account. All stimuli were generated at a rate of 48.8 kHz.

Binaural beat stimuli consisted of a pair of pure tones, one presented to each ear. The frequencies presented to the ipsilateral ear varied between 100 Hz and 1600 Hz in 100-Hz steps; in two experiments the step size was reduced to 50 Hz. The frequencies presented to the contralateral ear were always 4 Hz above that of the ipsilateral tone. The tones were presented simultaneously to the two ears, lasted 6 or 9 s, including 3-ms \cos^2 onset and offset ramps, and were separated by 1500-ms silent intervals. The initial stimulus level was 60 or 70 dB SPL. If time permitted, additional recordings were performed using additional intensities between 10 and 80 dB SPL in 10 dB steps, and monaural responses were obtained by setting the amplitude of the tone presented to either ear to zero.

Detailed analysis of in vivo recordings

Acceptance criteria, windowing and conditioning of the responses, detection of APs and EPSPs, periodicity analysis (Figures 3C, D; 4A; 5A, C, D, F) and extraction of metrics (vectors strength, CF, instantaneous firing rate) are detailed in *Supplemental Experimental Procedures*.

Acknowledgements

This work was supported by an FP6 European Union grant (EUSynapse) and by the Dutch Fund for Economic Structure Reinforcement (FES, 0908 'NeuroBasic PharmaPhenomics project') (J.G.G.B.), and NIH grants DC006788 and DC011403 (N.L.G.).

Supplemental Information

Supplemental Data

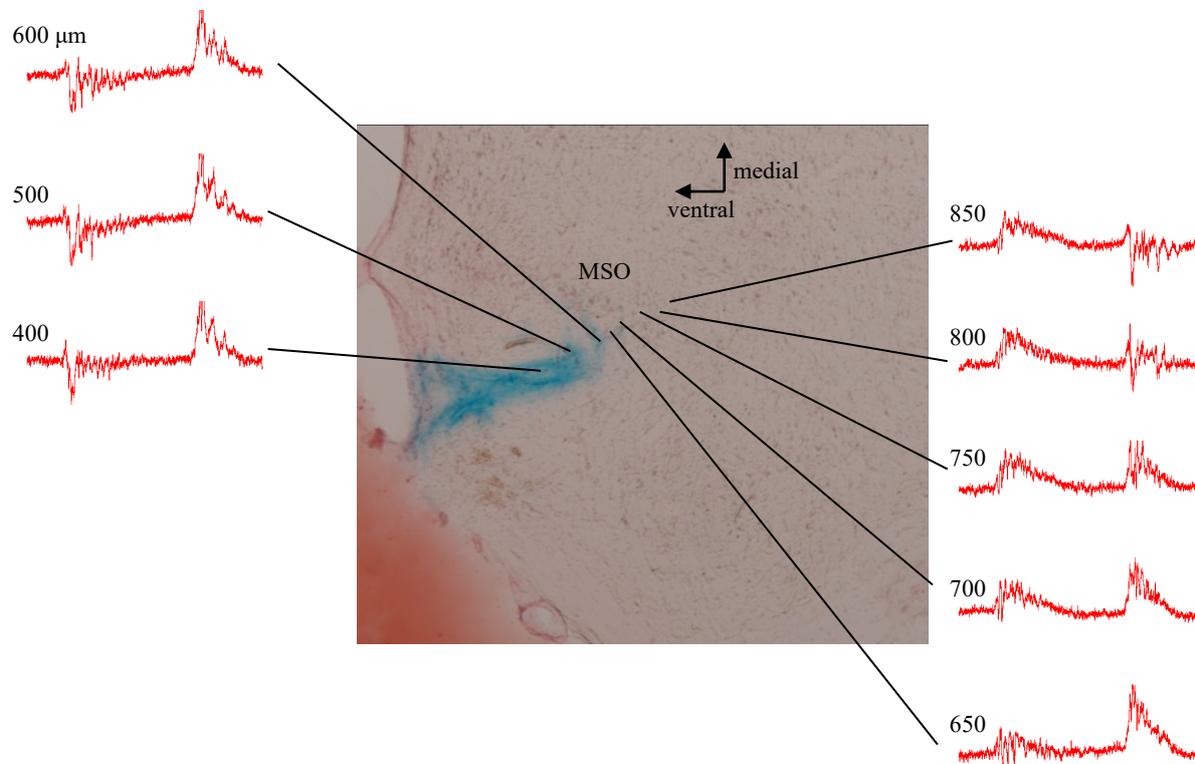


Figure S1 (related to Figure 1). Identification of the somatic layer in the gerbil MSO using a ventral approach.

Section of the ventral brainstem with the electrode track marked in blue by pressure ejection of Alcian Blue. Red traces show the local field potential response ('neurophonics') during alternating monaural click stimuli to the left and right ear (Biedenbach and Freeman 1964; Clark and Dunlop 1968; Galambos et al. 1959). Numbers indicate penetration depth from surface. The somatic layer ('MSO') lies between the point where the response to contralateral tones reverses (~650 μm) and the point where the ipsilateral click response reverses (~850 μm).

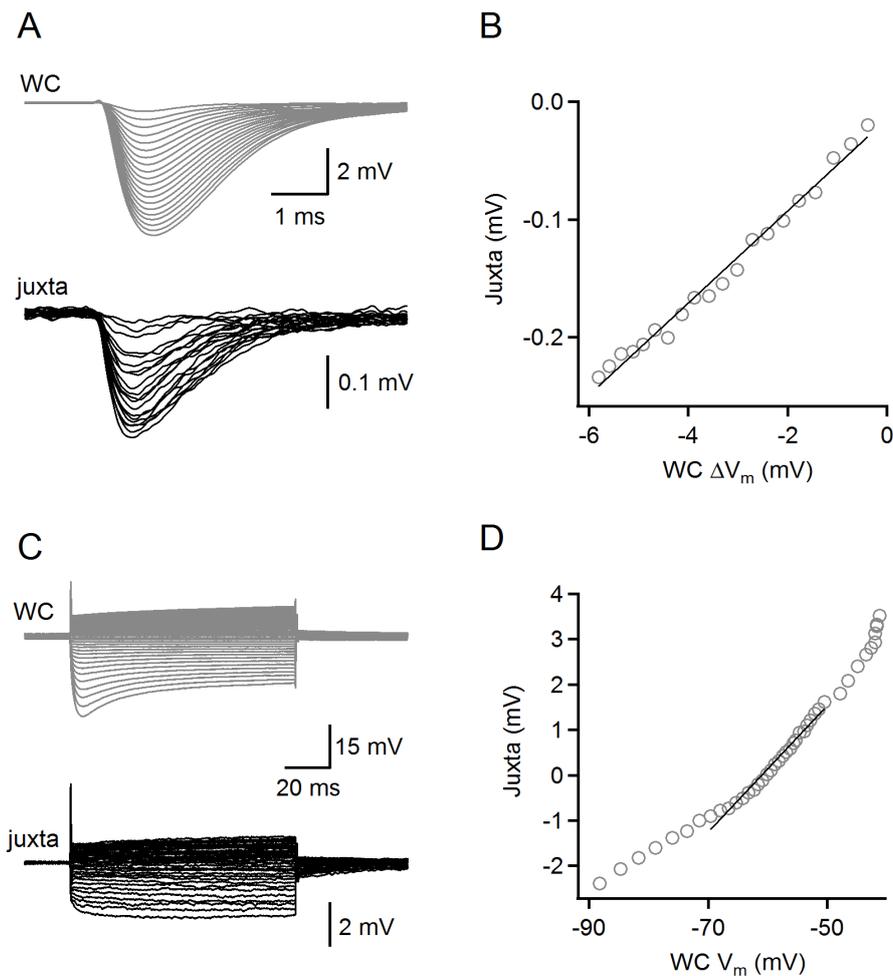


Figure S2 (related to Figure 2). Relation between juxtacellular and whole-cell potentials is linear for IPSPs and for current steps within the physiological membrane potential range.

(A) Sample whole cell and juxtacellular responses to IPSPs generated by conductance clamp simulation of an IPSP over a range of 20 conductance levels. Data are averages of 20 trials. (B) Linear correlation between peak IPSP amplitudes measured with juxtacellular and whole cell recordings from the cell shown in (A) ($r = 0.994$). (C) Sample whole cell and juxtacellular responses to current steps ranging from -5000 to 5000 pA in 200 pA increments. Data are averages of 6 trials. (D) Over the physiologically important voltage range (-70 to -50 mV), peak responses to current steps were linearly correlated between juxtacellular and whole cell recordings from the cell in (C) ($r = 0.989$).

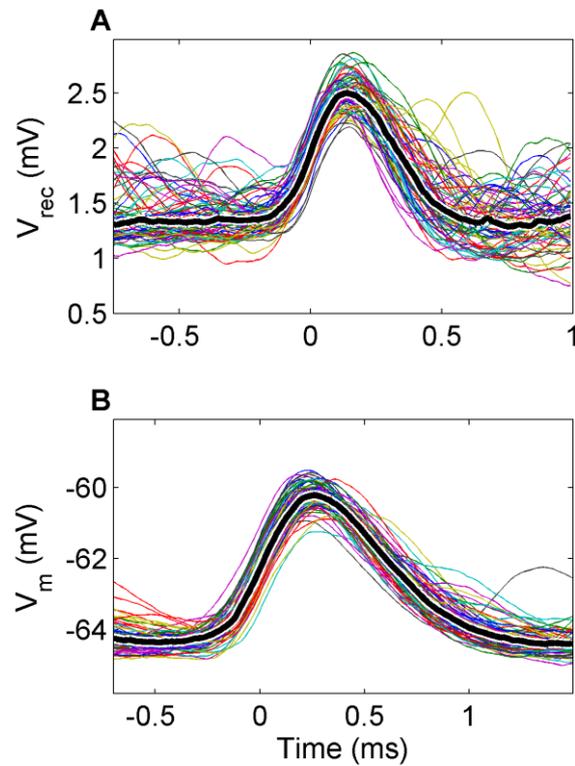


Figure S3 (related to Figure 2). Absence of downward-going phase preceding EPSPs.

Subthreshold events were extracted from the MSO recordings based on their steep upward slopes. Each panel shows the 70 largest events (colored thin lines) and the median value (thick black line) obtained from a different MSO neuron. The events are aligned on their steepest rising phase (time zero). **(A)** Juxtacellular recording; CF = 300 Hz ; **(B)** Whole-cell recording; CF = 780 Hz.

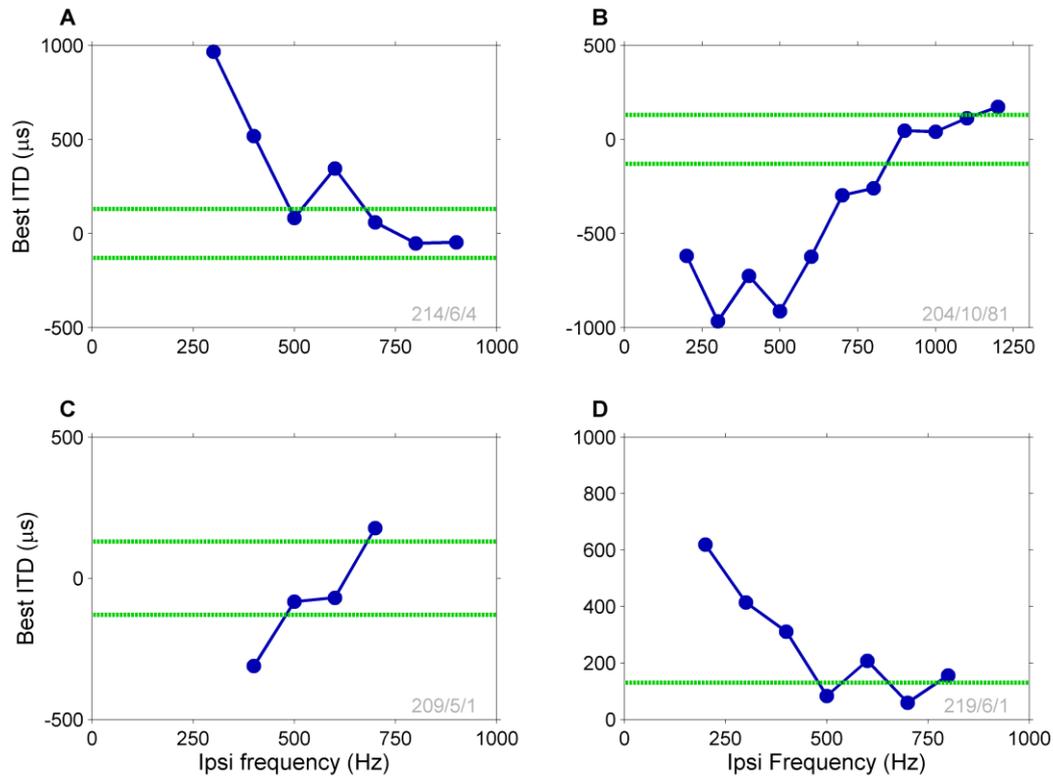


Figure S4 (related to Figure 3). Variation of ITD with stimulus frequency.

The best ITD of many MSO neurons varied considerably with stimulus frequency. Best ITD was evaluated from the peak of the AP cycle histogram derived from responses to 60-dB-SPL, 4-Hz binaural beat stimuli. The four panels show data from four different MSO neurons. Horizontal green lines mark the $\pm 130 \mu\text{s}$ physiological range of ITDs determined by the head size of the gerbil. CFs: 680, 760, 560, 330 Hz.

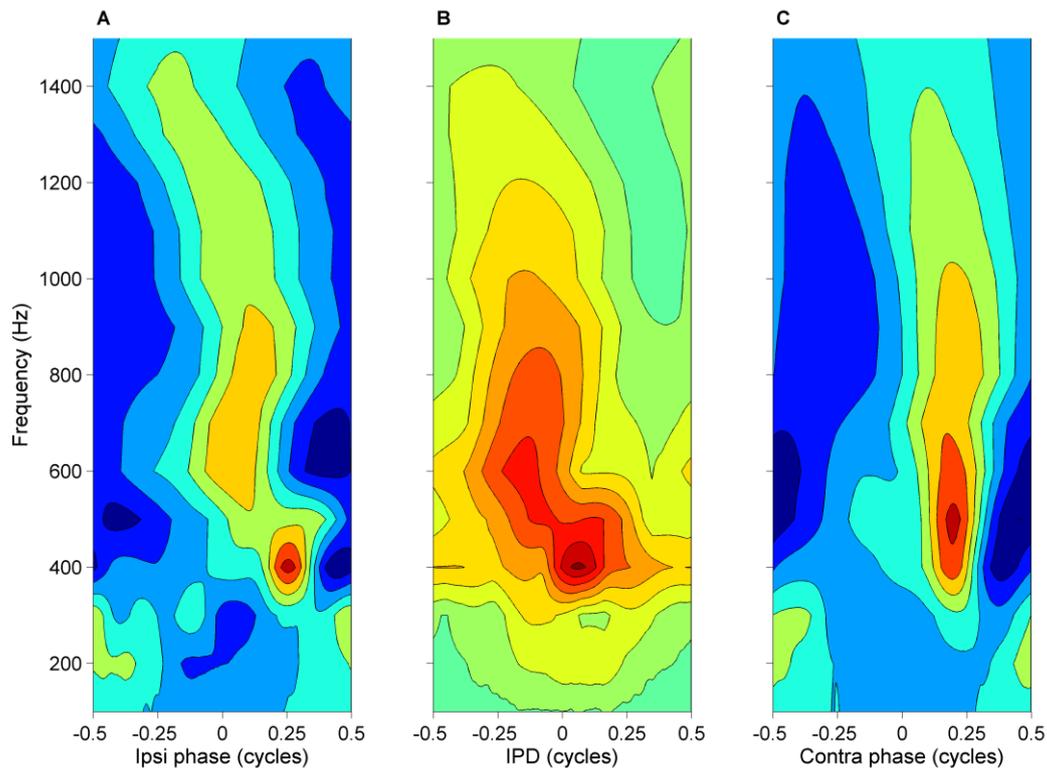


Figure S5 (related to Figure 4). Frequency-dependent monaural and binaural tuning.

(A) and (C) show the monaural cycle-averaged waveforms (a subset of which is shown in Figure 4A) in the form of contour plots (0.2-mV spacing) as a function of both stimulus frequency and the monaural phase relative to the stimulus onset. Monaural phases were compensated for a 5.5-ms overall stimulus-response latency delay estimated from the onset latency of the recording. The subthreshold is phase-locked over the entire range of frequencies, reflected by systematic vertical “ridges.” The steep transition of the monaural phase pattern between 300 Hz and 400 Hz produces a complex phase-frequency relation, which, however, is matched across the two ears. (B) shows the reduction of these monaural phase data to the IPD axis (cf. the green line in Figures 3A and 4B). The resulting “binaural receptive field” shows the tuning of the subthreshold input as a function of frequency and IPD. Contour spacing, 0.1 mV. The interaural matching of the monaural transition causes the binaural receptive field to be more regular than the monaural phase-frequency plots in A and C. The binaural tuning is still complex, and cannot be captured by a single, frequency-independent delay, because that would cause the ridge to be on a straight line through the origin (zero IPD and zero frequency) See also (Day and Semple 2011). CF: 680 Hz.

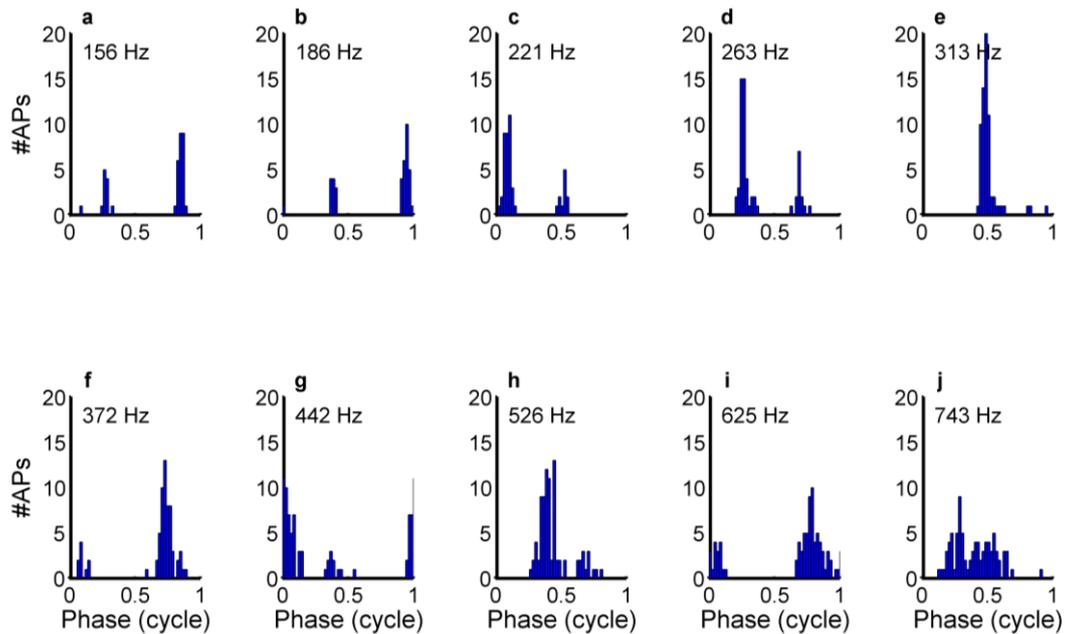


Figure S6 (related to Figure 4). Multiple peaks in cycle histograms of spherical bushy cells (SBCs) of the gerbil.

Unpublished data from unit RG0981/1u1 of (Kuenzel et al. 2011). Cycle histograms of APs in response to 50-ms, 80-dB-SPL single tones, analyzed as described in (Kuenzel et al. 2011). Such “peak splitting”, was a common observation in the SBC responses to low-frequency (<500 Hz) tones at intensities of 50-80 dB SPL. Courtesy of Dr. Thomas Kuenzel, Aachen.

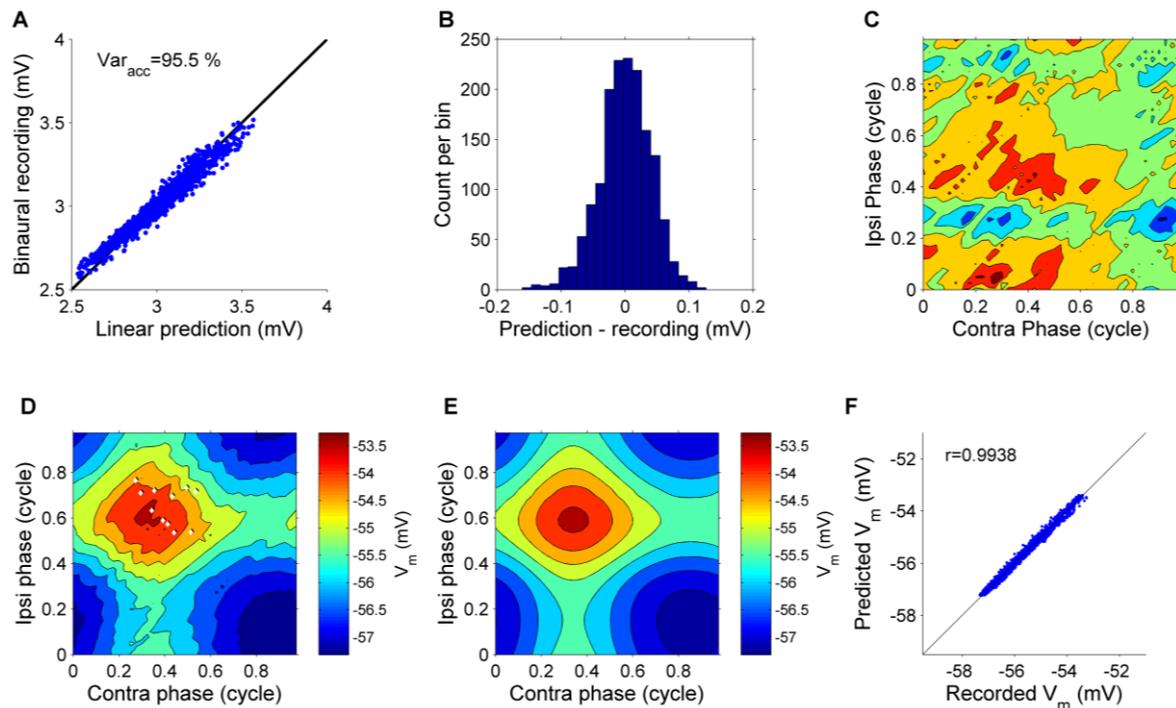


Figure S7 (related to Figure 5). Linear summation of monaural inputs.

(A) Scatter plot comparing linear prediction of the subthreshold input obtained from summing monaural recordings (Figure 5D) with the actual binaural response (Figure 5A). CF: 680 Hz. The black line indicates unity. The linear prediction accounted for 95.5% of the variance in the data. (B) Histogram of the deviations, i.e. data minus prediction. (C) Contour plot of the deviations. Blue patches correspond to data points below the predictions. Contour spacing 0.05 mV. The trend of the predictions to overestimate the highest values (A) the skewness of -0.32 of the distribution (B), and the concentration of negative deviations (blue patches in C) on the peaks of the subthreshold response (compare to Figure 4A) are all consistent with a contribution of a small bias toward smaller EPSPs in the binaural recording caused by the removal of those events that actually triggered APs. The removed segments added up to 0.076% of the total duration of the waveform analyzed. (D) 2-D representation of subthreshold input as in Fig. 2, but now for an in vivo whole cell recording. Stimulus was a 6-s, 700/704-Hz binaural beat at 40 dB SPL. *White dots* indicate action potentials. (E) The prediction from linear summation of the monaural contributions. (F) The linear prediction plotted against the recorded subthreshold response, with the correlation coefficient r indicated in the graph.

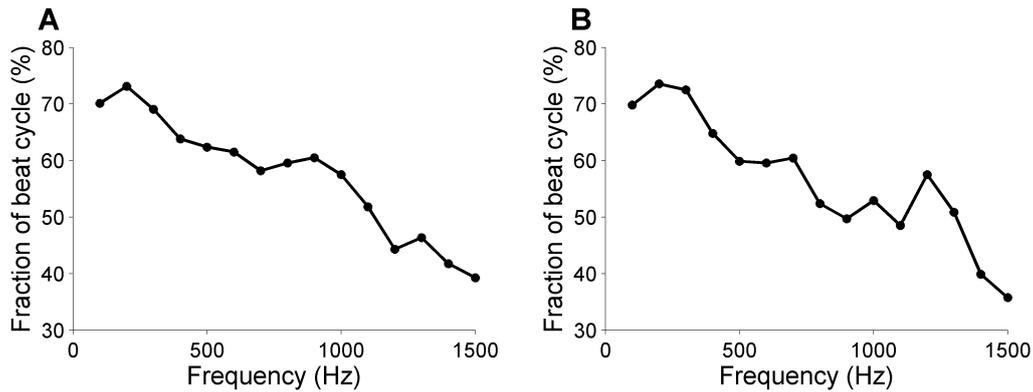


Figure S8 (related to Figure 6). Effect of frequency on the periodic reduction of subthreshold input.

(A) In response to 70-dB-SPL binaural beats having increasing monaural frequencies (*abscissa*), the dips in the variance of the subthreshold input became less pronounced, as reflected by the increasingly smaller fraction of the beat cycle during which the variance lies below the spontaneous value (*ordinate*). (B) A similar reduction of the fraction of below-spontaneous values of the variance was observed in the response to 70-dB-SPL monaural stimulation. The frequency dependence is in agreement with the assumption that the dips in the subthreshold potential are caused by the periodic absence of spikes from the monaural inputs; silent intervals will become less prominent for high stimulus frequencies, for which the phase-locking of the monaural inputs is known to degrade. CF: 680 Hz.

Movie S1 (related to Figure 5). Linear summation of monaural contributions predicts binaural subthreshold input.

The monaural contributions to the 400-Hz binaural beat stimulus were extracted by averaging over the respective stimulus periods of the two ears (cf. Figure 3C). The lower trace of the movie shows the ongoing ipsilateral (blue) and contralateral (red) inputs resulting from concatenating these single-cycle segments. The 4-Hz interaural frequency difference causes a running phase disparity between the inputs from the two ears. The upper trace shows the prediction (green) of the binaural subthreshold input obtained by the linear summation of the monaural waveforms. Its waxing and waning reflects the “beating” of the monaural inputs due to their slightly different periodicities. Along with the prediction the actual binaural subthreshold input, averaged over the 250-ms beat cycle, is shown (grey). The good match between the linear prediction and the measured response reveals that the response to the binaural beat can be well described as a beating between monaural inputs. The movie, which is best played in loop mode, provides a dynamic alternative to the static, two-dimensional mode of display of Figure 5A.

Supplemental Experimental Procedures

Data acquisition

Data were sampled at an interval of 10 μ s with a 16-bit A/D converter (Digidata 1322A). Potentials were filtered at 10 kHz low-pass (four-pole Bessel filter). In 2 recordings data were additionally high-pass filtered at 1 Hz (one-pole Bessel filter). To enable online analysis during recordings, the recorded potentials were filtered (0.1 Hz to 30 kHz) and amplified (gain 10x) with a PARC 113 pre-amp. Spike and very large EPSPs were sorted based on amplitude of either a positive or negative peak with a discrimination system controller (BAK electronics). Online analysis was done in parallel with the MultiClamp recordings and revealed auditory tuning of cells to which the experiments could be specified.

Whole cell in vivo recordings

For whole-cell recordings, pipettes were advanced with positive pressure up to 10 μ m until resistance approximately doubled, after which pressure was released, and suction was applied to obtain the cell-attached configuration. The whole-cell configuration was established by suction pulses. Whole-cell recordings were compensated for an estimated junction potential of -15 mV. Based on the fact that we only recorded from the somatic layer, as judged from the field potentials, most likely almost all juxtacellular and whole-cell recordings were somatic, although we cannot exclude that some recordings were made from proximal dendrites.

Analysis of in vivo recordings

The recordings were divided into 6500-ms or 9500-ms responses to single binaural-beat (or monaural tone) stimuli, each including a 500-ms spontaneous baseline immediately preceding stimulus onset. In order for a recorded response to be accepted for analysis, the RMS value of the 500-ms spontaneous portion had to exceed 0.1 mV and vary by no more than 25% across the individual responses by the same neuron. Recorded waveforms were filtered by a notch filter (50-Hz centre frequency; Q = 40, MATLAB function iirnotch) to remove any interference from AC power lines. Low-frequency artifacts caused by breathing and heart

beat were attenuated by subtracting a low-pass filtered version of the waveform, which was obtained by downsampling to a sample interval of 250 ms, followed by upsampling to the original sample rate. This method of high-pass filtering ensures that frequency components at 4 Hz (the binaural beat frequency) and above are unaffected. Any effects on the cycle averages used in the subsequent analysis are thus excluded, while still reducing the spurious low-frequency variance. To avoid transient effects, the first 500 ms following stimulus onset (i.e., the first two beat cycles) were excluded from analysis.

The eAPs were identified based on the steep (~ 10 V/s) downward slopes following their peak. The time derivative of the recorded waveform was computed by convolution with a 50- μ s-wide Hann window followed by subtracting adjacent samples. The local minima (within a moving 750- μ s-long time window) were determined, and the resulting histograms (Figure 1D) were visually inspected. Only those recording showing a clear bimodal distribution of local slope minima were accepted for further analysis (19 out of 25 cells). Spike arrival times were obtained by determining the temporal position of the peaks immediately preceding the steep downward slope. The timing of the EPSPs was obtained by either determining the time of the peak immediately preceding the eAP peak, or, when no such peak was present in the 500 μ s preceding the eAP, by finding the local minimum of the second time derivative (“inflection point”) in that interval (Lorteije et al. 2009).

Cycle-averages (Figures 3C, D; 4A; 5A, C, D, F) were obtained by dividing the waveform into $1/f$ -long segments (f is the stimulus frequency tested, i.e., the ipsilateral or contralateral stimulus frequency or their 4-Hz difference) and computing the average across the corresponding time points of the segments. Any samples lying within 750 μ s from an eAPs peak (typically comprising a small percentage of the total analysis window: median 0.15; inter-quartile range 0.9%) were excluded from this averaging process. Across-cycle variance (Figure 5E) was determined by the same procedure.

Vector strength of eAPs was determined as described in (Goldberg and Brown 1969); its significance was evaluated with the Rayleigh test (Mardia and Jupp 2000); $p < 0.001$). All reported correlations are Pearson’s linear correlation values.

Characteristic frequency (CF) was estimated by determining peak spike rate in the responses to the binaural beats when varying carrier frequency, using the lowest SPL tested that gave a

significant response (30 dB SPL, $n = 4$; 40 dB SPL, $n = 3$; 50 dB SPL, $n = 6$; 60 dB SPL, $n = 4$; 70 dB SPL, $n = 2$). The location of the peak on the rate-frequency graph was estimated by fitting a parabola to the five data points nearest to the peak value. Low tail responses to high-CF neurons can be immediately excluded for all cells (17 of 19) that responded to 500 Hz tones at 60 dB SPL or less (Figure 5 of (Ohlemiller and Echterler 1990)).

Instantaneous firing rate (Figure 7C) was computed as follows. The beat-cycle average $V_B(t)$ subthreshold input potential was determined as described above, with t the time from the start of the 250-ms beat cycle. A histogram $H_0(V_B)$ of all $V_B(t)$ values was compiled using 15 equally spaced bins. The spike arrival times τ_k from the same recording were determined as described above, and the values of the cycle-averaged subthreshold potential at the corresponding times within the beat cycle, $V_B(\tau_k)$, were used to compile a histogram $H_{AP}(V_B)$, using the same binning. The conditional instantaneous firing rate $p(V)$ was computed by

$$p(V)dt = H_{AP}(V)/H_0(V),$$

with dt the sample period of $V_B(t)$. The same relation was used to predict monaural spike rates from binaural instantaneous firing rates, this time combining the $p(V)$ derived from the binaural recording with the amplitude statistics $H_0(V)$ derived from the monaural recording.

In addition to the general acceptance criteria for analysis given above, the compilation of population statistics demanded specific criteria determined by the availability of recordings suited for a particular analysis. The two cases are listed below.

1. *Binaural tuning required.* The criterion for this set of recordings was a Rayleigh test significance ($p < 0.001$) of APs to the 4-Hz beat frequency, yielding $n = 296$ beat responses, 22 cells, 3 of which were recorded in whole-cell configuration.

2. *Monaural recordings from both ears required.* This criterion yielded $n = 550$ binaural beat responses, 18 cells, 2 of which were recorded in whole-cell configuration.

Results are reported as mean \pm standard deviation.

Chapter 3

Predicting binaural responses from monaural responses in the gerbil medial superior olive

Andrius Plauška, J. Gerard G. Borst, Marcel van der Heijden

J Neurophysiol. 2016; 115:2950-2963

Abstract

Accurate sound source localization of low-frequency sounds in the horizontal plane depends critically on the comparison of arrival times at both ears. A specialized brainstem circuit containing the principal neurons of the medial superior olive (MSO) is dedicated to this comparison. MSO neurons are innervated by segregated inputs from both ears. The coincident arrival of excitatory inputs from both ears is thought to trigger action potentials, with differences in internal delays creating a unique sensitivity to interaural time differences (ITDs) for each cell. How the inputs from both ears are integrated by the MSO neurons is still debated. Using juxtacellular recordings, we tested to what extent MSO neurons from anesthetized Mongolian gerbils function as simple crosscorrelators of their bilateral inputs. From the measured subthreshold responses to monaural wideband stimuli we predicted the rate-ITD functions obtained from the same MSO neuron, which have a damped oscillatory shape. The rate of the oscillations and the position of the peaks and troughs were accurately predicted. The amplitude ratio between dominant and secondary peaks of the rate-ITD function, captured in the width of its envelope, was not always exactly reproduced. This minor imperfection pointed to the methodological limitation of using a linear representation of the monaural inputs, which disregards any temporal sharpening occurring in the cochlear nucleus. The successful prediction of the major aspects of rate-ITD curves support a simple scheme in which the ITD sensitivity of MSO neurons is realized by the coincidence detection of excitatory monaural inputs.

New and Noteworthy

We recorded from principal neurons of the gerbil Medial Superior Olive (MSO), measuring how firing rate varied with interaural correlation and Interaural Time Delay (ITD). We also recorded subthreshold monaural responses to wideband tone complexes. Wideband rate-ITD functions were predicted in considerable detail by cross-correlating the monaural responses, suggesting that to explain ITD sensitivity in MSO one need not postulate a major computational role of binaural interaction beyond a simple linear addition of monaural inputs.

Introduction

Accurate sound source localization can be essential for survival. The sub-millisecond differences in arrival time between the ears, the interaural time differences (ITDs), are a dominant cue for localization in the horizontal plane, especially at low frequencies. The medial superior olive (MSO) plays a key role in this process (Brown and May 2005). Two bilateral excitatory and two bilateral inhibitory inputs make up the core afferent circuitry of the MSO. The role of inhibitory inputs, one from the ipsilateral lateral nucleus of the trapezoid body, and the other from the contralateral medial nucleus of the trapezoid body neurons, is still debated (Franken et al. 2015; Myoga et al. 2014; Pecka et al. 2008; Roberts et al. 2013; van der Heijden et al. 2013). The excitatory inputs originate from spherical bushy cells in the anteroventral cochlear nuclei on both sides (Thompson and Schofield 2000) and their role in sound localization is understood in greater detail.

It is commonly accepted that MSO neurons act as coincidence detectors of their monaural excitatory inputs (Goldberg and Brown 1969; Yin and Chan 1990). The internal delays of both excitatory inputs give each MSO cell its own "best ITD", the interaural delay in the stimulus leading to maximal excitation. Although the origin of internal delays is debated, the general features of ITD selectivity in the MSO are in agreement with coincidence detection as proposed by Jeffress (1948), reviewed in Vonderschen and Wagner (2014). The integration of the monaural inputs can be alternatively described by schemes of coincidence detection or crosscorrelation, or by linear summation combined with a nonlinear expansive excitability (Colburn et al. 1990; van der Heijden et al. 2013). Such schemes, which produce similar predictions of ITD selectivity, correctly predict the best ITD of MSO neurons from the timing of the excitatory inputs as assessed from the neuron's response to monaural stimuli (Goldberg and Brown 1969; van der Heijden et al. 2013; Yin and Chan 1990). Several studies, however, suggest that such crosscorrelation schemes may fail to reproduce other aspects of binaural responses (Batra et al. 1997a; Batra and Yin 2004; Franken et al. 2015). All of these earlier tests were performed using tonal stimuli, and although they may provide insight into the underlying mechanisms of ITD tuning, the relevance of tonal responses to everyday sound localization is limited.

In this study we explore whether the prediction of binaural MSO responses from their monaural inputs can be generalized beyond predicting best ITDs for tones. The idea is as follows. If a simple crosscorrelation between monaural inputs suffices to accurately predict binaural responses of a more general and common type, there is little need to postulate a major computational role for somatic integration in MSO cells beyond the linear summation of ipsi- and contralateral inputs found by van der Heijden et al. (2013). Such a major computational role has been suggested in previous studies (Agmon-Snir et al. 1998; Brand et al. 2002; Franken et al. 2015; Jercog et al. 2010; Pecka et al. 2008; van der Heijden et al. 2013; Zhou et al. 2005). If, on the other hand, systematic deviations between crosscorrelation predictions and actual ITD tuning are observed, these may point to the nature and relevance of possible nonlinear interactions between ipsi- and contralateral inputs. We characterized the ITD selectivity of MSO cells by measuring firing rate in response to wideband stimuli with varying ITD. The resulting wideband rate-ITD functions in MSO have a damped oscillatory shape (Yin and Chan 1990), a pattern resembling the autocorrelation function of bandlimited waveforms. Yin and Chan (1990) showed a reasonable similarity between an example wideband rate-ITD curve and the cross-correlogram of spike trains obtained with monaural stimulation, thus providing a first, basic test of the predictability of wideband ITD tuning from monaural data. Here we report a systematic and extensive test, based on subthreshold monaural responses, which allowed us to use moderate sound intensities and to overcome the limitations of using monaurally evoked action potentials as a proxy for monaural inputs (Batra and Yin 2004). Using the measured responses to monaural wideband stimuli as inputs to a crosscorrelation model, we predicted the wideband rate-ITD function, which we then compared to the wideband rate-ITD functions recorded from the same MSO neurons.

Materials and Methods

Animal Procedures

All experiments were conducted in accordance with the European Communities Council Directive (86/609/EEC) and approved by the institutional animal ethics committee. Young-adult female Mongolian gerbils (average body mass 60 g) were anesthetized intraperitoneally

with a ketamine-xylazine solution (114 and 17 mg/kg body weight, respectively). Reflexive state was monitored throughout the experiment by performing the hind-paw pinch-reflex test and was maintained by regular administration of ketamine-xylazine solution (one third of induction dose). Body temperature was maintained at 37 °C using an electrical heating pad.

The head was fixed using a metal head plate that was glued to an exposed rostral-dorsal part of the skull. Both pinnae were removed, providing access to the ear canal for placing the tubes for delivering sound stimuli. The animal was fixed in a supine position. The skin, connective tissue, salivary glands and lymph nodes above the trachea were surgically removed, followed by a tracheotomy after which the animal kept breathing independently. Both bullae were exposed by removing the overlying muscles and making openings in the bone using a scalpel and forceps. The opening of both bullae equalizes the effect on low-frequency middle transfer (Ravicz et al. 1992). A 0.7 to 0.9 mm diameter craniotomy was made on the right side using a drill. The craniotomy was located approximately 2 mm rostrally from the stapedia artery and in the middle between the cochlea and medial wall of the skull. When needed the craniotomy was extended by drilling in the vicinity of the first craniotomy. The meninges were left intact. The angle of the electrode insertion point could be changed using a fixed-pivotal-point, custom-built positioning device on which the animal rested throughout the experiment.

In Vivo Juxtacellular Recordings

Recordings were made with thick-walled borosilicate glass microelectrodes having a resistance of 4 – 7 M Ω when filled with recording solution. Pipettes were filled with solution which contained K-gluconate 138, KCl 8, Na₂-phosphocreatine 10, MgATP 4, Na₂GTP 0.3, EGTA 0.5, HEPES 10 mM (pH 7.2 with KOH). Less than 10% of the cells were recorded from using extracellular solution which contained NaCl 135, KCl 5.4, MgCl₂ 1, CaCl₂ 1.8, HEPES 5 mM (pH 7.2 with NaOH). Some of the electrodes had biocytin (0.1%) added to the solution. No specific differences could be found between the responses of cells recorded with either solution, so we pooled them for the analysis.

Pipettes had high positive pressure (~100 mbar) during brain surface penetration. Immediately after successful penetration of the brain surface, the pressure was lowered to 20-30 mbar and we waited for a few minutes before making a recording to minimize the

impact of brain tissue movements relative to the electrode. The location of the MSO somatic layer was identified based on the local field potential (Biedenbach and Freeman 1964; Clark and Dunlop 1968; Galambos et al. 1959; Mc Laughlin et al. 2010), as described previously (van der Heijden et al. 2013). Field potentials were evoked using monaural click stimuli (2-ms duration) presented alternately to either ear. Once the somatic layer was reached, the pipette was advanced slowly and its resistance was monitored closely. Contacting a neuron resulted in a gradual increase in resistance, after which we released positive pressure. The electrode was then forwarded up to 10 μm to increase and stabilize seal resistance. Typically the resistance reached a value of 20-40 M Ω . All the recordings were done in current-clamp mode, while regularly monitoring the seal resistance. In case of evident changes in cell response the recordings were stopped.

Data were acquired using a Multiclamp 700B (Molecular Devices, Foster City, CA) amplifier using custom software written in MATLAB (The MathWorks, Natick, MA, USA).

Stimuli

Auditory stimuli were generated using custom software written in MATLAB and realized through a 24-bit D/A-channel (RX6; Tucker Davis Technologies (TDT), Alachua, FL, USA; 111.6 kHz), programmable attenuator (PA5; TDT) and an amplifier (SA1; TDT). Stimuli were delivered to the ear canals in a close-field fashion through Shure speakers (frequency range 22 Hz to 17.5 kHz) and a pair of small (\sim 11 cm length) tubes.

Three types of stimuli were used for this study: irregular tone complexes presented monaurally, irregular tone complexes presented binaurally (noise delay stimulus) and broadband, Gaussian noise with varied interaural correlation. The frequency range of all stimulus types was 50-3000 Hz; the tone complexes consisted of 30 components. The sound pressure level (SPL) range for monaural and ITD delay stimuli was 20–40 dB above hearing threshold. The total SPL of the noise was 15 dB higher than the SPL per tone of the tone complex stimuli in order to achieve the same effective SPL level. Hearing threshold was defined as the lowest SPL of monaural click stimulus to which neurophonics could be evoked.

Monaurally presented stimuli were frozen within recordings from one cell; their 22-s duration and 2950-Hz bandwidth ensure that the waveforms were representative of the statistical

ensemble of these wideband stimuli. The waveforms for all the other stimuli were recomputed for a different instance of their presentation, but were kept the same within one instance for all different stimulus conditions and repetitions.

Monaural stimulation

Monaural responses were evoked using ‘zwuis’, an irregular tone complex stimulus (van der Heijden and Joris 2003; 2006) with the following properties. Thirty distinct pure tone components with an irregular spacing in the frequency range of 50-3000 Hz were chosen in such a way that no second-order or third-order distortion products present in the response could have a frequency of any of the stimulus components. Duration of presentation was ~22 seconds (including 1 s each for pre- and post-stimulus baselines). All stimulus components had the same amplitude and a random phase (van der Heijden and Joris 2006; Versteegh and van der Heijden 2012).

Delay functions

Rate-ITD functions were obtained by presenting the zwuis stimulus binaurally with a systematically varying time delay between the ears. ITD was varied over a range chosen after evaluation of the frequency response of the neuron: for neurons preferring lower frequencies wider ranges were applied than for the higher frequency ones. Stimuli were delayed in equal steps in a random order with 20 repetitions for each condition. Positive ITDs corresponded to the signal leading at the contralateral ear. The duration of a single presentation was 300 ms, followed by a 100 ms silence period, resulting in a total stimulus duration of 170 s.

For a large subset of cells an inverted rate-ITD function was obtained by inverting the polarity of the stimulus in the contralateral ear. To distinguish between the responses to the two stimuli we call the former “normal rate-ITD function” and the latter, “inverted rate-ITD function”.

Difcor

To emphasize the fine structure component of the neuron's ITD sensitivity we plotted difcor function (Joris 2003; Joris et al. 2006) when possible. This was achieved by subtracting responses of inverted rate-ITD function from normal rate-ITD function.

Rate vs interaural correlation functions

Rate vs interaural correlation functions were obtained from a binaurally presented broadband Gaussian noise (50 – 3000 Hz bandwidth, 300 ms burst for each condition, 20 different correlation values (conditions) and 20 repetitions, total duration of stimulus - 162 s). For a detailed explanation of the method see Louage et al. (2006). Briefly, a single noise token is generated by mixing two independently drawn Gaussian broadband tokens. The same stimulus is always presented at the contralateral side. For the ipsilateral ear for every new condition a new noise token is generated by mixing the same two independent tokens, but with varying weighting coefficients. The difference between the coefficients of contralateral and ipsilateral presentation composites determines the normalized correlation between the stimuli and covers the full correlation range from $\rho = -1$ (anti-correlated) to $\rho = 1$ (correlated). The correlation stimulus was presented at the best ITD.

Admission/selection criteria for the data

Criteria for a successful MSO juxtacellular recording

Field-potential reversal is an established method for detecting the MSO somatic layer (Biedenbach and Freeman 1964; Galambos et al. 1959; Mc Laughlin et al. 2010). In all our experiments MSO was approached from the ventrolateral to the dorsomedial direction. During this approach the polarities of the field potentials indicated the location of electrode as follows: 1) the ipsilateral potential was negative, indicating a local sink, and the contralateral potential was positive before reaching the somatic layer; 2) the ipsilateral potential reversed as the electrode entered the somatic layer; 3) while in the somatic layer the field potentials had the same polarity; 4) leaving the somatic layer resulted in the reversal of the contralateral field potential. We considered a cell an MSO cell if the field potentials

were of the same polarity just before the electrode made contact with the cell. In the same animal it was always verified that further penetration led to a second reversal. For thirty-three experiments, the electrode penetration track was marked using either biocytin, Alcian blue or Dil at the end of the experiment. All sixteen penetrations that could be recovered histologically indicated that electrodes penetrated the somatic layer of the MSO.

Juxtacellular recordings have been shown to be suitable for resolving both sub- and suprathreshold activity of a single MSO unit (van der Heijden et al. 2013). The resistance between the electrode and the cell varied across different cells. Only those recordings for which the initial seal resistance was between 20 and 70 M Ω were accepted as valid juxtacellular recordings.

Baseline-based criteria for recording stability

All the stimuli included either one- or two-second duration pre- and post-stimulus baselines. We used these baseline periods to evaluate the stability of the recordings over time (Fig. 1). Stable recordings are indispensable when attempting to predict one set of responses from another. A power spectrum was generated for each baseline period of all the recordings. Most neural events are in the millisecond range, resulting in a peak of the power spectrum at around 1 kHz (Fig. 1BI). For all the recordings from a cell we determined the baseline power spectra amplitudes at 1 kHz and then plotted them sequentially (from the first recording till the last one). This allowed us to observe whether there were any changes in the recorded spontaneous activity of the cell which in turn represents the recording stability (Fig. 1C). Recordings were considered stable if during the time course the power spectra amplitudes did not deviate more than 5 dB from an arbitrary threshold, which was different for each cell. The threshold was chosen in such a way that the 5 dB window would include the largest number of sequential recordings from a cell. These recordings were used for further analysis. Field recordings, unsuccessful juxtacellular configuration with a cell and unknown objects that increased electrode resistance, but showed no signs of electrophysiological activity, exhibited no peaks in their baseline power spectra (Fig. 1BII, 1BIII). These cases were excluded from further analysis.

Criteria for data quality

To confirm that recordings were made from ITD-sensitive cells we performed a single-sided ANOVA test (function *anova1* in MATLAB) for the rate-ITD functions. $P < 0.01$ was considered statistically significant. Those cells for which the responses had significant differences in their means of different conditions (time delays), were considered binaural and the data were used for further analysis. These criteria were applied only to difcor data (see below).

To eliminate possible stimulus-independent changes in a cell's response to noise delay stimuli we set an additional criterion for the difcor functions. After fitting a difcor with a Gabor function (see below) we estimated the ratio between the offset and the amplitude of the envelope and set the threshold for this ratio at 0.2. This binaural consistency criterion selects against cases in which the cell response changed between direct and inverted noise delay stimuli presentations. Some of the cells showed highly reduced ITD sensitivity or changes in the spontaneous firing rates, which will affect difcor. Only difcors for which these ratios were < 0.2 were used for further analysis.

For rate vs interaural correlation functions we used a 'dynamic range' criterion. The dynamic range of a correlation function was defined as the difference between the average spike rates at the three highest stimulus correlation values and the average spike rate for the five correlation values centered around ρ equal to zero. The rationale behind this criterion was to discern the cells with poor response to the stimulus and exclude them from predictions that were based on the rate vs interaural correlation functions without excluding the cells which could be envelope-sensitive. After comparing the variances accounted for by the power function fit on the data, we determined a reasonable dynamic range threshold to be five spikes/s. Rate vs interaural correlation functions with a dynamic range below five spikes/s were considered unreliable for prediction modelling and were not considered further.

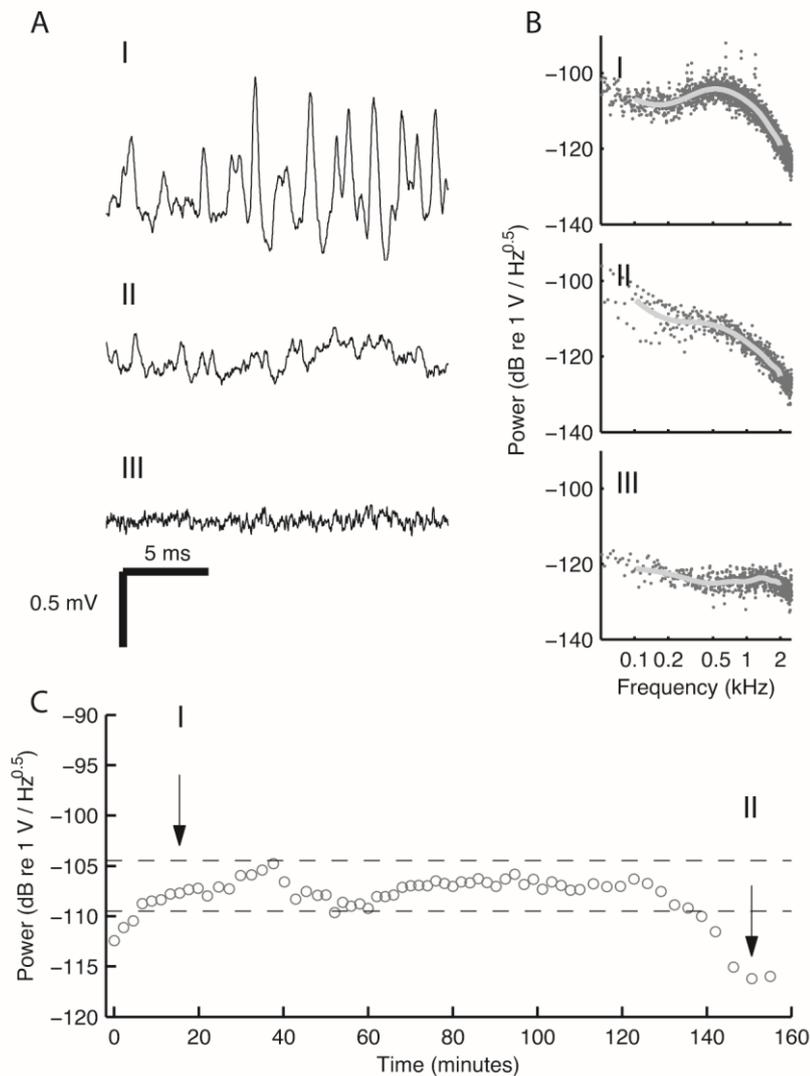


Figure 1. Evaluating recording stability using pre-stimulus baselines.

(A) Three examples of recorded pre-stimulus baselines illustrating the juxtacellular configuration (I), reduced contact (II) and lack of direct contact (field recording) (III). (B) Power spectra of the waveforms shown in A. Spontaneous MSO activity comes about as a peak at around 0.5 kHz on the power spectrum (BI). This peak is not present when either the contact with the cell is poor (BII) or the electrode is not in contact with a cell (BIII). Grey line shows the fit on the power spectrum between 0.1 and 2 kHz. (C) Changes in the power spectrum value at 1 kHz in the time course of recordings from one MSO cell. The total duration of the electrode contact with the cell was 155 minutes. Horizontal dashed lines indicate a 5 dB window used as a criteria of recording selection. Arrows with numbers indicate which baselines from the time course are shown in AI and AII. The measured resistances for configurations I, II and III were 36, 21 and 7 M Ω respectively.

Data analysis

Event detection

Action potential detection was done offline based on the statistics of the negative peak values in the derivative of recorded waveforms (van der Heijden et al. 2013). Only if the histogram of local minima exhibited a clear bimodality, the separation between the modes defined a threshold derivative which could be used to identify the action potentials; if the distribution did not show a clear bimodality, the recording was rejected from further analysis. The characterization of the monaural inputs (which were used to make the binaural predictions) is based on the long-term spectra of subthreshold responses. We therefore also restricted the analysis of binaural responses to the sustained part of the responses, cutting out the first 10 ms of the waveforms obtained with binaural stimulus presentations.

Fitting functions

Binaural data were fitted using two functions as described below. Such fits allowed us to compare measured data with predictions. Every fit was optimized using the least squares method (function *lsqcurvefit* in MATLAB) and the variance accounted for was estimated for each instance. A one-dimensional Gabor function (here simply referred to as Gabor function) has been used in fitting binaural models (Leibold and van Hemmen 2005) and data (Franken et al. 2014) before.

Linear representation of monaural inputs

Response amplitude and phase were determined by Fourier analysis of the recorded responses to monaural zwuis tone complexes, leading to a representation of the effective monaural input to the neuron that is a linearly filtered version of the stimulus. Statistical significance of each spectral component was tested by a Rayleigh test ($p < 0.001$) applied to the phase values obtained by segmenting the response waveform into 10 equally long, non-overlapping segments (Versteegh and van der Heijden 2012). The significant frequency components were plotted as a relative signal gain over the stimulation intensity. The phases

of significant components were presented versus the respective stimulus components. Normalized cross-correlation functions were obtained from these monaural transfer functions by computing the normalized cross-spectrum and converting it to the time domain by a Fourier transform. More information about stimulus generation, analysis and interpretation of results is provided in van der Heijden and Joris (2006); Versteegh and van der Heijden (2012). In order to check the effects of monaurally evoked action potentials on the predictions we compared predictions from monaural responses with truncated action potentials with predictions from responses with action potentials included. The presence of the relatively small numbers of monaurally evoked action potentials, which contributed on average at most 0.5 % of the whole response duration, did not affect the cross-spectrum and predictions. The predictions presented in this study were therefore based on the waveforms with the spikes present.

Gabor fit

Difcor functions were fitted with a Gabor function

$$G(\tau) = A + B e^{-0.5 \left(\frac{\tau - \tau_0}{w} \right)^2} \cos(2\pi(f\tau - \varphi_0)), \quad (1)$$

where τ is time (ms), A the offset (spikes/s), B the amplitude (spikes/s), τ_0 the peak time of the envelope (ms), w the width of the envelope (ms), f the frequency (kHz), and φ_0 the start phase of the cosine (cycle).

Rate-ITD functions were fit with a modified Gabor function with an additional power parameter which allowed to reshape the Gabor function before fitting it on the data. The modified function (“pGabor”) can be written as

$$G_p(\tau) = A + B g(p, e^{-0.5 \left(\frac{\tau - \tau_0}{w} \right)^2} \cos 2\pi(f\tau - \varphi_0)) \quad (2)$$

where $g(p, x) = \frac{(x+1)^p}{2}$

and p is the additional parameter, the power of the modifying function. This modification results in an asymmetric envelope of the fit, which more faithfully represents the measured rate-ITD functions with their “rectified” appearance.

Power function fit

Rate vs interaural correlation functions were fitted using a generalized version of the power function (Shackleton et al. 2005), where the new addition is the ‘parabola’ component with its weighting factor C . This function can also be used to fit non-monotonic rate vs interaural correlation functions that have an envelope or “polarity-tolerant” component (Joris 2003).

$$y(x) = A + B \left(\frac{1+x}{2} \right)^p + Cx^2, \quad (3)$$

with A , B , C , and p corresponding to the offset, amplitude, weighting parabola coefficient and power parameters, respectively.

Results

We presented multitone stimuli to juxtacellularly recorded MSO cells in anesthetized Mongolian gerbils to test to what extent binaural responses could be predicted from monaural responses. The stimuli were presented at 20-40 dB SPL above hearing threshold to ensure that MSO neurons were most sensitive to frequencies close to their characteristic frequency (CF). Best frequencies (BFs) were on average 0.95 ± 0.32 kHz (mean \pm standard deviation; range 0.27 - 1.35 kHz; $N = 60$). All MSO neurons in this study are therefore ‘low-frequency’ neurons. The low intensity of stimuli resulted in subthreshold-dominated monaural responses and varying response types upon binaural stimulation. Binaural responses of the neurons were evaluated only on the basis of action potentials. We looked into the extent to which an MSO neuron functions as a cross-correlator of its monaural inputs by cross-correlating its monaural responses and comparing predicted with measured binaural responses. As further detailed in the Methods, we used two different methods, difcor and rate-ITD (rITD) functions, to predict binaural ITD sensitivity from monaural responses, depending on the available data. Difcor predictions look at the neuron’s sensitivity to the fine

structure of the stimulus. Rate-ITD functions show a neuron's general ITD sensitivity regardless of which frequencies the neuron responds to. First we present the method of difcor prediction and its comparison with the data and after that the data and predictions of rITD functions.

Linearized rate-ITD curves

Wideband rITD functions of binaural neurons generally show a mix of sensitivity to stimulus fine structure and envelope, and the two contributions can be largely separated by subtracting and summing the curves obtained with different interaural polarities (Joris 2003). Denoting the regular and polarity inverted versions of the rITD curve by rITD+ and rITD−, respectively, their difference rITD_{DIFF} isolates the fine structure component of ITD sensitivity. It is a linearized version of the rITD function in that it asymptotes to zero for large positive and negative ITD values and has a more symmetrical envelope than the raw rITD+ and rITD− (Joris 2003; Joris et al. 2006). In correlograms of monaural responses, the difference correlogram ("difcor") closely resembles the cross-correlation function of the effective (filtered) waveforms to the two ears (Mc Laughlin et al. 2014), and this resemblance motivated the current approach. We derived the transfer characteristics of the monaural inputs to each MSO neuron from its subthreshold response to monaural tone complexes, computed the cross-correlation function and compared it to the measured rITD_{DIFF} data (see Methods).

From a total of 87 cells, both normal and inverted rITD functions were obtained, allowing an analysis of the fine-structure, oscillatory component of wideband ITD tuning using rITD_{DIFF} curves. Out of these 87 cells, 62 passed the statistical criteria for binaurality (ANOVA; $p < 0.01$). Application of the binaural consistency criterion (See Methods, Criteria for data quality) led to a collection of 47 cells from 31 animals.

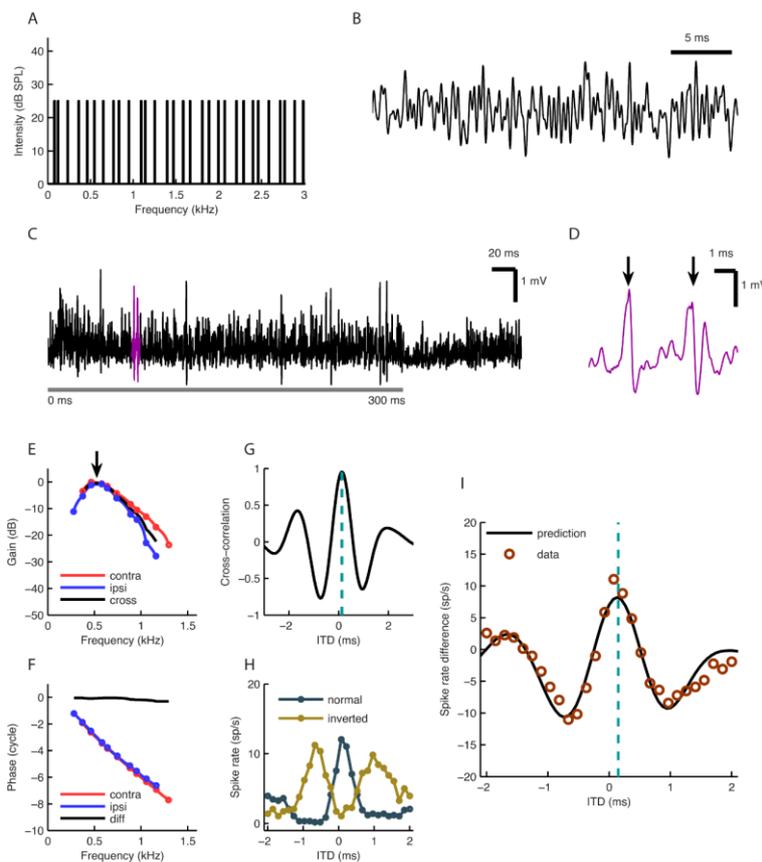


Figure 2. Predicting difcor of an MSO neuron.

(A) Power spectrum of irregular tone complex stimulus ('zweis'). The 30 components have equal magnitudes and span 3 kHz bandwidth in irregular intervals. (B) An excerpt of the zweis stimulus waveform - the result of mixing the 30 frequency components shown in A. (C) A typical MSO neuron response, recorded juxtacellularly, to a noise delay stimulus. Stimulus starts at 0 ms time point (with a 3 ms ramp) and lasts for 300 ms (grey bar). The highlighted segment is shown separately

in D. (D) A segment from a juxta-cellular recording of an MSO neuron, shown in C. The two largest events are action potentials (arrows) that were identified offline by thresholding the derivative of the response. (E) MSO neuron's frequency response to monaural zweis (tone complex) stimulation. Blue and red colors represent ipsi- and contralateral ear responses, respectively. The circles are data points, the lines interpolated values. Best frequencies (BF) of contralateral (0.54 kHz) and ipsilateral ear (0.51 kHz) are indicated with an arrow. The response magnitudes are estimated with respect to the magnitude of the stimulus SPL (30 dB for this cell) and normalized to the dominating monaural response. Black line represents the cross spectrum of monaural responses (see Results). (F) Phases to which the neuron 'locks' monaurally for each frequency component. Color coding is the same as in E. The circles are data points, the lines interpolated values. Black line shows the phase difference (contra minus ipsi). (G) Cross-correlation function ("crosscorr") of the neuron, showing how ipsilateral and contralateral ear transfer functions correlate linearly. The cyan bar indicates the peak of the function – the expected best ITD (0.13 ms). (H) Rate-ITD functions. The stimuli presented to both ears had either the same ('normal', dark blue line) or opposite polarity ('inverted', dark yellow line). (I) Comparison of measured and predicted difcor functions. The circles represent a difcor function which is obtained by subtracting the inverted rate-ITD function from the normal rate-ITD function. The prediction (solid black line) is estimated by scaling the crosscorr in ordinate direction (indirect conversion from cross-correlation to spike rates) to fit the difcor points by least squares. Note that such scaling does not alter the extrema locations of the prediction. Cyan bar indicates the expected best ITD (0.13 ms). The best ITD of the difcor was 0.12 ms. The data variance explained by the model was 91%. All data are from the same cell.

Responses to monaural stimulation can be used to predict fine structure sensitivity to ITDs

The basis of our model is the responses of MSO cells to monaurally presented tone complexes (zweis stimulus) presented at either 20, 30 or 40 dB SPL per tone. Fourier analysis of the raw waveforms, which were dominated by subthreshold EPSPs, yielded a representation of the monaural response characteristics of each ear in terms of linear transfer functions (see Methods, Linear representation of monaural inputs).

Figure 2A shows the power spectrum of the zweis stimulus with 30 equal amplitude components spaced irregularly within a 3 kHz bandwidth. A segment of the zweis waveform is shown in Figure 2B. An example of the response of a juxtacellularly recorded MSO neuron to a noise delay stimulus is depicted in Figure 2C. In this example the stimulus started at 0 ms time point with a 3 ms ramp and lasted for 300 ms (grey bar), which was followed by a 100 ms silent period. The highlighted segment of the recording is expanded and shown in Figure 2D. During stimulus presentation the neuron responded with action potentials (Fig. 2D, arrows). Figure 2E (circles) shows the frequency responses obtained from contralateral (red) and ipsilateral (blue) stimulation for a typical MSO cell. These responses provide a measure for the phase-locked response of the neuron to the 30 stimulus components. The peak response is the BF. For the neuron presented in Figure 2 the BF of the contralateral ear was 0.51 kHz and that of the ipsilateral ear 0.53 kHz. The depicted frequency responses can be seen as linear transfer functions from the ear canal to the subthreshold responses of the neuron.

The corresponding phase curves are shown in Figure 2F. The phase difference (contralateral minus ipsilateral) is plotted as a black solid line. For each monaural response a complex transfer function was calculated. A cross spectrum was then produced by multiplying the contralateral ear transfer function with a complex conjugate of the ipsilateral ear transfer function (Fig. 2E, solid black line). The inverse Fourier transform of the cross spectrum yielded the cross-correlation function or “crosscorr” (Fig. 2G). Such a function represents the normalized correlation between the two linearized monaural responses. The peak of the crosscorr is indicated by the cyan line (0.13 ms).

Normal and inverted rITD functions (Fig. 2H) were measured by varying the ITD of a wideband multitone stimulus. The stimulus of the normal rITD function was identical in the two ears

(apart from the ITD); the stimulus for the inverted rITD function was obtained by inverting the stimulus polarity in one ear. Subtraction of the inverted rITD function from the normal one resulted in a difcor function (Fig. 2I, circles). The difcor function reflects the dependence of ITD sensitivity on the fine structure, and has a tendency to linearize the dependence on interaural correlation (Joris 2003). Binaural responses were determined using only action potentials. Since there is no direct way to convert cross-correlation values into spike rate differences, we scaled the cross-correlation function by a single factor for optimal overlap with the measured difcor function in the least-squares sense (Fig. 2I, black line). Note that such rescaling does not affect the shape of the function or the location of its extrema. The cyan bar in Figure 2I indicates a close match to the best ITD of Figure 2G and to the measured best ITD of this cell, which was 0.12 ms.

Figure 3 shows monaural and binaural responses with the predictions for six additional MSO neurons (1-6). The color scheme and the methods are the same as in Figure 2. Monaural response magnitudes (A) and their phases (B) were used to produce a difcor prediction (D, solid line) as described above. The predictions were compared with the measured difcor (D, circles) by estimating the variance of the data accounted for by the model (percentages in D). For these six cells the variances accounted for ranged from 72% to 96%. Figure 3C shows the phase difference dependence on the frequency. The irregular shapes of many of these curves shows that the relation between the two inputs is not a simple delay, which would yield a straight line through the origin. Similar irregularities were previously found in the MSO of the gerbil (Fig. 10 of (Day and Semple 2011); Figs. S4 and S5 of (Day and Semple 2011); van der Heijden et al. (2013)) and in the MSO of the cat (Fig. 10C of Yin and Chan (1990); Figs. 1-4 of (Batra et al. 1997a)). Some of the pairs of magnitude-frequency curves in Fig. 2A show an interaural difference in frequency tuning, again showing that a simple delay is not sufficient to describe the interaural disparity of the inputs. A detailed analysis of asymmetric frequency tuning is beyond the scope of the present study.

For a more exhaustive comparison between predictions and data we fitted both to a standard Gabor function (Fig. 4A, solid black line) using least squares (see Methods). The Gabor function is a sinusoid multiplied by a Gaussian envelope. The binaurally relevant parameters of the Gabor function are the starting phase and oscillation frequency, and the temporal

position of the envelope peak and its width. The best ITD can be derived from these four parameters.

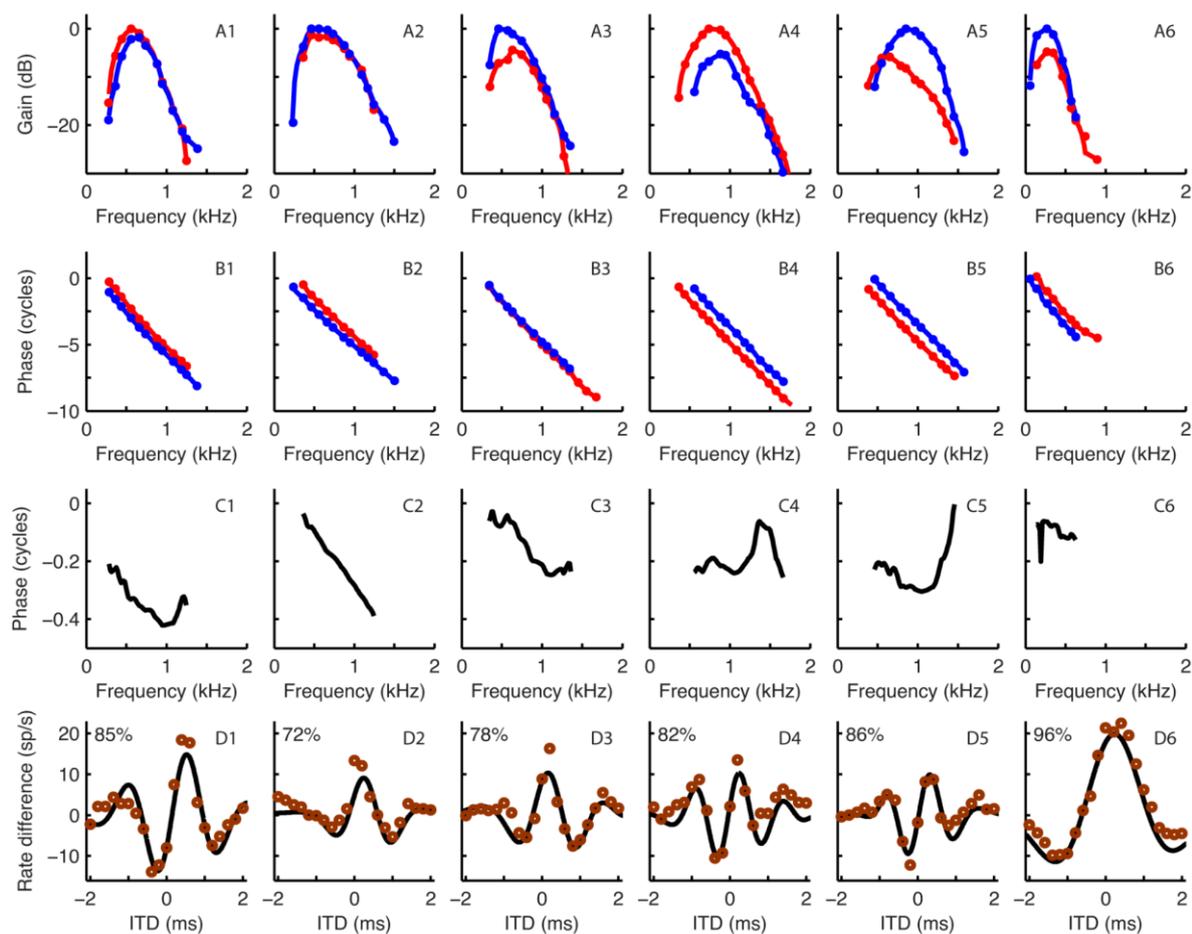


Figure 3. Examples of difcor predictions for six MSO cells.

Magnitudes (**A**) and phases (**B**) of responses to monaural stimulation together with phase differences (**C**) and measured difcors (circles, **D**) and their predictions (solid line, **D**) for six MSO neurons (1-6). Color coding and methods are the same as in Figure 2. For cells 2-6 the axes scaling is the same as for cell 1. The percentage values in **D** indicate the variance explained by the model on the measured data. Best frequencies for contralateral and ipsilateral ear and best ITDs for each cell respectively are as follows: Cell 1 – 0.54 and 0.61 kHz, 0.47 ms; Cell 2 – 0.56 and 0.54 kHz, 0.14 ms; Cell 3 – 0.65 and 0.53 kHz, 0.18 ms; Cell 4 - 0.80 and 0.84 kHz, 0.21 ms; Cell 5 – 0.62 and 0.89 kHz, 0.28 ms; Cell 6 – 0.29 and 0.24 kHz, 0.28 ms.

Figure 4B-F shows correlations between the parameters of fits on data and on predictions: best ITD (Fig. 4B), frequency (Fig. 4C), phase (Fig. 4D), envelope peak time (Fig. 4E) and envelope width (Fig. 4F). DC offset and envelope size are not informative for comparisons

between data and model and are thus not included in this study. Pearson's correlation coefficient is indicated in the top left of each plot. Grey lines are the identity lines. Note that the model is systematically predicting a somewhat wider envelope of the response than seen in the data (Fig. 4F). Apparently the central peak of a difcor is higher (relative to the secondary peaks) than what a cross-correlation of monaural inputs would produce (see Fig. 3D, 1-4). The other parameters show no obvious biases.

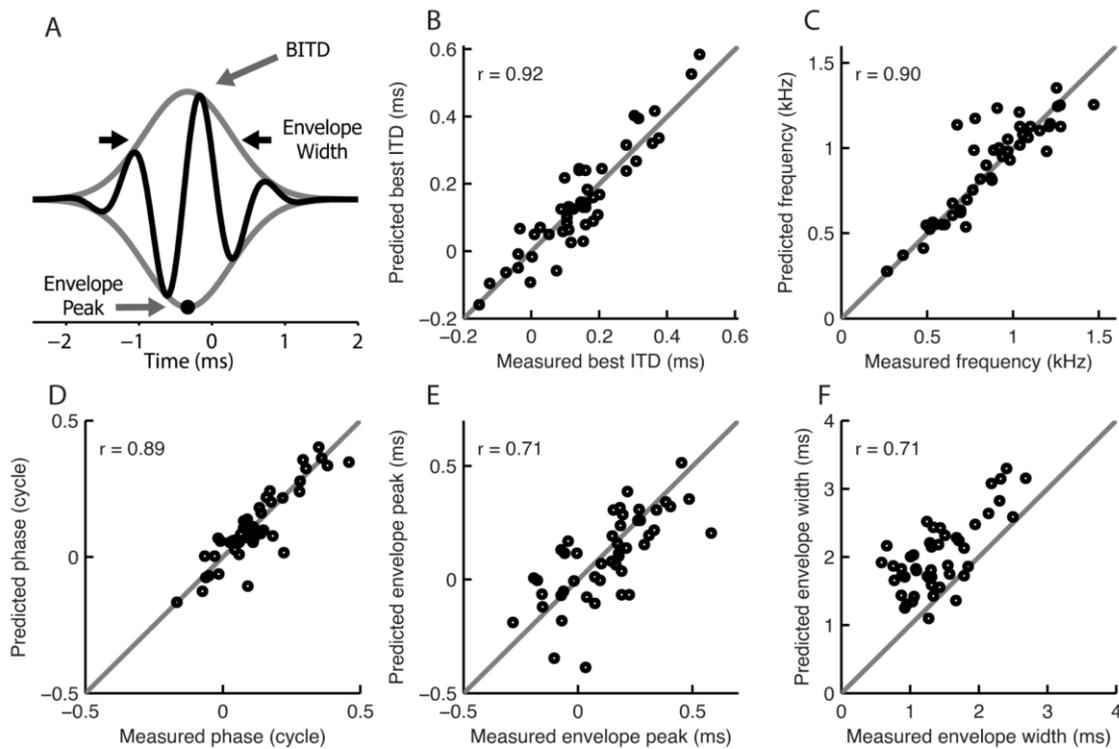


Figure 4. Comparing model and difcor fit parameters on a population level.

(A) Data and models were fit with a Gabor function (black line) which is defined by a phase offset, oscillation frequency, full width at half maximum of the envelope (black arrows) and envelope peak time (black dot). Best ITD is determined by the starting phase and oscillation frequency. Two additional parameters – offset and amplitude (distance between envelope peak and trough) are not indicated in the picture as they are not considered in this study. (B-F), Scatter plots of Gabor fit parameters (B – best ITD; C – frequency; D – starting phase; E – envelope peak time; F – full width at half maximum of the envelope) for difcors and their predictions of the MSO neurons for which difcor predictions were available ($N = 47$). Grey lines indicate identity. Values in each plot show the Pearson's correlation coefficient.

Rate-ITD function predictions

From a total of 84 cells for which responses to noise with varied interaural correlation were recorded, 32 (from 24 animals) passed the dynamic range criterion for rate vs interaural correlation (rIC) functions (see Methods, Criteria for data quality). Most cases that did not pass the criterion had low overall spike rates at the moderate intensities used, which did not allow a sufficiently accurate characterization of the rIC functions.

Responses to monaural stimulation can also be used to predict neuron's rate-ITD function

When possible, we recorded the response of MSO neurons to changing interaural correlation using a noise-mixing technique (see Methods). The resulting rIC function allowed us to predict the rITD function. An example is shown in Figure 5. The waveforms presented to the ipsi- and contralateral ear for this stimulus are illustrated in Figure 5A. The stimulus is anti-correlated when the polarities are opposing (left), uncorrelated when the polarities are unrelated in any systematic way (middle), and correlated when the polarities are identical (right). The small shift between waveforms in the correlated case was imposed intentionally for illustrational purposes to demonstrate that both stimuli are identical. Figure 5B-D show the same information as described for Figure 2, namely the monaural frequency (Fig. 5B) and phase (Fig. 5C) responses and the resulting cross-correlation function (Fig. 5D). Figure 5E depicts the measured relationship between the output rate of the neuron and the correlation between left and right ear stimuli; solid line is a power function fit (see Methods). For all the cells the increase in correlation from uncorrelated ($\rho = 0$) to perfectly correlated ($\rho = 1$) resulted in an increase in firing rate while 7 out of 32 cells showed higher output rates for anti-correlated stimuli ($\rho = -1$) than for uncorrelated ones ($\rho = 0$). Such non-monotonic behavior indicates a "polarity-tolerant" or envelope component in the response. An example of a "polarity-tolerant" cell is shown in Figure 6C4. For 28 rIC functions the power fit p parameter values were 3.1 ± 1.4 (mean \pm standard deviation; range 1.3 - 6.9).

Using the relation between the interaural correlation and output rate (Fig. 5E), we translated the crosscorr function into a predicted (expected) rITD function (Fig. 5F, solid line) and compared it with the measured one (Fig. 5F, circles). For conversion from interaural

correlation responses we used the power fit function. The expected best ITD for the cell was 0.15 ms, while the noise delay function peaked at 0.14 ms.

Four additional examples of rate-ITD function predictions are shown in Figure 6. The color scheme is the same as for Figure 5. For each cell (1-4), magnitudes (Fig. 6A) and phases (Fig. 6B) of monaural responses are plotted together with rIC functions (Fig. 6C). Figure 6D shows the measured (circles) and predicted (solid line) rate-ITD functions. Values indicate the percentage of the variance of the data accounted for by the prediction, ranging from 46% to 94%. For the assessment of the accuracy of the prediction we used a modified Gabor function as a fit. Since rITD functions, unlike difcors, do not have comparable sizes of peaks and troughs owing to the spike rates being only positive, we modified the Gabor function by adding a power parameter p , analogous to the power value in the functions used for fitting the rIC function data. A value $p = 1$ correspond to an "undistorted" Gabor function; increasingly higher values of p corresponds to an increasing degree of "rectification" by enhancing the peaks and reducing the troughs. We fitted both the data and predictions with this modified Gabor function. Some of the rITD functions had only one peak (or trough) across the measured ITD range, and fitting such data with the modified Gabor fit was over-determined because no envelope could be meaningfully determined. For that reason we excluded 9 cells from the Gabor parameter comparison, but we did include them in the best ITD comparison. Figure 7 compares the parameters of the fits for the data and predictions: the best ITD (Fig. 7A), frequency (Fig. 7B), phase (Fig. 7C), power value (Fig. 7D), envelope peak (Fig. 7E) and envelope width (Fig. 7F). Phases and frequencies once again showed the strongest correlations, while the envelope width, the worst. That is most likely due to the Gabor fit having difficulties in finding the most suitable envelope for rITD functions since they show less oscillatory behavior than difcors; note that the biggest discrepancies arise for the values of the largest envelope widths on the data fits. The largest outlier in Figure 7D and two largest outliers in 7E and 7F come from the same pair of cells. The underlying cause of the parameter mismatch for these cells is the discrepancies between locations of secondary peaks in measured and predicted rITD functions. These discrepancies gave rise to Gabor fits with very different envelopes which resulted in envelope-related parameter mismatch (power, envelope peak and envelope width). The Pearson's correlation coefficient ($r = 0.61$) in Figure 7D is reported with the outlier being excluded. We found that the underestimation of the

envelope width (Fig. 7F), rightmost points) happened only for the cells with very low average best frequency (the mean of BFs for contralateral and ipsilateral ears); mean BF for those cells was below 600 Hz.

Figures 4B and 7A show the relationships between the measured and predicted best ITDs from difcor (Fig. 4B, N = 43) and rITD functions (Fig. 7A, N = 31). For difcor best-ITDs, the Pearson's correlation coefficient was 0.92, the same as for rITD functions. We estimated absolute mismatches between the predicted and measured best ITDs; on average they were 43 ± 35 (standard deviation, s.d.) μs and 50 ± 37 μs for difcor and rITD function predictions, respectively.

Discussion

Summary of findings

We recorded sub- and suprathreshold responses of MSO neurons to wideband stimuli at moderate levels. The cross-correlation functions of monaural, subthreshold-dominated responses were used to directly predict the binaural rITD_{DIFF} data (Fig. 2) and the rITD function data by incorporating the nonlinear relation between interaural correlation and firing rate, measured with a mixed-noise technique (Fig. 5). Predicted ITD sensitivity was compared with the measured responses through a standard fitting function. We found that cross-correlation accurately accounted for binaural temporal parameters (phase and frequency), but did less well reproducing the exact envelope of measured ITD sensitivity functions.

Predicting ITD tuning by the cross-correlation of monaural inputs

Best ITDs were accurately predicted from cross-correlating subthreshold monaural inputs (Fig. 4B and Fig. 7A), in agreement with previous work based on monaurally evoked action potentials (Goldberg and Brown 1969; Moushegian et al. 1975; Spitzer and Semple 1995; Yin and Chan 1990) and subthreshold responses (van der Heijden et al. 2013). Mismatches between predicted and measured best ITDs were not systematic and generally smaller than 50 μs . Based on physiologically-relevant ITD range measurements (Maki and Furukawa 2005),

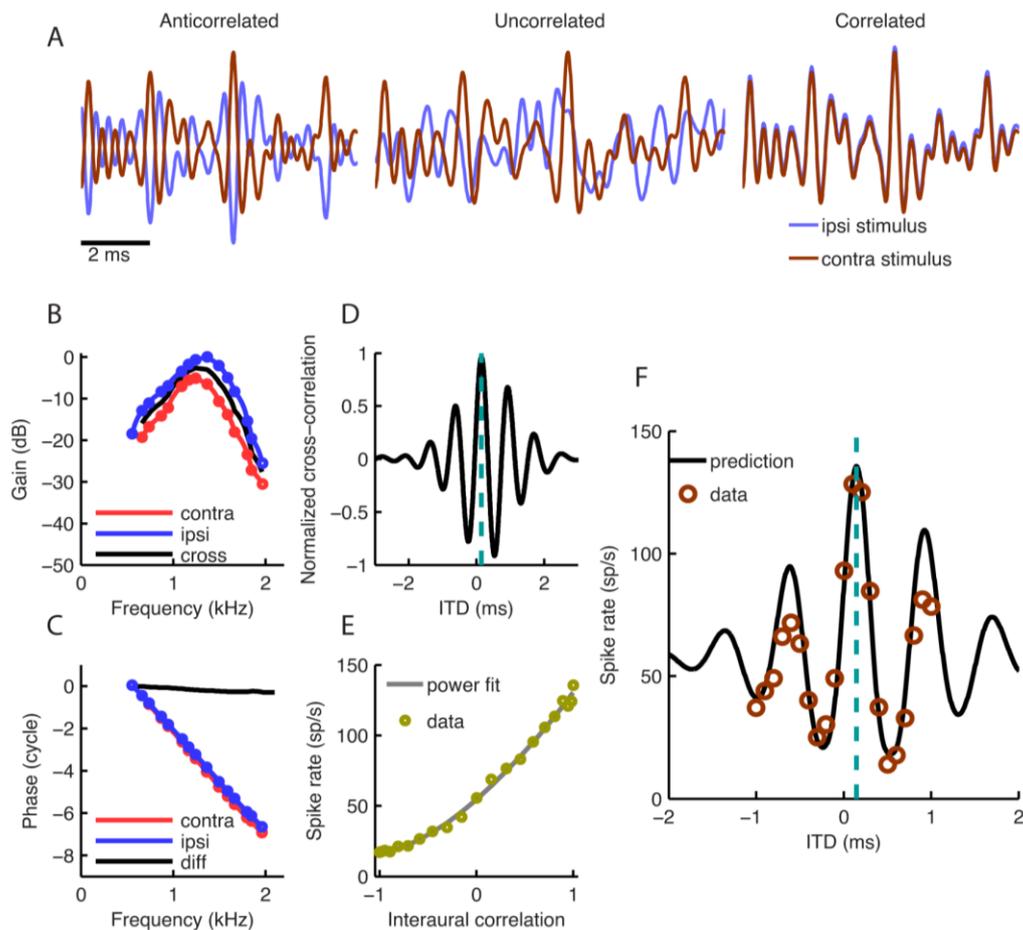


Figure 5. Predicting binaural responses of an MSO neuron.

(A) Excerpts of stimulus waveforms with varied interaural correlation when the presentations between ipsilateral and contralateral ears are anti-correlated (*left*), uncorrelated (*middle*) and correlated (*right*). The slight mismatch between waveforms in the example of correlated stimuli was imposed only for illustrational purposes. (B) Frequency response of an MSO neuron to monaural zwuis stimulation. Blue and red circles represent ipsilateral and contralateral ear responses, respectively. Best frequencies of the contralateral and ipsilateral ears were 1.23 kHz and 1.33 kHz, respectively. The response magnitudes are estimated with respect to the magnitude of the stimulus SPL (20 dB for this cell) and normalized to the maximum of the dominating monaural response. Black line represents the cross spectrum of monaural responses (see Results). (C) Phases to which the neuron ‘locks’ monaurally for each frequency component. Color coding is the same as in A. Black line shows the phase difference (contra minus ipsi). (D) A cross-correlation function (“crosscorr”) of the neuron. A crosscorr shows how ipsilateral and contralateral ear transfer functions correlate linearly. The cyan bar indicates the peak of the function (0.15 ms). (E) Neuron’s response to noise with varying interaural correlation (see Methods). Data points are indicated as yellow circles and the solid line is a power fit. (F) Comparison between the measured and the predicted ITD rate functions of the neuron. Using the spike rate values for each interaural correlation (E) the cross-correlation function (D) was converted into a rate-ITD function prediction (black solid line). The circles represent measured data; best ITD of the neuron was 0.14 ms. Cyan bar indicates the expected best ITD (0.15 ms). All data is from the same cell.

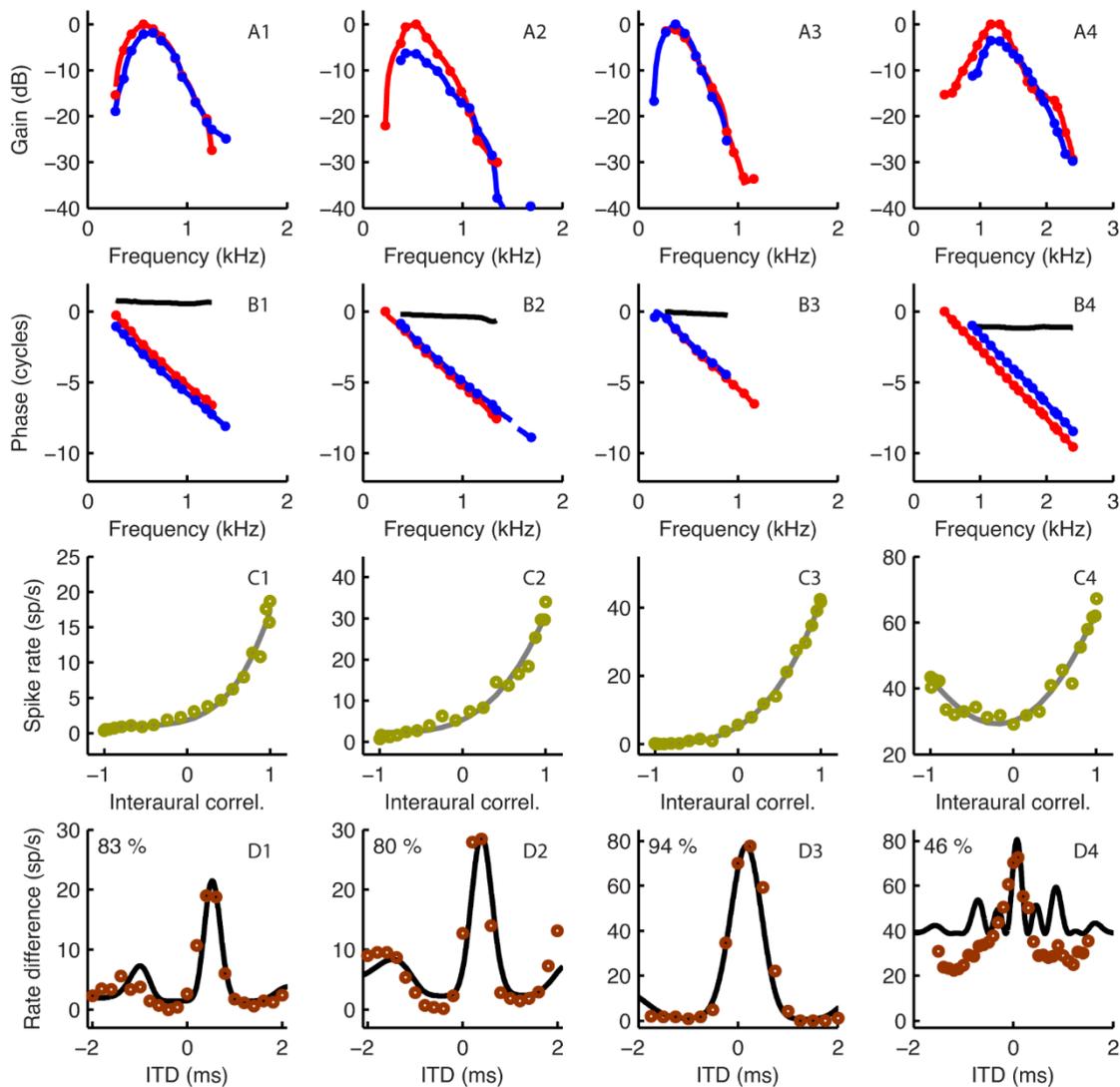


Figure 6. Examples of rate-ITD function predictions for four MSO cells.

For four cells (1-4), magnitudes (**A**) and phases (**B**) of monaural responses together with interaural correlation sensitivity (**C**) are used to predict the rate-ITD function (**D**, solid line) and compare it with the measured one (**D**, circles) (see Figure 5 and Results for more details). The solid line in **C** is a power fit function. Color coding is the same as in Figure 5. The percentage values in **D** indicate the variance of the measured data explained by the model. Best frequencies for contralateral and ipsilateral ears, best ITD and power value of power fit (p , see Methods) for each cell respectively are as follows: Cell 1 – 0.54 and 0.61 kHz/0.47 ms/4.38; Cell 2 – 0.49 and 0.47 kHz/0.3 ms/3.25; Cell 3 – 0.33 and 0.36 kHz/0.2 ms/3.12; Cell 4 – 1.23 and 1.31 kHz/0.08 ms/0.68. Cell 1 is also shown in Figure 3 as Cell 1. Cell 4 is an example of an MSO neuron which is prominently sensitive to the envelope component of the stimulus.

a 50- μ s delay at low frequencies translates into sound source angles of 20°-30°, comparable with the minimum resolvable angle in gerbils, which is at least 25° (Carney et al. 2011). Thus

our overall prediction accuracy is comparable with the gerbil's behavioral acuity in sound localization.

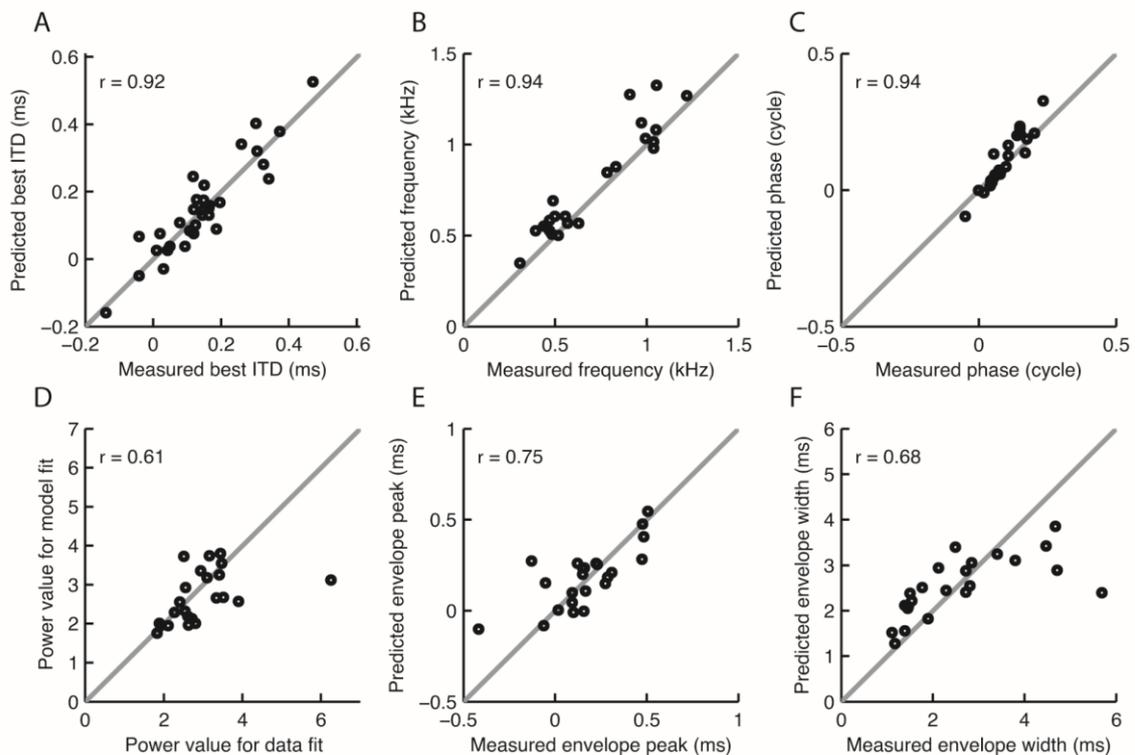


Figure 7. Comparing model and rate-ITD function fit parameters on a population level.

(**A-F**) Scatter plots of power-modified Gabor fit parameters (**A** – best ITD; **B** – frequency; **C** – starting phase; **D** – power; **E** – envelope peak time; **F** - full width at half maximum of the envelope) for rate-ITD functions and their predictions of the MSO neurons for which the predictions and fits were available ($N = 32$ for **A**, $N = 23$ for **B-F**). Grey lines indicate identity. Values in each plot show the Pearson's correlation coefficient r . Correlation coefficient in **D** is reported with the outlier excluded.

Our $rITD_{DIFF}$ functions had a damped oscillatory shape that was well captured by Gabor functions and that resembled $rITD_{DIFF}$ functions obtained in the inferior colliculus of cats (Mc Laughlin et al. 2008). Oscillation rate, oscillation phase offset and envelope position of the $rITD_{DIFF}$ functions were well predicted (Fig. 4C-E). Envelope width was systematically overestimated (Fig. 4F), corresponding to a slight underestimation of the main peak height in individual cases (Fig. 3). The underestimation may reflect the failure of $rITD_{DIFF}$ functions to fully linearize the relation between stimulus correlation and spike rate. When this relation is

highly supralinear ($\text{power} \gg 2$), the subtraction of normal and polarity-inverted curves will flatten, but not fully straighten it. Such higher p values were regularly encountered in the present study (see Results), and the range of p values was similar to that found in the IC of the guinea pig (Shackleton et al. 2005) and cat (Mc Laughlin et al. 2014). Another, related factor that has likely contributed to the underestimation of the central peak is the sharpening of phase locking in the cochlear nucleus compared to the auditory nerve (AN). Cross-correlation functions of the effective stimuli produce accurate predictions of linearized correlograms (difcors) in the AN, but are often too symmetric to match well the difcors from bushy cells with their sharp central peak (Joris et al. 2006; Louage et al. 2005). Figure 9C of Louage et al. (2005) shows the sharpness of the central peaks that cannot be matched by Gabor functions (compare our Figure 3). Thus, the mismatches in Figure 3D, rather than undermining the adequacy of coincidence detection to characterize MSO responses, reveal the limitations of modeling monaural inputs by a linearly filtered version of the acoustic stimulus. Overcoming this limitation would require an event-based analysis in which the cross-correlation functions are replaced by correlograms, but this is beyond the scope of the current study.

The raw wideband rITD functions (Fig. 5 and Fig. 6) had a damped oscillatory shape similar to the rITD_{DIFF} functions, but in addition showed a substantial degree of "rectification": a selective enhancement of peaks and flattening of troughs. They strongly resemble wideband rITD functions in cat MSO (Yin and Chan 1990). Comparison of modified Gabor fits with a power parameter that captured the amount of nonlinearity (Fig. 7) revealed that the monaural predictions accounted well for the starting phase and oscillation rate (Fig. 7B-C), but did less well predicting the envelope characteristics (Fig. 7D-F). The poorer match of the envelope parameters may originate in part from using more fit parameters and the risk of "over-fitting" the data; in particular, envelope width (Fig. 7F) and power value (Fig. 7D) likely interact, as they both control peakedness. This tradeoff may account for the spread in Fig. 7D-F. Interestingly, the systematic underestimation in peakedness in the rITD_{DIFF} predictions (Fig. 3C and Fig. 4F) did not reoccur in the predictions of the raw rITD functions, indicating that our use of rIC functions (Fig. 6C) adequately captured the nonlinearities that had limited the rITD_{DIFF} predictions. This finding is the physiological counterpart of the successful use of

interaural correlation in modeling a large body of psychophysical data (Colburn 1996; Sayers and Cherry 1957).

Lack of evidence for asymmetries in processing binaural inputs

The main aim of our study was to test to what extent binaural responses could be predicted from monaural responses. The monaural responses are dominated by fast fluctuations, which are the excitatory EPSPs, but additionally, IPSPs are present as well (Franken et al. 2015; Grothe and Sanes 1994; Roberts et al. 2013; 2014). These monaural synaptic inputs are modified as they interact with voltage-dependent ion channels (Franken et al. 2015; Khurana et al. 2011; Mathews et al. 2010; Scott et al. 2010). Even though our somatic recordings do not allow us to further dissect the monaural signals, the possibility to control the ipsi- and contralateral inputs separately, in combination with the anatomical segregation of the excitatory inputs provides a unique opportunity to study the somatic integration of these monaural inputs. Brand et al. (2002) observed a broadening of the window for ITD detection and a shift of best ITDs towards 0 μ sec in the presence of the glycine receptor antagonist strychnine. They proposed a model in which contralateral IPSPs affected the early phase of ipsilateral EPSPs more than their late phase. Subsequent modeling and experimental studies showed that to create the differential impact on early and late phases of the EPSP, fast kinetics and/or high conductance were needed for the glycinergic inputs, and it is still debated to what extent these requirements match the physiological characteristics of these inputs (Leibold 2010; Myoga et al. 2014; Roberts et al. 2014; Zhou et al. 2005). Recently, Franken et al. (2015) presented evidence suggesting that the strychnine-induced shift originally observed by Brand et al. (2002) and Pecka et al. (2008) was caused by an off-target effect on I_h channels. The experiments presented here also do not support this well-timed inhibition theory, since we did not find evidence for the predicted differential filtering of ipsi- and contralateral inputs during their interaction. Our ability to predict binaural responses from monaural responses argues against strong shifts resulting from nonlinear interactions between ipsi- and contralateral inputs. The successful prediction of entire rITD functions confirms and strengthens previous evidence in favor of coincidence detection as the basic mechanism of ITD tuning (Goldberg and Brown 1969; Yin and Chan 1990), as originally proposed by Jeffress (1948), even though this apparently simple scheme may involve complex interactions

between I_h , Kv1, Na channels and both excitatory and inhibitory conductances (Khurana et al. 2011; Roberts et al. 2014; Scott et al. 2010), and we cannot exclude that some of the small deviations we found in our predictions are caused by these interactions (Franken et al. 2015). Similarly, *in vivo* recordings did not yield evidence for asymmetries in the rising phase of EPSPs from either side (Franken et al. 2015; van der Heijden et al. 2013). Asymmetries in the origin of the axon, as proposed by Zhou et al. (2005), seem not to be borne out by the typically somatic axonal origin in most MSO neurons (Kuwabara and Zook 1999; Rautenberg et al. 2009; Scott et al. 2007; Scott et al. 2005). Franken et al. (2015) found evidence for systematic shifts in ITD in some gerbil MSO neurons when cross-correlating monaural EPSP period histograms, showing that their ITD tuning often deviated from instantaneous coincidence detection. A direct comparison with their results is not easy, since they used high-intensity binaural beats at frequencies typically much below CF, whereas we used broadband stimuli at intensities not far above the absolute hearing threshold.

Comparison to avian ITD tuning

In overall methodology the present study resembles that of (Fischer et al. 2008; Fischer et al. 2011) in the nucleus laminaris of the barn owl. Both studies used juxtacellular recordings to obtain a linear representation of monaural inputs, which in turn led to predictions of ITD tuning. The major common finding between the avian studies and our study is the successful prediction of ITD tuning from cross-correlating the monaural responses, indicating that nonlinear binaural interactions are unlikely to make a strong contribution to their ITD tuning. The major difference is the adequacy of a linear representation of the monaural inputs. The avian rITD functions are symmetric and lack the strong rectification and dominance of a central peak that characterizes the mammalian rITD functions. The comparatively linear character of the avian system may reflect both physiological and anatomical aspects (Macleod and Carr 2011). The high CFs of the avian AN fibers in the barn owl (up to 8 kHz) tend to linearize the neural representation of the acoustic of individual inputs, resulting in cycle histograms of high-CF avian AN fibers that are sinusoidal, i.e., faithfully reflect the stimulus waveform (Köppl 1997), and a similar faithfulness is preserved in the nucleus laminaris (Funabiki et al. 2011). In contrast, the lower-CF mammalian cycle histograms are strongly peaked in the AN and become further sharpened at the level of the cochlear nucleus (Joris et

al. 1994a). The sharpening and rectification is clearly visible in the periodograms of monaural MSO responses in van der Heijden et al. (2013). In the avian system the large number of monaural inputs to each nucleus laminaris neuron is likely to further linearize the neural representation of the monaural waveforms, in particular if the timing differs somewhat among individual inputs (Macleod and Carr 2011). In contrast, mammalian MSO neurons have a comparatively small number of monaural inputs (Couchman et al. 2010).

Conclusion

In sum, not only the best ITDs but the entire ITD tuning of MSO neurons in response to wideband sounds is well predicted by a straightforward cross-correlation of monaural responses. This supports Jeffress (1948) hypothesis that ITD sensitivity is realized by a simple coincidence detection of monaural inputs. The imperfections in the predictions most likely mainly originate from modeling the monaural inputs by the bandpass-filtered acoustic waveform: unlike the avian system, where the inputs carry a faithful representation of the (bandpass-filtered) acoustic waveform, monaural inputs to MSO show strong nonlinear distortion in the form of rectification and temporal sharpening. We conclude that an improved understanding of the major aspects of ITD tuning (temporal acuity, range of best ITDs, etc.) should not primarily be sought in the details of processing by MSO neurons themselves, but in a better characterization of the monaural inputs and their convergence onto MSO.

Grants

This work was supported by the Dutch Fund for Economic Structure Reinforcement (FES, 0908 'NeuroBasic PharmaPhenomics project').

Chapter 4

A test of the stereausis hypothesis for sound localization in mammals

Andrius Plauška, Marcel van der Heijden, J. Gerard G. Borst

The Journal of Neuroscience. 2017; 37(30):7278-7289

Abstract

The relative arrival times of sounds at both ears constitute an important cue for localization of low-frequency sounds in the horizontal plane. The binaural neurons of the medial superior olive (MSO) act as coincidence detectors that fire when inputs from both ears arrive near simultaneously. Each principal neuron in the MSO is tuned to its own best interaural time difference (ITD), indicating the presence of an internal delay, a difference in the travel times from either ear to the MSO. According to the stereausis hypothesis, differences in wave propagation along the cochlea could provide the delays necessary for coincidence detection if the ipsi- and contralateral inputs originated from different cochlear positions, with different frequency tuning. We therefore investigated the relation between interaural mismatches in frequency tuning and ITD tuning during *in vivo* loose-patch (juxtacellular) recordings from principal neurons of the MSO of anesthetized female gerbils. Cochlear delays can be bypassed by directly stimulating the auditory nerve; in agreement with the stereausis hypothesis, tuning for timing differences during bilateral electrical stimulation of the round windows differed markedly from ITD tuning in the same cells. Moreover, some neurons showed a frequency tuning mismatch that was sufficiently large to have a potential impact on ITD tuning. However, we did not find a correlation between frequency tuning mismatches and best ITDs. Our data thus suggest that axonal delays dominate ITD tuning.

Significance Statement

Neurons in the medial superior olive (MSO) play a unique role in sound localization owing to their ability to compare the relative arrival time of low frequency sounds at both ears. They fire maximally when the difference in sound arrival time exactly compensates for the internal delay: the difference in travel time from either ear to the MSO neuron. We tested whether differences in cochlear delay systematically contribute to the total travel time by comparing for individual MSO neurons the best difference in arrival times, as predicted from the frequency tuning for either ear, and the actual best difference. No systematic relation was observed, emphasizing the dominant contribution of axonal delays to the internal delay.

Introduction

Sound sources that are not straight ahead or behind travel different distances to both ears. As a result, interaural time differences (ITDs) are created. These ITDs are an important cue to calculate the location of a sound in azimuth, and in humans ITD cues are more important than differences in intensity (Wightman and Kistler 1992). These computations are performed within the CNS, by a specialized system in the brainstem. A central role in these computations is played by the principal neurons of the medial superior olive (MSO), which are excited by inputs from spherical bushy cells (SBCs) of the cochlear nucleus from both the ipsi- and the contralateral side (Thompson and Schofield 2000). MSO neurons have a bipolar dendritic orientation, and the ipsi- and contralateral inputs are segregated to the two main branches of these neurons (Smith et al. 1993; Stotler 1953). They function as coincidence detectors; their firing probability depends on the relative arrival time of the inputs from both ears; when the inputs arrive coincidentally, the firing probability of the MSO neurons reaches its maximum (Goldberg and Brown 1969; Moushegian et al. 1975; Spitzer and Semple 1995; van der Heijden et al. 2013; Yin and Chan 1990). Importantly, this so-called best ITD (BITD) varies between cells, and is typically different from 0 μ sec, indicating that the travel time for the signals from both ears to the MSO neurons must be different, with for most neurons a shorter travel time for ipsilateral than for contralateral signals (Crow et al. 1978; Goldberg and Brown 1969; Moushegian et al. 1975; Pecka et al. 2008; Spitzer and Semple 1995).

In a classical model by Jeffress, the internal delay was postulated to be entirely due to differences in axonal travel time (Jeffress 1948). Efforts to test this hypothesis have met with mixed success. In birds, there is good evidence that a difference in axonal delay can explain to a large part the distribution of BITDs (Carr et al. 2015; Seidl et al. 2014; Seidl et al. 2010). In mammals, evidence is more equivocal (Beckius et al. 1999; Karino et al. 2011; Smith et al. 1993). An alternative to the Jeffress' hypothesis is the so-called stereausis hypothesis, which proposes that differences in wave propagation time along the basilar membrane due to asymmetric innervation can provide the necessary delays (Bonham and Lewis 1999; Schroeder 1977; Shamma et al. 1989). This alternative is based on the finite speed of the

traveling wave in the cochlea, which travels from the base to the apex. For low frequency waves in the apex, travel time differences between neighboring locations are large compared to typically measured BITDs. Hence, small interaural differences in characteristic frequency (CF), the frequency for which auditory thresholds are lowest, suffice to create substantial differences in internal delay (Bonham and Lewis 1999; Joris et al. 2006). Reports of frequency tuning for individual MSO neurons have shown anecdotal evidence for disparities in best frequency (van der Heijden et al. 2013). There is some evidence for frequency-dependent internal delays (Day and Semple 2011). However, because of the difficulties in recording from MSO neurons, in combination with the often sparse firing induced by monaural sound stimulation, a direct, comprehensive test of the stereausis hypothesis has not yet been performed in mammals.

We recently showed that loose-patch (juxtacellular) recordings can be used to record synaptic inputs to the gerbil MSO neurons *in vivo* (van der Heijden et al. 2013). Gerbils have unusually good low-frequency hearing, and, in contrast to rats and mice, a well-developed MSO (Irving and Harrison 1967; Rautenberg et al. 2009). This opens up the possibility for a direct test of the stereausis theory. Here, we therefore compared mismatches in frequency tuning with binaural tuning in the same MSO neuron; in a subset of experiments we also compared auditory and electrical binaural tuning, thus testing key predictions of the stereausis theory.

Materials and Methods

Animal Procedures

All experiments were conducted in accordance with the European Communities Council Directive (86/609/EEC), and were approved by the institutional animal ethics committee. Young-adult female Mongolian gerbils with an average body mass of 60 g were anesthetized intraperitoneally with a ketamine-xylazine solution (114 and 17 mg/kg body weight, respectively). Anesthesia was maintained by administering one third of the induction volume at regular time intervals; the animal remained in areflexic state throughout the experiment. Rectal temperature was maintained at 37 °C using an electrical heating pad.

The surgical approach to the MSO has been described in detail previously (Kuwada et al. 1984; Plauška et al. 2016). Briefly, the head was fixed with a metal head-plate and the animal was positioned in supine position. Both pinnae were removed to expose the middle ear cavities and speakers were attached to both ears using a thin tube. The skin, connective tissue, salivary glands, lymph nodes and muscles covering both bullae were surgically removed. The animal was intubated and kept breathing independently. The right bulla was opened as wide as possible. A hole was also made in the left bulla to maintain the same pressure conditions in both middle ears. A ~1 mm diameter craniotomy was made to expose the brainstem on the right side. The meninges were left intact. The electrode insertion angle with respect to the craniotomy could be changed using a fixed-pivotal-point, custom-built, positioning device on which the animal was laying during the experiment.

For electrical round window stimulation experiments, both facial nerves were cut to prevent facial muscle activation by the electrical stimulation, and a silver-ball electrode was placed in contact with each round window through a lateral opening in the bulla. Electrodes were fixed to the skull with Histoacryl (Braun; Ann Harbor, MI, USA) and wax (Sticky Wax; Kerr, Bioggio, Switzerland). Local ground electrodes were placed close to the bony part of the external acoustic meatus.

In Vivo Loose-patch Recordings

Thick-walled borosilicate glass microelectrodes (4 – 7 M Ω resistance; 1 – 1.5 μ m tip diameter) were used for loose-patch (juxtacellular) recordings. Most of the recordings were done using pipettes filled with intracellular solution containing (in mM): K-gluconate 138, KCl 8, Na₂-phosphocreatine 10, MgATP 4, Na₂GTP 0.3, EGTA 0.5, HEPES 10 mM (pH adjusted to 7.2 with NaOH). Less than 12% of the cells reported here were recorded using normal rat Ringer's solution containing (in mM): NaCl 135, KCl 5.4, MgCl₂ 1, CaCl₂ 1.8, HEPES 5 mM (pH adjusted to 7.2 with NaOH). No clear differences could be found between the responses of cells recorded with either solution so the recordings were pooled.

High positive pressure (70 – 100 mbar) was applied to the pipette during brain surface penetration. After successful penetration the pressure was lowered to 20 – 30 mbar, and we

waited for a few minutes before further advancing the electrode to minimize the impact of brain tissue movements relative to the electrode during recordings.

The somatic layer of the MSO was identified by monitoring local field potentials (“neurophonics”; (Biedenbach and Freeman 1964; Clark and Dunlop 1968; Galambos et al. 1959; Mc Laughlin et al. 2010), which were evoked by 2 ms alternating clicks to both ears (Kuwada et al. 1984). Upon reaching the somatic layer, the electrode was advanced in small steps and its resistance was closely monitored. A gradual increase in resistance together with the appearance of electrophysiological activity indicated contact with a neuron. Subsequently the positive pressure on the electrode was released and the electrode was advanced up to 10 μm to establish the loose-patch recording configuration. Recordings were done in current clamp mode. In case of changes in cell response, the recordings were stopped. Data were acquired using a Multiclamp 700B (Molecular Devices, Foster City, CA) amplifier using custom software written in MATLAB (7.6.0; The MathWorks, Natick, MA, USA).

Auditory Stimulation

Auditory stimuli were generated using custom software written in MATLAB and realized through a 24-bit D/A-channel (RX6; Tucker Davis Technologies, TDT, Alachua, FL, USA; 111.6 kHz), programmable attenuator (PA5; TDT) and an amplifier (SA1; TDT). Stimuli were delivered to the ear canals in a close-field fashion through Shure speakers (frequency range 22 Hz to 17.5 kHz) and a pair of small (~11 cm length) tubes. The lowest SPL at which click stimuli evoked neurophonics was considered to be the animal’s hearing threshold. Three types of auditory stimuli were used in this study: monaurally and binaurally presented irregular tone complexes and binaurally presented pure tones.

The type of multitone stimulus used in this study (‘zwuis’) has been described in detail elsewhere (van der Heijden and Joris 2003; 2006). In short, zwuis stimuli are produced by summing multiple irregularly spaced frequency components with the same amplitude and a random phase, choosing the frequencies in such a way that neither second-order nor third-order distortion products in the response match any of the components themselves. The stimuli in this study contain 30 components spanning a frequency range of 50-3000 Hz.

For monaural stimulation, the zwuis stimulus was presented in 300 ms bursts; its intensity ranged from -10 to 50 dB SPL per component in 10 dB steps; twenty repetitions were presented for each condition; each zwuis stimulus was followed by 100 ms silence; total duration was 58 s. Binaural zwuis, which was presented at the same burst intervals as the monaural stimulus, had systematic time delays between the ears ('noise delay stimulus'). Each condition was presented twenty times. Pure tone stimuli were presented binaurally with different ITDs and frequencies. The frequency range for each pure tone stimulus was centered on the estimated best frequency (BF) of that neuron, and was presented in 100 Hz steps. Each presentation started with a 20 ms silent period, followed by a 70 ms burst and another 50 ms silent period. All conditions were presented ten times. For both binaural zwuis and pure tone stimuli, ITD ranges depended on the frequency sensitivity of the neuron; stimuli intensity levels were between 20 – 70 dB SPL. Sound intensity levels for binaural stimuli were chosen to be 20 – 30 dB above hearing threshold. Stimuli waveforms were recomputed for a different instance of their presentation, but were kept the same within one instance for all different stimulus conditions and repetitions.

To observe secondary peaks in rITD functions of low-frequency (< 700 Hz) neurons, we increased the ITD range of the stimuli. Consequentially, to conserve the recording duration, we increased the stimulus step size; 8 out of 68 neurons were stimulated binaurally with ITD steps >0.2 ms up to a maximum of 0.4 ms.

Electrical Stimulation

Electrical pulses for round window stimulation were generated using a homemade bipolar current stimulator. Pulses were 100 μ s in duration. Current intensities at both ears were varied between 0.2 – 0.8 mA with the aim of finding levels at which a subthreshold response was evoked by stimulating at either ear, but suprathreshold responses when stimulating at both ears. Binaural electrical stimulation was presented with time delays ranging from -2 to 2 ms in 0.2 ms steps. For all electrical stimulation recordings, the stimulation window was 100 ms, preceded and followed by a 50 ms period without stimulation. Stimulus frequencies ranged from 20– 120 pulses/s. No obvious differences in binaural stimulation were obtained at the different frequencies, and responses were therefore pooled within a cell.

MSO cell admission

Our previous study showed that loose-patch (juxtacellular) recordings are suitable to resolve synaptic events in MSO neurons (van der Heijden et al. 2013). We observed that the best quality loose-patch recordings were made when the seal resistance was between 20 and 70 M Ω ; at lower resistances the contribution of field potentials became too high, higher resistances caused strong waveform filtering. In this study only recordings with seal resistances between 20 and 70 M Ω were accepted as valid loose-patch recordings.

Each new stimulus block was preceded by a silent period of 1 or 5 s. These baseline periods were used to judge recording stability. Details of the method are presented in Plauška et al. (2016). Briefly, a power spectrum of pre-stimulus baselines was estimated for all recordings from a cell. Its value at 1 kHz, which predominantly reflected the spontaneous activity of the neuron, had to remain within a 5 dB window for the stimulus block to be included in the analysis.

Action potentials were detected offline based on a threshold criterion for the maximum repolarization rate of individual events (van der Heijden et al. 2013). Only cells for which histograms of negative peak sizes showed clear bimodality were accepted for further analysis. For rITD functions, the first 10 ms of the response to each stimulus presentation were excluded from analysis to avoid onset effects.

Responses evoked by electrical round window stimulation were more difficult to analyze because of the presence of larger contamination by field potentials, presumably due to the hypersynchronous excitation. The analysis window was restricted to latencies of 2.5-8 ms from the contralateral electrode. At shorter latencies unambiguous AP identification was typically not possible.

Experimental Design and Statistical Analysis

To test the stereausis hypothesis we investigated whether BITD and characteristic frequency (CF) mismatches were systematically correlated. Apart from the criteria given in the previous section, to be included in this comparison cells had to exhibit significant differences in the means of responses to different time delays in their rate ITD (rITD) functions, which was assessed with a single-sided ANOVA test (function *anova1* in MATLAB) with the threshold for

binaural sensitivity at $p < 0.01$. For composite rITD functions, a dynamic range criterion was defined as half of the difference between the sums of the two highest and the two lowest spike counts in the response. Based on visual inspection of all composite rITD functions, only composite rITD functions with a minimum dynamic range of 10 action potentials were accepted. To obtain BITDs, rITD functions were fit with a modified Gabor function:

$$G_p(\tau) = A + Bg\left(p, e^{-0.5\left(\frac{\tau-\tau_0}{w}\right)^2} \cos(2\pi(f\tau - \phi_0))\right) \quad (\text{Eq. 1})$$

where, $g(p, x) = \frac{(x+1)^p}{2}$, τ is time (ms), A – offset (spikes/s), B – amplitude (spikes/s), τ_0 – peak time of the envelope (ms), w – width of the envelope (ms), f – frequency (kHz), ϕ_0 – start phase of cosine (cycle), the power parameter p modifies the envelope of the Gabor function to account for the typically asymmetric peak and trough relationship. BITD was defined as the ITD of the dominating peak in the case of wideband rITD functions, and the ITD of the central peak for composite rITD functions (Yin and Chan 1990). If fits of the wideband rITD functions with the modified Gabor function could account for <80% of the variance in the data, typically owing to low numbers of spikes, the wideband rITD functions were not included in further ITD-related analysis.

To obtain CF, responses to monaural zwuis stimuli were processed using Fourier spectral analysis. From the power spectrum, the baseline spectrum, obtained from the 1 s period preceding and following the stimulus, and an additional 5 dB were subtracted. The subtraction of the baseline spectrum serves to compensate for the spectral coloring caused by the finite bandwidth of individual elementary events responsible for the neural activity. If the events themselves were randomly timed, i.e., originated from a generating process having a flat spectrum, the spectrum of the resulting waveform equals that of the elementary event. More generally, the resultant waveform is a convolution of the waveforms of the point process and that of the elementary event, and their long term power spectra multiply (Ashida et al. 2013; Campbell 1909; Fesce 1990). The information about CF is contained in the former (the timing of the events, not their shape), which is retrieved by dividing out the baseline spectrum, i.e., subtracting it in the log domain. Monaural receptive fields were built by interpolating intermediate values (MATLAB function *contourf*). Responses to monaural

stimulation were fit with a weighted 6th order polynomial. Only responses with >10 (one third of total) signal components above the noise floor (i.e. signal-to-noise ratio > 1), were used. The noise floor was obtained from the spectral analysis of the recordings in the absence of sound stimuli. Best frequencies (BF) were defined as the fit peak; CF as the BF at the lowest sound intensity. The contour plots, polynomial fits and thresholding are illustrated in Figure 5.

Estimates for the standard deviation of the CF mismatch or of the BITD were obtained using bootstrap methods. The 20 responses to identical stimulus presentations were randomly divided in two groups of 10; BITDs for both groups were computed as described above, and the difference D between these two BITD estimates was computed. This procedure was repeated $N=25$ times using independent random divisions, resulting in N values for D . The reported standard deviation STD for the BITD estimates based on the complete set of 20 responses equals

$$STD = \sqrt{\frac{1}{2N} \sum_{k=1}^N D_k^2}, \quad (\text{Eq. 2})$$

which is based on applying the same procedure to sets of 20 independent numbers drawn from a normal distribution having unity variance. Standard deviations for CF mismatches were obtained by the same method, but now simultaneously subdividing the responses to repeated contra- and ipsilateral stimulation.

We used a t-statistic to test whether Pearson's r differed significantly from zero (function *corrcoef* in MATLAB).

CF mismatches were converted to predicted BITDs as described in the next section, and the correlation between BITDs and predicted BITDs was assessed using a bootstrap analysis as described in the Results and the legend to Figure 9.

Conversion of CF mismatches to predicted BITDs

Ipsi- and contralateral CF estimates were first converted to cochlear location (distance from the base) using the tonotopic map for the gerbil cochlea (Müller 1996). A given interaural mismatch DX (in m) in cochlear location corresponds to a travel time difference DT equal to

$$DT = \frac{DX}{c(CF)}, \quad (\text{Eq. 3})$$

where $c(CF)$ is the phase velocity of the traveling wave of frequency CF at its best site, which is related to wavelength λ at CF by

$$\lambda(CF) = \frac{c(CF)}{CF}, \quad (\text{Eq. 4})$$

Values of either $c(CF)$ or $\lambda(CF)$ in the apex were obtained from studies reporting large populations of auditory nerve responses to identical stimuli in chinchilla (Figure 3 of (Temchin et al. 2012), cat (Figure 9B of (van der Heijden and Joris 2006), and guinea pig (Figure 5B at 50 dB SPL of (Palmer and Shackleton 2009). The wavelength dependence on CF was found to be similar across species (Figure 1A), supporting the application to the gerbil, for which no data are available, and showed an approximately linear relation between λ and $\log(CF)$. We therefore pooled the data from the three studies and used linear regression to characterize this dependence. Application of equations 3-4 produced the predicted BITD from the measured CF mismatches. Note that the steepness of this dependency is greater for lower CF s (Figure 1B).

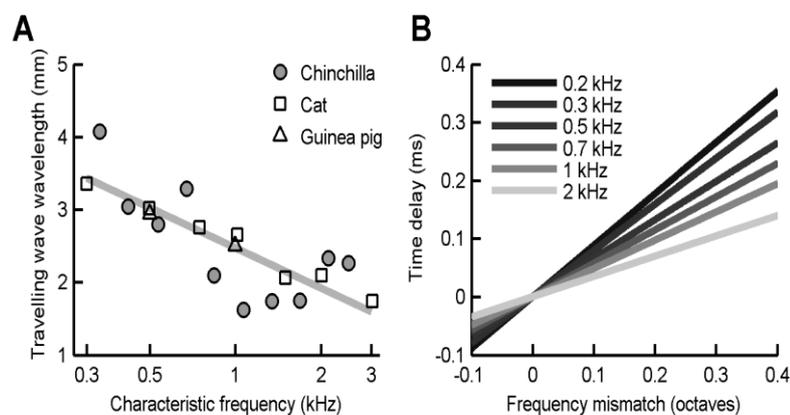


Figure 1. Cochlear time delays.

(A) Travelling wave wavelength dependence on the characteristic frequency in the cochlea for three different species: chinchilla (Temchin et al. 2012), cat (van der Heijden and Joris 2006) and guinea pig (Palmer and Shackleton 2009); circles, squares and triangles, respectively). Solid line shows the linear fit for all three species. **(B)** Theoretical cochlear time delay dependence on frequency mismatch between the ears for six different characteristic frequencies. Relationships were derived from linear fit in A.

Results

The stereausis theory makes several clear predictions. First, there should be distinct mismatches in frequency tuning for ipsi- and contralateral inputs, where the expectation would be a general bias for contralateral tuning to be to lower frequencies than ipsilateral tuning, thus creating an extra cochlear delay for contralateral sounds. A second prediction is that the central internal delay obtained by direct electrical stimulation of both cochleae at varying latencies can be different from the total internal delay for acoustic stimuli as reflected by the BITD for the same neuron. A third prediction is a systematic correlation between ITD tuning and frequency tuning mismatches for individual cells. We used a tailored, multitone stimulus to investigate the distribution of mismatches in frequency tuning in MSO neurons using loose-patch recordings, and to compare in individual neurons these mismatches with binaural tuning; in a subset of experiments we also compared auditory and electrical binaural tuning.

ITD tuning

We compared frequency and ITD tuning of principal neurons in the MSO of anesthetized gerbils using loose-patch recordings with the aim of testing whether there exists a systematic relation between the two, as predicted by the stereausis model. We will first present an overview of the ITD tuning, which was studied using binaural stimulation with irregularly spaced multi-tone ('zwuis') stimuli or with simple tone stimuli. Figure 2A shows an example of the response of an MSO neuron to binaural zwuis stimulation, illustrating prominent subthreshold activity and action potentials (asterisks). Figure 2B shows action potential rates at different ITDs for a low frequency MSO neuron. This rate ITD function (rITDf) was obtained by systematically varying the time delay between the zwuis stimuli to both ears. Spike rates during these measurements were 25 ± 27 sp/s (mean \pm SEM; range 0.3-129 sp/s; $N = 68$). The rITDf was fit with a modified Gabor function (see Methods) from which we extracted the BITD as the ITD at which the fit function had its maximum. The vertical bar in Figure 2B indicates the BITD of the neuron, which was 0.17 ms. On average, BITDs were 0.12 ± 0.12 ms (Figure 2C; $N = 68$), indicating a bias for contralateral ear leading, as observed previously for both

gerbils (Brand et al. 2002; Pecka et al. 2008; Spitzer and Semple 1995) and other species (Crow et al. 1978; Goldberg and Brown 1969; Moushegian et al. 1975). The grey area in Figure 2C demarcates the ecological ITD range of gerbils, which is about ± 0.13 ms (Maki and Furukawa 2005). More than half (57%) of the BITDs fell within the ecological ITD range.

To obtain an alternative estimate for the BITD, we also presented binaural pure tone stimuli at frequencies around the BF. In the example shown in Figure 3A, the different frequencies ranged from 0.4 – 0.9 kHz, each of which was presented in 0.2 ms steps between -2 and 2 ms at a stimulus intensity of 40 dB SPL. Different traces in Figure 3A correspond to spike count ITD functions in response to different frequencies. This neuron was most sensitive to 0.7 and 0.8 kHz tone stimuli, as expected from its BF, which was 0.78 kHz. We summed all the spike count ITD functions to produce a composite ITD function (Figure 3B; (Yin and Chan 1990). The BITD was obtained as the most central peak (vertical bar, 0.11 ms) of the Gabor fit (black line). A comparison of BITDs obtained with zwuis and pure tone stimuli showed good agreement for most cells ($N = 26$; Figure 3C; $r = 0.85$; $p < 0.0001$).

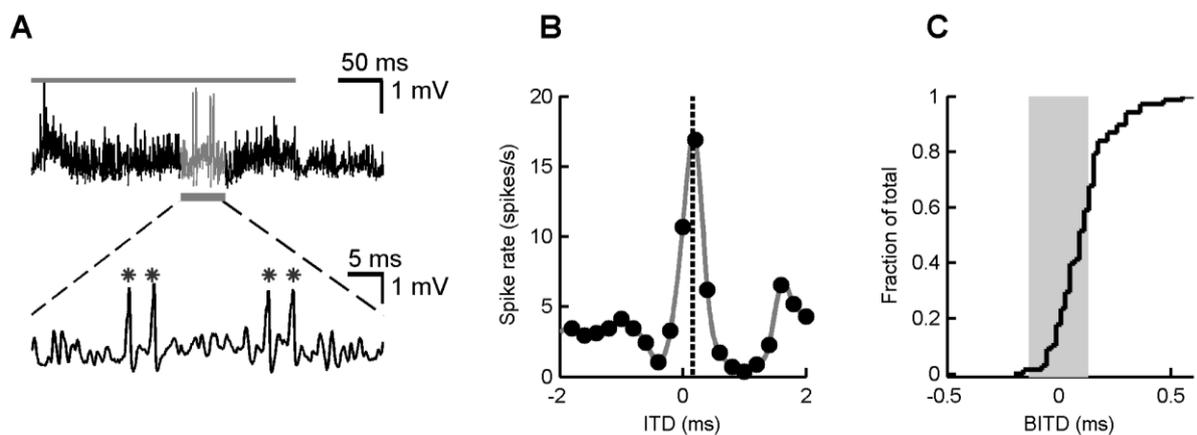


Figure 2. Determining BITDs of MSO neurons.

(A) Loose-patch (juxtacellular) recording of an MSO neuron during binaural stimulation with 300 ms ‘zwuis’ stimulus (grey bar) at 30 dB SPL. The grey portion of the waveform is shown at higher time resolution below, revealing subthreshold events and action potentials (*) during stimulation. (B) Example of a rate-ITD function (rITDf), showing firing rate as a function of ITD (positive ITD values: contralateral leading). Circles indicate measured spike rates; grey line is a cubic spline through the data points. BITD was 0.17 ms (vertical line). Same cell as A. (C) Cumulative histogram plots of BITDs. The grey area indicates the physiological ITD range for a gerbil (± 0.13 ms). In 60 of 68 cells, BITDs were biased towards contralateral ear leading.

Binaural MSO sensitivity to electrical round window stimulation

One prediction of the stereausis hypothesis is that ITD tuning may be different for auditory stimuli and when the cochlea is bypassed by electrical stimuli. To test this prediction, we recorded MSO responses to round window electrical stimulation, and compared the responses to those obtained by auditory stimulation within the same cell (Figure 4A). Electrical stimuli were presented both monaurally and binaurally. We kept increasing current intensity for each cochlea individually until we reliably started seeing subthreshold responses upon monaural stimulation and action potentials during binaural electrical stimulation. In some cells we also varied stimulation rate, however, this did not obviously affect BITD estimates and responses obtained at the same set of current intensities were pooled (see Methods). Current stimulation typically evoked a complex response consisting of a series of peaks and troughs with latencies ranging from <3 ms to >8 ms. The short-latency peaks varied little between trials and generally did not seem to evoke action potentials, suggesting that they were field potentials originating from more proximally located areas. Figure 4B shows an example of the MSO response to electrical round window stimulation. Blue and red traces correspond to ipsilateral and contralateral ear responses, respectively. Grey traces show responses to binaural stimulation at the electrical BITD for this recording. Black traces show averages of 10 repetitions under the same condition. The comparison of monaural and binaural stimulation illustrates that summation of electrically-evoked synaptic potentials at a latency of ~5.6 ms evoked the most action potentials for this cell during binaural stimulation.

Because of the large size of the field potentials and electrically induced movements at high stimulus intensities it was possible in only three out of a total of 18 MSO cells to compare an electrical BITD with the auditory BITD (Figure 4C). For these three cells CFs were not available. The minimum sound-evoked latencies (7-9 ms) were clearly longer than the electrical latencies (3.5-6 ms) in the same experiments. For all three cells there was a mismatch between auditory and current-evoked BITD: -0.12 vs -0.32 ms; cell 2, -0.01 vs 0.21 ms; cell 3, 0.38 vs 0.12 ms. These data therefore suggest that a difference in ipsi- and contralateral cochlear delay can make a substantial contribution to ITD tuning, as predicted by the stereausis hypothesis.

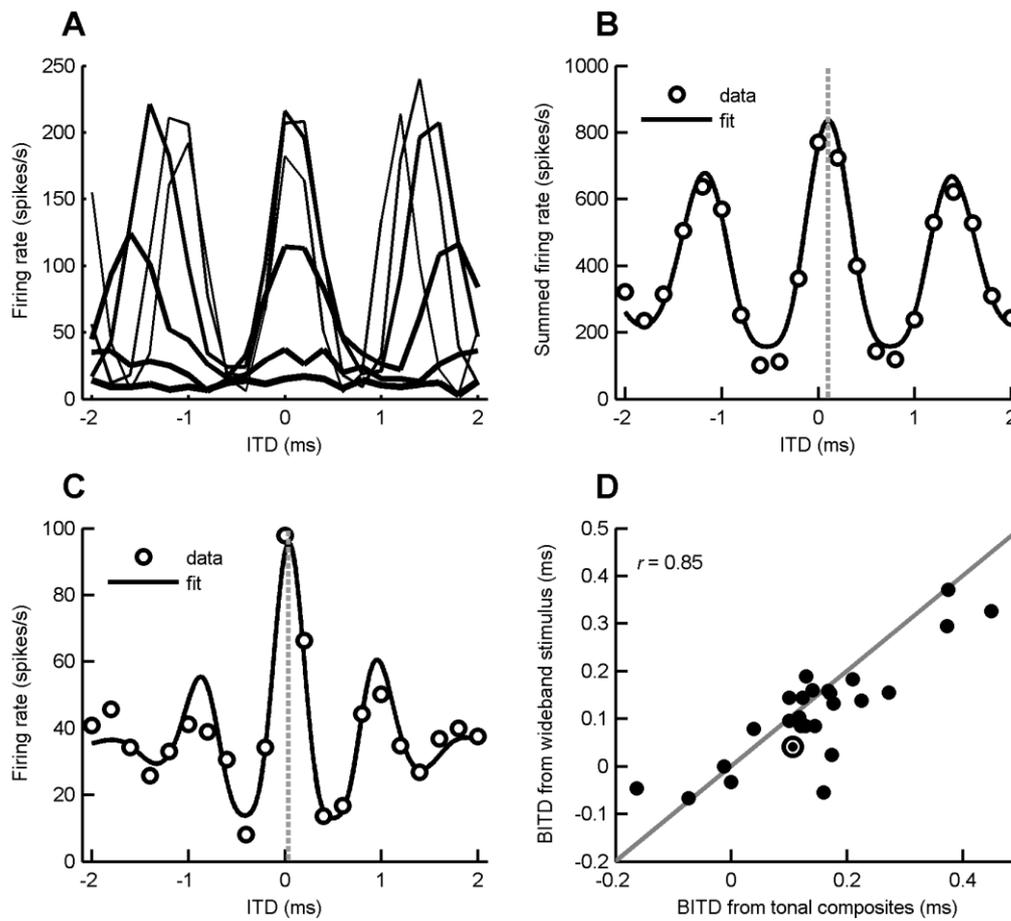


Figure 3. Composite ITD curves from tonal data.

(A) Superimposed tonal ITD curves at six different frequencies. Average BF of the neuron at 40 dB SPL was 0.78 kHz. Line thickness indicates tone frequency, varying in 100-Hz steps from 400 Hz (thickest line) to 900 Hz. (B) Composite ITD curve (circles) was obtained by adding the six tone responses shown in A. Solid line is fit with a Gabor function. Vertical line indicates the BITD (0.11 ms). (C) Wideband ITD curve (circles) and fitted Gabor function (solid line) for the same cell as shown in panel B. Vertical line indicates BITD (0.04 ms). (D) Comparison of BITDs from wideband rITDFs and from tonal stimulation composite ITD curves ($N = 26$; $r = 0.85$; $p < 0.0001$). Grey line indicates identity; only BITDs with estimated standard deviation < 0.25 ms were used. Highlighted symbol corresponds to the cell shown in panels B, C.

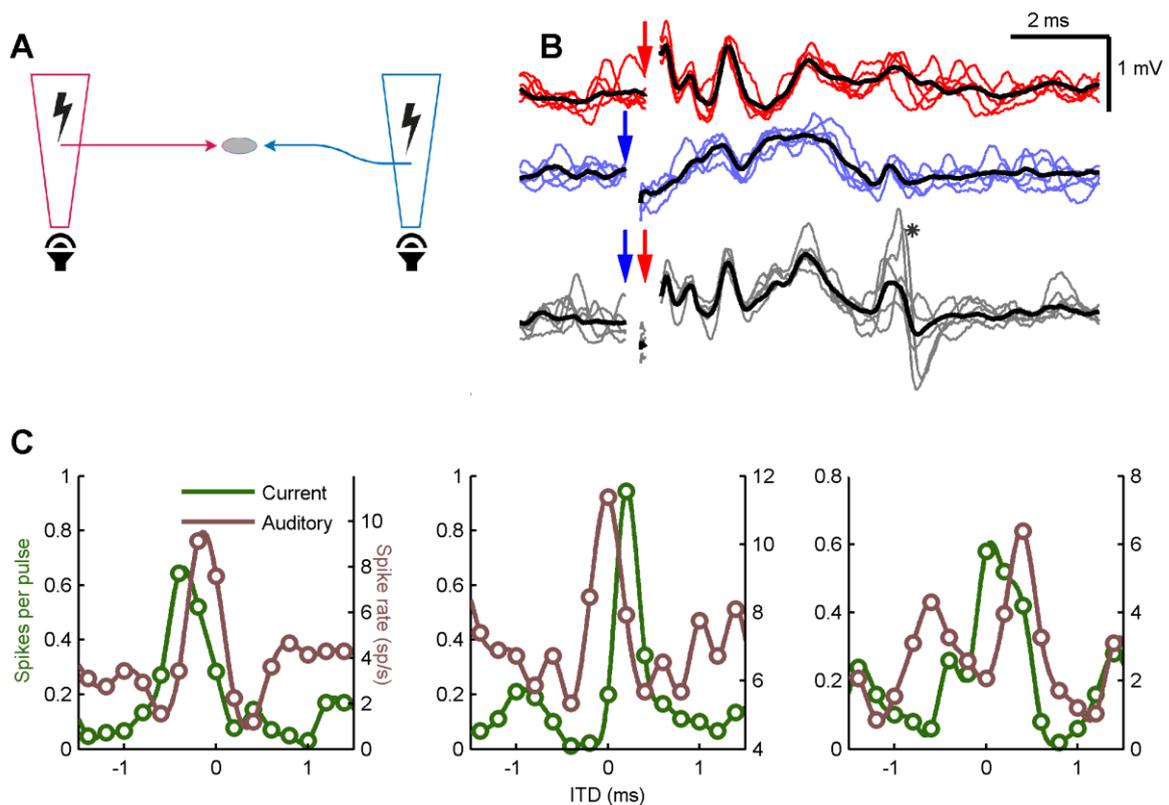


Figure 4. Binaural MSO sensitivity to electrical round window stimulation.

(A) Schematic representation of the two stimulation methods. Blue and red trapezoids symbolize ipsilateral and contralateral cochleae, respectively. Speaker icons indicate sound stimulation at the base of cochlea; bolt icons indicate electrical stimulation which bypasses the traveling wave. Two arrows pointing towards MSO cell represent neural pathways converging onto MSO. (B) Monaural and binaural MSO responses to electrical round window stimulation. Red, blue and grey traces show responses to contralateral-only, ipsilateral-only and both ear stimulation, respectively. Each set of traces shows 5 instances of individual responses and the black trace is an average of a total 10 repetitions. Large vertical deflections on the left are the stimuli artefacts. The three groups of traces were displaced with respect to each other in the vertical direction for visual clarity. The asterisk on the bottom trace indicates the location of two evoked action potentials. Arrows indicate the beginning of stimulus; stimulus artefacts were cut out for demonstrational purposes. (C) Comparison of rITD functions from wideband auditory and electrical round window stimulation data. Circles indicate data points, lines are interpolated values. The y-axis on the left is applicable for current stimulation rITD function and shows how many spikes on average were evoked by a single stimulus. The y-axis on the right is that of auditory stimulation rITD function and tells the spike rates evoked by the stimulus. BITDs for the three cells from left to right (auditory vs electrical): -0.12 vs -0.32 ms, -0.01 vs 0.21 ms, 0.38 vs 0.12 ms.

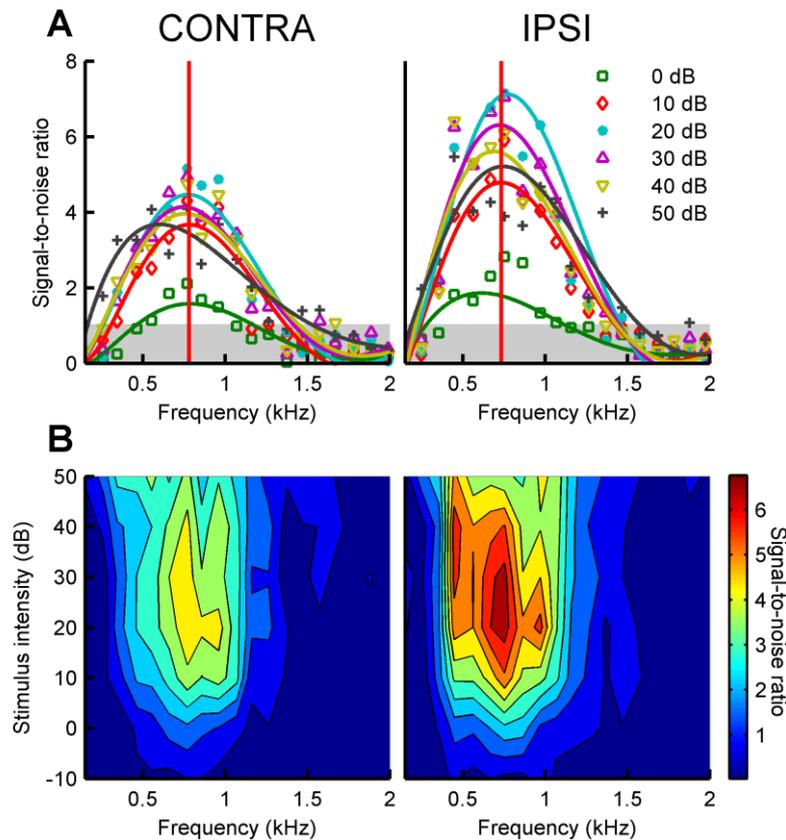


Figure 5. Determining the characteristic frequency of an MSO neuron.

(A) Individual monaural responses to zwuis stimuli presented at different sound intensity levels. Responses to different SPLs are represented by different colors; the numbers indicate SPL per tone component. Symbols show the measured data points, solid lines – the fits. Left and right plots show responses for contralateral and ipsilateral ear, respectively. Grey area demarcates the threshold where the response cannot be distinguished from the noise floor (see Methods). Red vertical lines indicate the characteristic frequencies, determined at 10 dB SPL stimulus intensity; CFs for contra- and ipsilateral ears were 0.78 and 0.74 kHz, respectively. (B) Monaural receptive fields of the MSO neuron determined using zwuis stimulus presented at different sound intensities (in 10-dB steps). Left and right plots show receptive fields for contralateral and ipsilateral ear, respectively.

Frequency tuning

Stereausis critically depends on a difference in frequency tuning for ipsi- and contralateral sound stimulation. We therefore compared frequency tuning for both ears using zwuis stimuli, presented at intensities ranging from -10 to 50 dB SPL per component. Figure 5A shows spectral amplitude components at the different frequency components of the monaural zwuis stimulus for an MSO neuron; solid lines show the fits with a polynomial

function. From these responses, the ipsi- and contralateral receptive fields of this cell were constructed (Figure 5B). Characteristic frequency (CF) was defined as the peak of the fit to the response function at the lowest stimulus intensity at which both ipsi- and contralateral ears responded with at least 10 significant components above a signal-to-noise ratio of 1. For this cell, this intensity was 10 dB SPL, yielding CFs for ipsi- and contralateral stimuli of 0.74 and 0.78 kHz, respectively (vertical lines in Figure 5A). Frequency tuning generally resembled the frequency tuning of SBCs (Caspary et al. 1994; Kopp-Scheinflug et al. 2002; Kuenzel et al. 2011). Mean CF was typically lower for deeper cells (*results not shown*), in agreement with the tonotopic organization of the MSO (Day and Semple 2011; Franken et al. 2015; Goldberg and Brown 1968; Guinan et al. 1972; Karino et al. 2011). The observed distribution of CFs was similar to frequency tuning based on binaural stimuli in some earlier studies (Brand et al. 2002; Day and Semple 2011), whereas in other studies in the gerbil MSO cells that were tuned to much higher frequencies were observed (Franken et al. 2015; Pecka et al. 2008). Recordings in the present study were typically made at a penetration depth of at least 300 μm ; we did not record from very superficial neurons, which constitute the neurons tuned to frequencies >2 kHz in the gerbil (Franken et al. 2015), since in superficial layers it was more difficult to ascertain that we recorded from the somatic layer using field potential recordings.

To estimate the difference in frequency tuning for responses to sound stimuli presented to either ear, we cross-correlated magnitude spectra of ipsi- and contralateral responses at the same intensity at which CF was determined (Figure 6A, B). The vertical bar in Figure 6B indicates the peak of the cross-correlation function; it was at 40 Hz, indicating that contralateral frequency tuning was to slightly higher frequencies than ipsilateral. This cross-correlation peak generally corresponded well with the difference in CFs between both ears, but we consider the former a more robust estimate of frequency tuning mismatch than the difference in CFs, since it takes the entire frequency curve into account. We thus estimated CF mismatches for 83 cells by cross-correlation (Figure 6C). Mismatches up to 400 Hz were observed, but most mismatches were much smaller. Even though the frequency mismatches were not very large, they were larger than the frequency tuning mismatches that were observed in the nucleus laminaris of the barn owl (Fischer and Peña 2009; Peña et al. 2001) or alligator (Carr et al. 2009). Average mismatch was -8 ± 115 Hz, suggesting that there was no overall preferred mismatch direction. Figure 6D shows the absence of a significant corre-

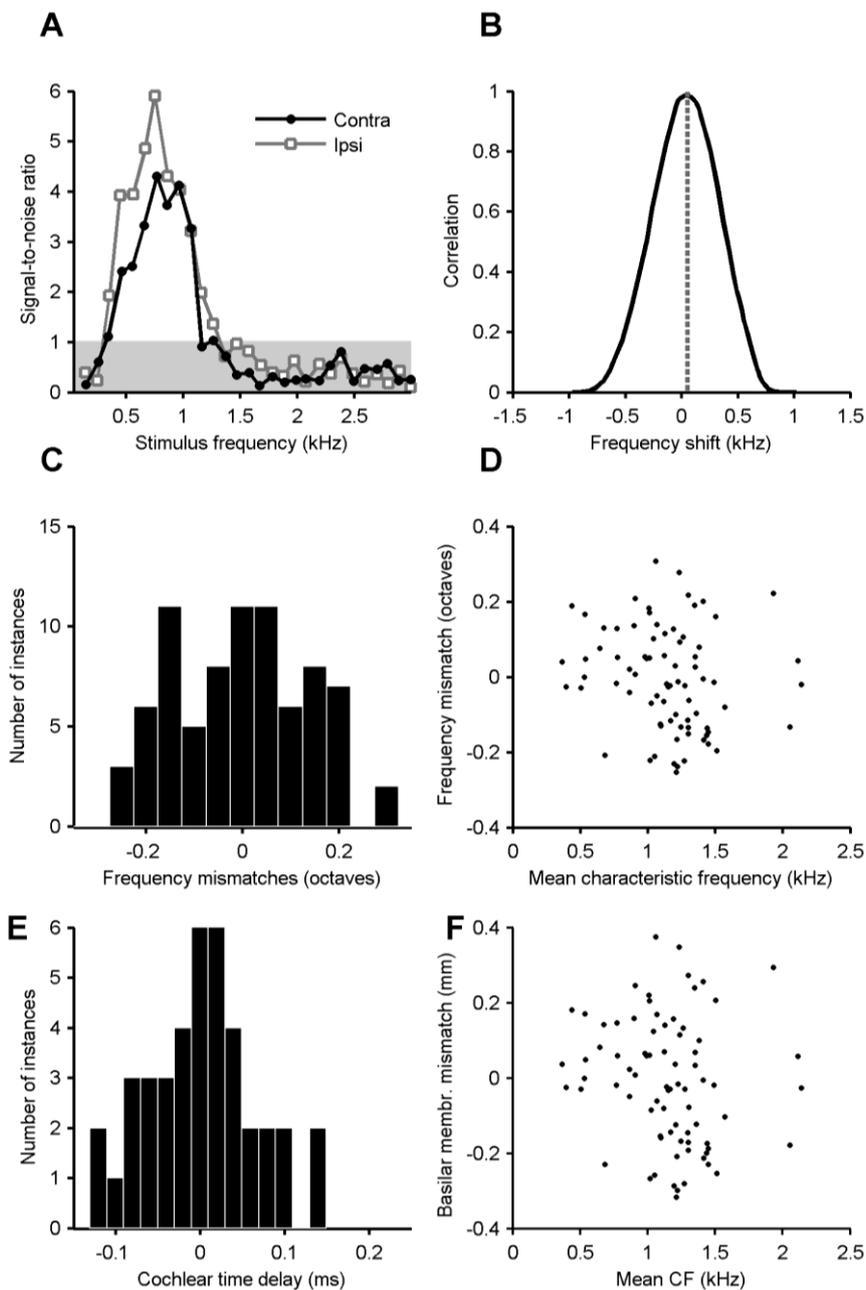


Figure 6. Frequency mismatch estimation.

(A) MSO neuron's frequency responses to contralateral (filled circles) and ipsilateral (open squares) stimulation at 10 dB SPL, estimated from the Fourier spectrum of the response waveform (cf. Figure 5A). Grey area indicates the noise floor. (B) Normalized cross-correlation function of the two frequency responses shown in A. The grey vertical bar indicates the peak of the curve, revealing that the contralateral ear has an estimated 56-Hz higher CF. The data for A and B were taken from the same cell as in Figure 5. (C) Histogram of characteristic frequency mismatches between ipsi- and contralateral stimulation ($N = 78$). (D) Relation between interaural CF mismatches and mean CF of both ears for all MSO cells ($N = 78$; $r = -0.17$; $p = 0.13$). (E) Distribution of cochlear time delays calculated from measured frequency mismatches ($N = 78$). (F) Relation between basilar membrane mismatches, calculated using Müller (1996), and CF ($N = 78$; $r = -0.10$; $p = 0.35$).

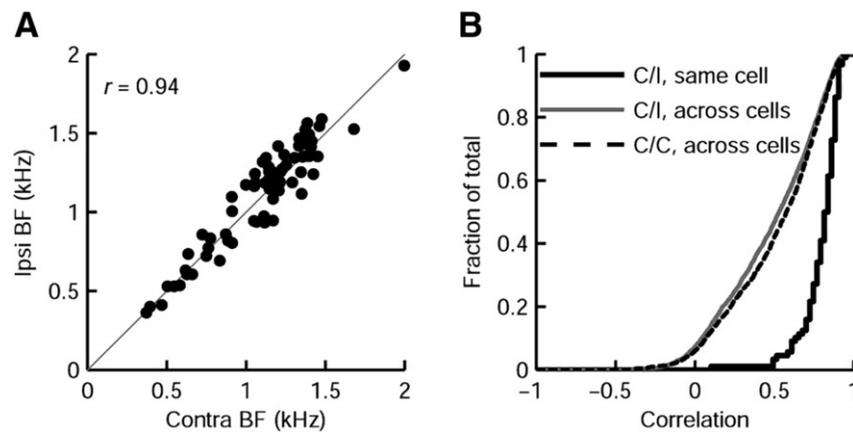


Figure 7. Correlation between monaural frequency tuning within and across MSO cells.

(A) Relation between contra- and ipsilateral BFs of MSO neurons at a stimulus intensity of 30 dB SPL ($N = 69$; $r = 0.94$; $p < 0.0001$). Seventeen neurons were not included as they were not sensitive to the 30 dB SPL stimulus. Grey line indicates identity. (B) Cumulative histogram plots of the normalized correlation coefficients between monaural receptive fields. Thick black line: contra- and ipsilateral receptive fields from the same MSO neuron ('C/I, same cell', $N = 84$). Grey line: contra- and ipsilateral fields from all pairs of MSO cells ('C/I, across cells'). Broken line: contralateral receptive fields between all pairs of MSO cells ('C/C, across cells').

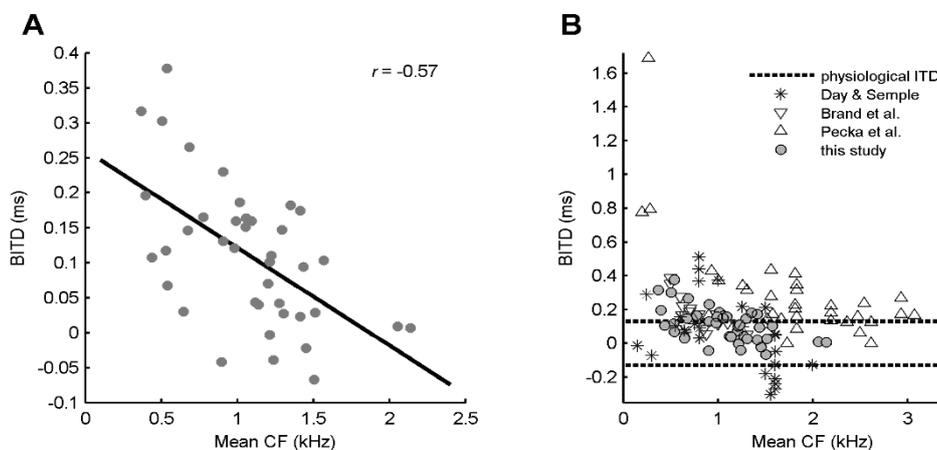


Figure 8. Inverse relation between BITD and mean CF.

(A) Cells tuned to low CFs tend to have more positive BITDs; black line shows linear regression ($N = 40$; $r = -0.57$; $p = 0.0004$). (B) Comparison of BITD and mean CF correlation in Mongolian gerbils for four studies (and their stimulus): Day and Semple (2011) (binaural beat tone stimuli at BF), Brand et al. (2002) (FM tones at BF), Pecka et al. (2008) (tones at BF) and this study (the same data as in A). Dashed lines indicate gerbil's physiological ITD range.

lation between CF mismatch (in octaves) and average CF for the same neurons ($r = -0.07$; $p = 0.53$). Similar results were obtained when frequency mismatches were translated into differences in basilar membrane location (Figure 6E, F; $r = -0.10$; $p = 0.35$; (Müller 1996).

BFs from ipsi- and contralateral inputs were well correlated ($r = 0.95$; Figure 7A). On average, a high correlation was also observed when the entire ipsi- and contralateral receptive fields were correlated for each neuron, yielding an average r of 0.81 ± 0.13 ('native CI', Figure 7B). For comparison we also correlated random pairs of ipsi- and contralateral ears ('random CI') and random pairs of contralateral ears ('random CC'), which yielded on average a much lower correlation of $r = 0.48 \pm 0.30$ and $r = 0.51 \pm 0.29$, respectively (Figure 7B). We thus conclude that ipsi- and contralateral receptive fields are generally similar within MSO neurons, which is in general agreement with earlier reports measuring CF or BF (Goldberg and Brown 1969; Guinan et al. 1972; Moushegian et al. 1964a) and with our earlier work in which we used mostly intense tones in a small number of cells (van der Heijden et al. 2013).

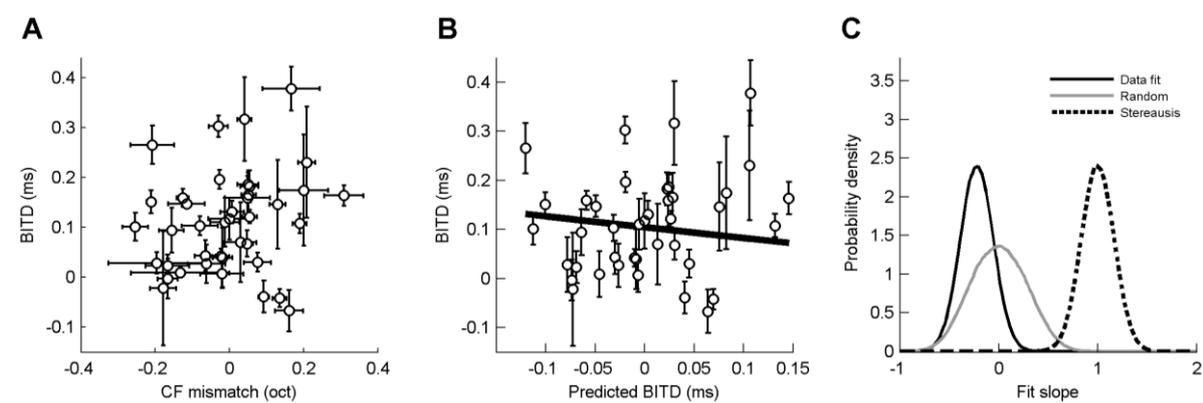


Figure 9. Relation between BITDs and frequency tuning mismatches of MSO neurons.

CF mismatches were determined using cross-correlation of responses to zwuis stimuli (Figure 6). (A) Relation between BITD and CF mismatch ($N = 40$). Bars indicate standard deviations. (B) Relation between BITD and predicted BITD. Predicted BITD was obtained from the relation between the travelling wavelength and CF (Figure 1A). Bars indicate combined standard deviation in BITD and predicted BITD. Line shows linear regression (slope -0.22 ms/ms; $r = -0.012$; $N = 40$). (C) Results of bootstrap analysis of the slope of the regression line of the relation between BITD and predicted BITD. Grey line ('random') shows distribution of slopes when BITD values were scrambled; black line ('data fit') shows fit slopes when data points were drawn from a distribution with the same mean as the measured BITD and its combined error; red line ('stereausis') shows distribution of fit slopes when data points were drawn from a distribution with the same mean as the predicted BITD and the combined error.

Relation between frequency and ITD tuning

We next compared ITD tuning and frequency tuning in the same cells. BITD and mean CF were inversely correlated (Figure 8A), in agreement with earlier work (Figure 8B). The estimates for the mismatch in monaural frequency tuning allowed us to test a key prediction of the stereausis theory, which is to investigate whether the frequency mismatch can predict ITD tuning. For binaural tuning we used BITDs obtained from wideband rITDs in order to have zwuis stimuli for both estimates. CF mismatches were converted into octave differences. No obvious correlation was observed ($N = 40$; Figure 9A). We next converted the CF mismatches to predicted BITDs based on the frequency-place map of the gerbil cochlea (Müller 1996) and estimates of propagation speed and wavelength of the traveling wave in the apex of the cochlea for several species (Figure 1; see Methods). Note that the predicted BITDs have no bias to either ipsi- or contralateral leading values (Figure 6E), in agreement with the lack of asymmetry in the distribution of CF mismatches (Figure 6C). No obvious correlation was observed between the observed and predicted BITDs ($r = -0.012$; slope = -0.22 ms/ms; $N = 40$); to avoid the nonlinearities associated with fitting bivariate regression when both variables have errors (Buonaccorsi 2010), the error estimates for both BITD and predicted BITD were combined into a single value. To get an estimate for the reliability of this conclusion that takes into account the number of cells and the precision of the individual measurements, we did a bootstrap analysis. To get an error for the estimate of the slope of the line fit, the individual BITD values were drawn from a distribution with the same mean and combined error as the original measurement; the resulting slope was -0.22 ± 0.17 ms/ms; if the association between the BITD and the predicted BITD was randomized, the average slope of the line fit became 0 ± 0.28 ms/ms; if individual values were drawn from a distribution with mean identical to the predicted BITD, conform the stereausis prediction, with combined error obtained from the measurement, the slope became 1.0 ± 0.17 (Figure 9C). From this we conclude that the measured slope was well inside the range of slopes that could be expected if the association between BITD and frequency mismatch were random, and well outside the range of slopes that could be expected if the stereausis prediction was accurate. Similar results were obtained when BITDs were first detrended for the inverse relation between BITD and mean CF illustrated in Figure 8A (*results not shown*).

Discussion

We tested three predictions of the stereausis theory. First, as predicted by the stereausis theory, we found evidence that tuning for bilateral electrical stimulation, in which the cochlea is presumably bypassed, was different from ITD tuning for auditory stimuli. Second, we found that even though frequency tuning for ipsi- and contralateral sounds were generally similar, the interaural differences were large enough to create substantial cochlear disparities in many cells. Most importantly, however, we did not find evidence for a direct correlation between BITDs and CF mismatches, even though the accuracy of our methods was sufficient to detect such a correlation in a scenario in which the mismatches were the dominant source of internal delays. We therefore failed to obtain critical support for the stereausis theory, suggesting that axonal delays are more important for determining internal delay than cochlear disparities.

Spatial tuning

A majority of neurons had a positive bias (contralateral leading) in BITDs, in agreement with previous results (Figure 8B). We observed that more than half of neurons had a BITD within the ecological range, similar to earlier work in gerbil and in cat (Day and Semple 2011; Yin and Chan 1990). In contrast, the observed distribution of BITDs was quite different from two other earlier studies in gerbil, where only about 20% of BITDs fell within the ecological range (Figure 8B; (Brand et al. 2002; Pecka et al. 2008). One possible cause for the difference between our results and some of the earlier results is that we used wideband stimuli at relatively low intensity. ITD tuning for tones can be quite complex in the gerbil MSO (Day and Semple 2011; van der Heijden et al. 2013), and especially at frequencies away from CF, or at very high intensities, preferred ITDs to tone stimuli can be quite variable. Most physiological sounds, however, are wideband, and sound intensities above 60 dB SPL are uncommon in nature.

The wideband methods we employed did not allow to take onset responses into account. However, for humans, ongoing ITDs of low frequencies are the main source of information to detect the sound source location in the horizontal plane (Wightman and Kistler 1992).

We did not target a specific area within the MSO, but since BITDs do not vary systematically along the rostrocaudal axis in the gerbil (Franken et al. 2015), it seems unlikely that our conclusions would depend on the exact rostrocaudal location within the MSO.

Electrical stimulation

To separate the delay caused by the traveling wave in the cochlea from the retrocochlear delay, we successfully obtained in three cells an estimate for BITD based on electrical stimulation at the round window in addition to the sound-evoked BITD. In each case we observed a difference of about 0.2 ms for both estimates. This illustrates that cochlear delays can contribute to the overall internal delay, as predicted by the stereausis hypothesis, but that electrical BITD does not predict sound-evoked BITD well, in agreement with the lack of a correlation of the sound-evoked BITD with the predicted BITD based on frequency tuning mismatch.

Our approach of electrically stimulating the round window had the advantage that it preserved hearing, thus allowing a comparison between sound-evoked ITD tuning and electrical ITD tuning. In larger animals such as guinea pigs or cats it has been shown that sound-evoked responses can be largely preserved following cochlear implant insertion into the basal cochlea (McAnally et al. 1997; Miller et al. 2006; Sato et al. 2016; van den Honert and Stypulkowski 1984), but this is a delicate procedure that would still not allow us to directly stimulate the apical cochlea, which supplies the low-frequency afferents involved in ITD tuning. In contrast, monopolar stimulation near the round window can excite the whole cochlea and thresholds are essentially independent of CF (Hartmann et al. 1984; Moxon 1971; van den Honert and Stypulkowski 1984; 1987). Our experiments thus complement earlier studies in slices showing that it is possible to measure an electrically evoked BITD in MSO neurons (Jercog et al. 2010; Roberts et al. 2013).

In all experiments peaks at different latencies were observed at the higher stimulation intensities. The underlying mechanism remains uncertain in our experiments, but is unlikely to involve a traveling wave, since this would be expected to induce ringing, and periodic responses with an interval determined by CF, similar to the response to a click stimulus (Goblick and Pfeiffer 1969; Moxon 1971; Recio and Rhode 2000; van den Honert and

Stypulkowski 1984; 1987). In addition, the minimum sound-evoked latencies were clearly longer than the electrical latencies in the same experiments. We therefore consider it likely that the traveling wave was bypassed by the electrical stimulation. Our observation that in each of the three cells there was a clear difference in the BITD for electrical and auditory stimulation is thus compatible with a contribution of cochlear delays to the overall internal delay, as predicted by the stereausis hypothesis.

Stereausis hypothesis

The ability to measure both ITD tuning and monaural frequency tuning in a large number of cells allowed us to test some of the predictions of the stereausis hypothesis. We used differences in CF to estimate differences in cochlear delay. Since we used wide-band stimuli at moderate sound levels, a substantial change in cochlear disparities compared to the threshold measurements seems unlikely to occur during our BITD measurements. For the positive bias (contralateral leading) in BITDs to originate from a difference in tuning, it would be needed that, on average, contralateral input would be tuned to lower frequencies than ipsilateral input, thus creating a larger cochlear delay for contralateral sounds. This is not what was observed. Neither did we observe evidence for the opposite, i.e. lower BFs in response to ipsi- than contralateral stimulation, as was observed in gerbil IC (Semple and Kitzes 1985).

To convert CF mismatches to predicted delays (Figure 1) we used data from the literature. Probably the biggest source of error in the conversion is the smaller cochlear length of the gerbil compared to cat, chinchilla and guinea pig (Greenwood 1990; Liberman 1982; Müller 1996). Note, however, that if we had scaled down the apical wavelength for the gerbil accordingly, the predicted BITDs would be 50-100% larger, making the expected size of the stereausis effects correspondingly larger. Previous assessments of stereausis (Day and Semple 2011; Shamma et al. 1989) were based on cochlear models (Holmes and Cole 1984; Tan and Carney 2003) rather than data, predicting an even larger contribution of CF mismatches to ITD tuning. Another possible source of error in the BITD predictions are the individual differences in cochlear length (standard deviation/mean ~5%; (Müller 1996; Plassmann et al. 1987), whereas length differences between left and right cochlea are small (Bohne and Carr 1979).

Predicted and measured delays were not significantly correlated. Our data thus agree with earlier tests in low frequency units of the nucleus laminaris of the barn owl (Fischer and Peña 2009; Peña et al. 2001) and alligator (Carr et al. 2009) and in the inferior colliculus of the barn owl (Singheiser et al. 2010). We conclude that for individual MSO cells, frequency mismatches are expected to make a substantial contribution to their internal delay, but that they are not the dominant determinant of BITD on a population level.

The origin of internal delays (Jeffress revisited)

Since our data do not support a role for a systematic contribution of frequency mismatches, the most likely remaining mechanism to create an internal delay is the presence of differences in the time it takes signals to travel from the cochlear nucleus to the MSO neurons. As SBCs typically innervate both ipsi- and contralateral MSO neurons (Thompson and Schofield 2000), it is hard to see how a systematic internal delay could be created in the time it takes to travel from cochlea to cochlear nucleus in the absence of a systematic shift in frequency tuning. In recent years several alternative theories to the classical Jeffress' model (Jeffress 1948) have been put forward that focus on delays created within the MSO itself, including a role for well-timed inhibition (Brand et al. 2002; Myoga et al. 2014; Pecka et al. 2008) or for asymmetric EPSPs (Jercog et al. 2010), but in recent studies these alternatives have not received much support (Day and Semple 2011; Franken et al. 2015; Roberts et al. 2013; van der Heijden et al. 2013; Zhou et al. 2005). We cannot exclude a role for intrinsic conductances in creating delays within the MSO, even though the reported effects might be smaller for wideband, low intensity stimuli as used here, than for high-intensity, low frequency tones that were used in Franken et al. (2015).

This leaves as the most likely possibility –by exclusion– the option that axonal delay lines are responsible for the internal delay, as originally proposed by Jeffress (1948). A direct test of the branching patterns of the axons of SBCs indicated that they cannot account for the frequency-dependent distribution of best delays in the cat (Karino et al. 2011). However, in addition to axonal length, differences in axonal conduction velocities can also contribute to the internal delay, as shown for birds (Seidl et al. 2014; Seidl et al. 2010). Some evidence for differences in conduction velocities for mammalian auditory brainstem axons has indeed

been found recently (Ford et al. 2015; Seidl and Rubel 2016). A combination of anatomical and physiological studies would thus allow to further test the Jeffress hypothesis.

Acknowledgments

This work was supported by the Dutch Fund for Economic Structure Reinforcement (FES, 0908 'NeuroBasic PharmaPhenomics project'). The authors declare no competing financial interests.

Chapter 5

DISCUSSION

Revisiting Jeffress

One of the most impactful theoretical models in auditory neuroscience was proposed by Lloyd A. Jeffress in 1948 (Jeffress 1948). In his publication Jeffress proposed a simple neuronal network that could perform sound localization. This network is composed of two groups of inputs converging from the opposing sides onto binaural neurons, which are sensitive to the timing of the input signal arrival. His theory consists of three core concepts:

Pillar 1: the computational units (i.e. the neurons) act as coincidence detectors with regard to the inputs originating from both ears.

Pillar 2: the network can convey timing information about the inputs.

Pillar 3: inputs converging onto the computational units from opposite sides are arranged in such a way that the responses of the computational units form a spatial gradient.

Below, I will discuss these three pillars of the Jeffress model in the context of recent theoretical and experimental work, including the work described in this thesis.

Pillar 1: Coincidence detection

From models to experiments

Goldberg and Brown (1969) showed experimentally that the MSO neurons do act as coincidence detectors. They found that each neuron fired preferentially at a certain ITD, the 'best' ITD (BITD), and that this BITD can be predicted from the firing response to monaural tones. They often observed binaural facilitation: at the BITD the binaural rates were often higher than the sum of the monaural rates. The output rate during unfavorable delays could drop below monaural stimulation firing rates. These findings were confirmed in later experiments (Batra et al. 1997b; Crow et al. 1978; Moushegian et al. 1975; Spitzer and Semple 1995; Yin and Chan 1990), thus strongly supporting the first pillar of the Jeffress' hypothesis, coincidence detection. An early modeling paper already showed that these findings could be successfully reproduced by a remarkably simple coincidence model, in which the binaural

neurons lacked dendrites and inhibitory inputs (Colburn et al. 1990). In this model monaural coincidences were limited by setting the action potential threshold higher than the maximal monaural input. The high threshold in combination with an exponential relaxation of the membrane potential allowed for coincidence detection between inputs originating from both ears. The model was, thus, largely implemented as an AND gate, without a possibility for monaural coincidences. Despite its simplicity, the model could reproduce key aspects of the Goldberg and Brown (1969) data remarkably well, including the finding that firing rates at the worst ITD were typically lower than during monaural stimulation. The model thus showed that synaptic inhibition was not necessary for the low rates observed during the worst ITD and a follow-up point-neuron model by Han and Colburn (1993) found no difference in MSO performance when comparing excitatory-only and excitatory and inhibitory input cases. Instead, the absence of input from the other, out-of-phase ear, was sufficient to explain the low firing rates at the worst ITD (Colburn et al. 1990). In Chapter 2, experimental evidence supporting this idea was obtained: the response variance to binaural beat stimuli tended to drop below the variance of spontaneous activity at the worst ITD.

In later simulations, it was shown that with passive dendrites, segregation of inputs from different ears to different dendrites limits local input saturation due to a reduction of driving force, thus favoring binaural (interdendritic) summation over monaural (intradendritic) summation (Agmon-Snir et al. 1998; Dasika et al. 2007). The reduction of driving force is especially apparent for thin dendrites, with high local input resistance. These simulations also showed the need for relatively compact dendrites in order to preserve timing information. A corollary of the results of Agmon-Snir et al. (1998) was that a non-stimulated dendrite can act as a sink for the other dendrite. As we observed a lack of impact of the activity in one ear on the size of the responses to the other ear using stimulation with binaural beats (Chapter 2), and we could predict the binaural responses from the monaural responses (Chapter 3), we concluded that this sink effect cannot be very strong. However, for a full test of the relative efficacy of monaural vs. binaural input summation, a much better control over inputs would be needed, or the ability to directly measure dendritic potentials, as recently done in slice recordings (Winters et al. 2017). Experimentally, we found that the nonlinear relation between EPSP size and AP probability provided a good explanation for the observed binaural facilitation. This nonlinear relationship is a general feature of neurons (Silver 2010), and in

combination with the linear summation of the EPSPs provided a very simple scheme allowing the MSO neurons to act as a coincidence detector (Chapters 2 and 3).

Cross-correlation

An additional test for coincidence detection of MSO neurons comes from experiments in which the monaural inputs were cross-correlated to test to what extent this matches the binaural responses. The first experimental test for cross-correlation in MSO was done by Yin and Chan (1990) (Figure 12) who showed that the ITD sensitivity of MSO in the cat was indeed comparable to the predicted ITD response obtained by cross-correlating monaural spike trains generated by frozen noise sound stimuli. However, while the measured and predicted BITD agreed well, the overall spike rates and secondary peak locations only partially matched the predictions. Batra et al. (1997b) performed a comparable test of linearity in rabbit superior olivary complex, and found that binaural vector strength matched the product of monaural VSs, as predicted for a cross-correlation operation.

Deviations from the predictions of a simple cross-correlation mechanism were first found by Batra and Yin (2004) in their modelling study. Using data from Yin and Chan (1990) and Batra et al. (1997a; 1997b) and the simple model of Colburn et al. (1990) discussed earlier, they found that monaural VSs derived from binaural-beat stimuli predicted the interaural-phase VS better than those from monaural stimuli. During monaural stimulus presentation, the other ear contributes only spontaneous synaptic inputs. During the binaural beat stimulus that same side will contribute a higher input rate (albeit phase-locked to a different stimulus frequency than the other side). Based on their model, they conclude that the VS measured during monaural stimulation includes APs generated by monaural coincidences, whereas their impact will be smaller during binaural stimulation. These findings led to the conclusion that MSO output produces a degraded version of monaural input cross-correlation. It is worth noting that the studies above used action potentials from MSO for monaural inputs and not subthreshold events, as well as that the stimuli were tones and binaural beats, but not wideband stimuli.

Another study that found deviations from a simple cross-correlation scheme was done by Franken et al. (2015). They employed two rITDf prediction methods: a) responses to monaural

stimuli were corrected for membrane potential, summed and thresholded to predict the elicitation of action potentials; b) monaural response EPSP period histograms were cross-correlated to estimate the predicted rITDf. Both methods yielded the same result – predicted and measured rITD functions were offset by more than 100 μ s in half of their datasets. Upon closer examination of membrane potential (V_m) during summation they found that ITD tuning was affected by ITD-dependent V_m changes preceding the coincidence. That is, the history of MSO membrane potential shapes ITD tuning and cross-correlation cannot capture this effect.

Experimental evidence that the linear sum of monaural responses is very close to the binaural response to the same stimuli (Chapter 2) points to a simple linear somatic summation of dendritic inputs within MSO cells, an operation which is compatible with a cross-correlation. Indeed, our findings reveal that cross-correlation can successfully account for the MSO operation as a coincidence detector. We utilized wideband stimuli at maximum 40 dB SPL above the hearing threshold. In contrast, Franken et al. (2015) evoked responses with binaural beats presented at 70 dB SPL at frequencies that were typically far from CF. Under these conditions, peak splitting becomes important (Chapter 2), and this may account to some extent for their results. We performed cross-correlation of monaural responses to wideband stimuli and compared the predicted rITDf with AP-based rITDf from binaural wideband stimulation. The similarity of prediction and data was measured through a comparison of parameters of the Gabor function which was used to fit both prediction and data. The temporal parameters – phase and frequency – of the two functions matched accurately, while this was less the case for the envelope features. Predicted and measured BITDs were offset by at most 50 μ s, which for the gerbil falls within its minimum resolvable angle of 25°.

In summary, it is clear that there are many factors that can influence MSO operation in coincidence detection: dendritic segregation, somatic inhibition, complex interactions between various ionic currents and somatic input integration. However, experimental evidence suggests that MSO seems to process these inputs in a linear fashion, as the key features of rITDf – BITD, oscillation frequency and phase – can be predicted from cross-correlation of monaural subthreshold inputs evoked with moderate sound levels without the need to include inhibition as a prominent player. We cannot exclude that interaction between excitatory and inhibitory conductances or the recent history of membrane potential contribute to the coincidence detection mechanism. In fact, these might contribute to the

discrepancies we saw in envelope features of rITDs. However, the overall MSO task, sensitivity to interaural differences, does seem to rely on linear processes.

Pillar 2: Internal delay

The second important component of the Jeffress' hypothesis is that the neuronal network that decodes information about sound localization has to be able to encode differences in arrival time at both ears. As discussed in the previous section, the neurons are known to be coincidence detectors, i.e. these cells are driven most effectively when inputs from both ears summate at the soma. This means that unless BITD is 0 μ s, there must be a difference in the travel time for signals from both ears to the place where they effectively come together to compensate for the difference in the arrival time at either ear. The anatomical location for this so-called internal delay has been debated. In the original paper by Jeffress (1948), the known "slow rate of conduction" of the axons inspired a mechanism involving different path lengths corresponding to different travel times. We investigated two alternatives, one postulates a difference in travel time within the cochlea ('stereausis'), the other within the MSO. For the latter, the most prominent invocation was based on the inhibitory inputs of the MSO neurons. I will start by discussing some of the implications of the Jeffress model for ITD tuning, then I will discuss the evidence favoring Jeffress' axonal delay model, followed by the well-timed inhibition model and, finally, the stereausis model.

Axonal delay model

The original proposal by Jeffress was that this delay arises from different axonal-dendritic innervation length across the dorso-ventral axis of the MSO. There is ample experimental evidence showing that avian ITD detection is well described by Jeffress' model (Carr and Konishi 1990; Köppl and Carr 2008; Overholt et al. 1992; Seidl et al. 2010; Young and Rubel 1983), and that axonal delay lines are responsible for internal delays. Two predictions arise from a pure axonal delay line mechanism: a) the range of measured BITDs should be compatible with axonal anatomy. b) BITDs should be frequency-independent. Anatomical studies (Beckius et al. 1999; Karino et al. 2011; Seidl and Rubel 2016; Smith et al. 1993) have been discussed in detail in the General Introduction. The main conclusion from these papers is that even though there are aspects of the axonal innervation that are in line with the

Jeffress' hypothesis, the overall axonal organization cannot account for the full range of observed best delays. However, there is evidence that contralateral and ipsilateral collaterals differ in their conduction velocity, which might help to account for best delays.

Interaural phase

Day and Semple (2011) found that within a large group of MSO neurons, BITD depended on stimulus frequency, which we observed in our experiments as well (Figure 4, Chapter 2). An anatomical correlate for this observation is lacking at present since Karino et al. (2011) found no evidence for a difference in axonal branching patterns for low and high CF fibers.

Since the inputs to the MSO show phase-locking (see Introduction) interaural time differences can also be translated to interaural phase differences (IPDs) for tonal stimuli. If the BITD of an MSO neuron is frequency-independent, then for all stimulus frequencies rITD functions will share a common BITD, also known as the characteristic delay (CD; Rose et al. 1966). On the other hand, if there is frequency-dependency in ITD sensitivity, rITD functions will have a common point (CD), which is not at the peak (BITD). This common point is the result of a constant phase delay – characteristic phase (CP) – being present for all the relevant frequencies. We can write the IPD dependence on stimulus frequency (f) as:

$$\text{IPD}(f) = \text{CP} + \text{CD} * f;$$

derivatively,

$$\text{BITD}(f) = \text{CD} + \text{CP}/f.$$

In theory, internal delay can arise due to both time and phase delays in the inputs of the MSO. However, axonal delays (Jeffress' prediction) imply a pure time delay mechanism, since the speed at which an AP travels within an axon is in principle independent of how it was evoked. This results in CP = 0 for all frequencies. In this case $\text{IPD}(f) = \text{CD} * f$; the relationship between interaural phase and frequency is linear, with intercept (CP) zero and slope equal to CD, BITD does not depend on stimulus frequency. On the other extreme, for pure phase delays CD = 0 and IPD = CP for all frequencies; BITD and frequency relationship is hyperbolic. See Vonderschen and Wagner (2014) for more detailed discussions of this topic.

Several studies found that MSO exhibits a mixture of both types of delays, showing the presence of non-zero CP (Batra et al. (1997a); Day and Semple (2011); Pecka et al. (2008); Yin and Chan (1990); this thesis, Chapter 3). Importantly, pure time delays, as proposed by Jeffress (1948), cannot account for this phenomenon. Furthermore, the phase-frequency relationship in some neurons deviates from a linear relationship (Day and Semple (2011); this thesis, Chapter 3). These deviations from the Jeffress' model have led to proposals for alternative mechanisms to create internal delay. Two prominent alternatives for axonal delay lines are the well-timed inhibition and the stereausis model.

Well-timed inhibition as a source for internal delay

One of the most hotly disputed topics of MSO operation is the role of glycinergic inhibition in coincidence detection. Two bilateral somatic inhibitory inputs originating from the ipsilateral MNTB and LNTB have been anatomically identified and described in detail (Adams and Mugnaini 1990; Clark 1969a; b; Grothe and Sanes 1993; 1994; Perkins 1973). The studies on their physiological role, however, have not been conclusive. As discussed in the previous section, several key aspects of MSO function could be accounted for with high accuracy using models that lacked inhibition.

The first physiological evidence for the role of inhibition in mammalian MSO was presented by Grothe and Sanes (1993; 1994). Using electrical stimulation of MSO afferents in slice recordings they showed that at high stimulus intensities and short time delay, responses to one 'ear' (unilateral) stimulation were suppressed by the stimulation of the other side; this effect could be explained by glycinergic inhibition, originating from either MSO or LNTB. It disappeared when the inhibition was suppressed by strychnine application. These studies were followed by in vivo experiments (Brand et al. 2002; Pecka et al. 2008). Application of strychnine (blocking inhibition) through iontophoresis onto MSO cells resulted in an increased MSO firing rate and BITD shift towards zero, an effect from which MSO cells recovered after several minutes. Similar results were observed when enhancing inhibition by tonic glycine application. The authors proposed a model featuring well-timed inhibition delaying spike timing, to shift the BITD to more positive (contra leading) values, analogous to an earlier model by Batra et al. (1997a). In support of this mechanism, there were models which incorporated well-timed inhibition to explain genesis of internal delay (Leibold 2010; Leibold

and van Hemmen 2005; Svirskis et al. 2003). However, an attempt by Franken et al. (2015) to replicate the in vivo experiment showed that iontophoretic application of strychnine not only shifted BITD towards zero, but also continued reducing firing rate of the MSO until ITD was insignificant (in less than 15 minutes). This was possibly due to iontophoresis causing high concentrations of strychnine in the nucleus. When strychnine was pressure-applied (to limit its buildup) no significant BITD shift was observed, while the increase in firing rate was still observed. Slice studies showed that at high concentrations, strychnine blocks I_h (Franken et al. 2015), further suggesting that the observed BITD shift was a non-specific effect. Other evidence against this model comes from the work presented in Chapter 2. Both Franken et al. (2015) and we (Chapter 2) did not observe the IPSP preceding the EPSPs proposed in the well-timed inhibition. Moreover, the lack of an effect of stimulation of one ear on the amplitude of the responses to the other ear (Chapter 2) is also not in agreement with a strong effect of inhibition, as the inhibitory inputs are localized at the soma, so their activation should affect both ears. Additional evidence against the well-timed inhibition model comes from modeling studies. Zhou et al. (2005) pointed out that the model by Brand et al. (2002) needed very fast inhibitory events (~ 0.1 ms), while Smith et al. (2000) showed that decay times of IPSCs in MSO were at least ~ 3.9 ms. Conductance clamp slice studies provided experimental evidence for the need for a nonphysiologically fast IPSP time course for the model to work (Roberts et al. 2013). Incorporating fast inhibition in their model, Day and Semple (2011) were unable to account for nonlinearities in the phase-frequency relationship that they saw in their own data. A slice conductance clamp study by Myoga et al. (2014), despite being generally supportive of timed inhibition effects, found that contralateral inhibition alone could account for only a ~ 50 μ s BITD shift, and only together with ipsilateral inhibition the shift could reach ~ 150 μ s. In conclusion, previously observed BITD shifts due to inhibition being blocked were most likely caused by nonspecific effects of strychnine at high concentrations and not the removal of inhibition. Even though it is likely that synaptic inhibition plays a substantial role within the MSO, it is unlikely to shift ITD tuning substantially, as proposed in the well-timed inhibition model.

Cochlear delay model

The cochlear delay model, also referred to as 'stereausis' model, was first proposed by Schroeder (1977) and detailed in a computational model by Shamma et al. (1989). It proposes

that ipsi- and contralateral afferents to MSO neurons originate from different locations in the cochlea; the resulting difference in travel times within the cochlea is proposed to make a significant, systematic contribution to the internal delay. Its main prediction is that there is a systematic relation between ipsi- and contralateral frequency tuning mismatches and BITD. The first experimental evidence for cochlear mismatches has been shown by Yin and Kuwada (1983) in the IC of the cat, followed by indirect observation of the phenomena in the MSO by Yin and Chan (1990) and SOC (Batra et al. 1997a; Spitzer and Semple 1995). Modelling studies have also confirmed that the stereausis model could account for BITD distribution in mammalian MSO (Bonham and Lewis 1999; Joris et al. 2006). A recent publication by Day and Semple (2011) found that modeled coincidence detection with cochlear delays could successfully reproduce their experimentally observed frequency-dependent delays due to small CF mismatch giving rise to large interaural cochlear delays.

We tested the stereausis hypothesis experimentally *in vivo* (Chapter 4) utilizing auditory and electrical stimulation. To bypass frequency dependent cochlear disparities, in some experiments we electrically stimulated the round windows of both cochleae. Electrically-evoked and acoustically-evoked rITD functions were not similar and showed a BITD mismatch of ~ 0.2 ms in either direction. This observation supported possible cochlear delay role in ITD tuning. From the responses to auditory monaural low SPL wideband stimuli, we showed that there is no systematic frequency mismatch bias towards either ear. As BITDs are biased towards the contralateral ear, the stereausis model predicts that the contralateral ear should be tuned to lower frequencies than the ipsilateral ear. Most importantly, there was no significant correlation between BITDs and CF mismatches in our data (Figure 9, Chapter 4), leading to the conclusion that while cochlear mismatches can substantially contribute to internal delay generation, they are not the dominant mechanism behind it.

Pillar 3: Map of BITDs

The third pillar of Jeffress' model implies the presence of anatomical BD map due to systematic delay line organization. Early work, based on a limited number of observations in the cat MSO, did hint towards a relation between rostrocaudal location and BITD (Yin and Chan 1990). A more extensive test of the presence of a spatial map for BITD was done in the gerbil MSO by Franken et al. (2015) (Figure 2e), who did not find evidence for a relation between rostrocaudal location of the MSO and BITD. Since the MSO is small in the medio-lateral direction and since the dorsoventral axis is devoted to frequency (Guinan et al. (1972), Chapter 2), this suggests that the mammalian MSO does not contain a spatial map of BITDs.

Jeffress Revisited

The evidence presented above shows that >70 years after its publication, the elegant Jeffress proposal continues to be an inspiration for experimental and modeling work into the sound localization circuitry. Of its three pillars, only coincidence detection has been well documented. My own work on the interaction between subthreshold events from both ears provides further evidence for this, and the nonlinear relationship between EPSP size and AP generation may provide an explanation for some of the uncertainties that have surrounded tests of coincidence detection in the past.

A specific set of ion channels – most importantly AMPA-type GluR channels, HCN, LVAPC and VGSC – enable MSO neurons to act as timing specialists. The interplay of these channels allows interactions of subthreshold events at the soma to be remarkably linear; however, we found that MSO firing rate depends supralinearly on subthreshold potential. In MSO neurons action potential backpropagation is quite limited (Scott et al. 2007), limiting the interaction of APs and EPSPs within dendrites. Two prominent somatic inhibitory inputs, relayed through fast synaptic pathways, have been proposed to delay AP initiation, but apart from slice and modelling studies, little convincing evidence has been shown to support this mechanism. Most of the evidence points to MSO acting as a cross-correlator, summing the converging (possibly modified but not dramatically reshaped by inhibition) inputs in a linear fashion. While some deviations from this linearity are possible, they do not seem to affect the key computational task of this nucleus – ITD sensitivity. We and Franken et al. (2015) also

addressed other mechanisms that might regulate ITD tuning at the level of the MSO itself, but did not find supporting evidence for them.

A recent proposal by van der Heijden (2018) suggests that inner hair cells could also account for internal phase shifts, but this hypothesis has not yet been experimentally tested.

Cochlear mismatches, as proposed in the stereausis model, could also feasibly contribute to the internal delay, yet it does not seem to be the dominant source of it.

This leaves, at present, the axonal delay lines envisioned by Jeffress. Despite some anatomical evidence being present in favor of these delay lines, the observation that MSO ITD tuning depends on stimulus frequency remains an observation that is not compatible with a pure axonal delay line mechanism.

Future directions

The biggest hurdle in MSO experiments arguably comes from its location deep in the brainstem. Because of the high degree of myelination and the unfavorable location, optical imaging techniques for MSO *in vivo* activity measurements do not yet seem feasible, neither from ventral nor from dorsal side, leaving the need for electrophysiological techniques. Experiments in Chapters 2-4 allowed us to record sub- and suprathreshold events from isolated MSO units under wideband auditory stimulation, providing a chance to investigate the operation of the MSO in higher detail than in previous *in vivo* studies. However, monaural subthreshold responses had to be inferred from the somatic recordings and we had to rely on juxtacellular configuration as the whole cell-patching success rate was very low. Though we showed that juxtacellular recordings can successfully capture subthreshold events, for very precise quantification of synaptic events in the MSO, a larger dataset of whole cell *in vivo* recordings is needed, ideally obtained at high recording bandwidth.

Early MSO studies used relatively simple, ‘unnatural’, auditory stimuli – clicks, pure tones or frozen noise. Though providing insight in basic operational features of MSO, they were limited in the questions they could tackle. We used a wideband stimulus, a mixture of 30 frequencies, in our experiments, and confirmed that input summation linearity works remarkably well. The most convincing approach would be to engage the early binaural apparatus using segments of natural environmental sounds as stimuli. In our experiments we presented the stimulus close to hearing threshold to study responses at the characteristic frequency of each neuron. However, naturally occurring sound cues vary in their sound intensities and spectral contents. Our unpublished results indicate that ITD tuning is robust against sound intensity variations, even though increasing sound levels recruit more input fibers as well as shift the frequency preferences for both ears. The question how MSO retains input matching under these conditions is a very important but as of yet largely unaddressed question, which merits further studies.

The role of inhibition in the MSO, though extensively studied, is still debated. Although Franken et al. (2015) showed that pharmacological blockage of inhibition does not affect ITD tuning, this approach could be used to investigate changes in subthreshold event integration.

In Chapter 3, we showed that some features of ITD functions could not be predicted by cross-correlation, possibly some aspects of ITD modulation are mediated by inhibitory events. Other than MNTB, LNTB as well is thought to provide inhibitory inputs to the MSO, though recently its role in MSO inhibition has been challenged by Franken et al. (2016) albeit with a small sample size. The functional roles of inputs from MNTB and LNTB are yet to be definitively investigated, and require larger sample anatomical as well as pharmacological studies.

The lack of anatomical evidence for axonal delay lines gave rise to alternative theories to explain internal delays. However, recent studies, as well as our findings, did not find strong evidence to support these alternatives, inviting a revision of the original Jeffress' postulate. As recently found by Seidl and Rubel (2016), conduction velocity regulation differs between inputs from both ears in gerbil and, despite no clear anatomical arrangement, the cochlear-to-MSO travelling time differences between the ears can be the long sought internal delay mechanism. These anatomical findings are yet to be correlated with physiological features. One way to test this hypothesis is using rabies virus tracing (Callaway and Luo 2015; Ghanem and Conzelmann 2015). After determining physiological features of an MSO neuron through juxtacellular configuration, the cell would have to be electrophoretically transfected with DNA encoding a receptor for the rabies virus and a fluorescent protein (Kim et al. 2016; Reardon et al. 2016). After an incubation period, and infection with rabies virus, retrogradely labeled inputs to the MSO can be morphologically recovered and quantified. Conduction velocity difference between ipsi- and contralateral projections could be retrieved by recording from SBCs while electrically stimulating both MSO nuclei. After successfully filling the cell with a dye, its morphological features can be recovered through histology. When coupled, the two experiments can relate measured BITDs to conduction velocities and answer the question whether the excitatory inputs to MSO are the underlying source of internal delays. These experiments would, thus, hopefully resolve the remaining controversial pillar of the Jeffress' hypothesis.

Electrophysiological properties of MSO neurons have been described for gerbil before and after hearing onset and their comparison revealed striking developmental changes during this critical period. Virtually no studies so far have focused on functional changes in the gerbil MSO during development *in vivo*. Questions and methods suggested above would yield the most answers when applied on animals of various ages, particularly around the critical hearing

period. Do developmental changes in MSO inputs contribute to input matching and at what time point does the matching occur? Is inhibition onto MSO more prominent before hearing onset and what role does it play during development?

A possible goal of binaural hearing research is not only to explain how sound localization is computed in the brain, but can also be to improve the quality of life for the many people that hear poorly. Bilateral cochlear implant (CI) users can to some extent utilize ILD cues for sound localization, whereas ITD cues are more ambiguous and less accessible (for review on the topic see Laback et al. (2015). ITD sensitivity is important for auditory perceptiveness in many aspects and is much poorer in CI users than in normal hearing people. Recent study on bilateral CIs in unanesthetized animals showed that ITD sensitivity in IC largely depends on carrier pulse rates used in CI processors (Chung et al. 2016). The primary center of binaural convergence, however, is not the IC but MSO. For our cochlear mismatch experiments we made a preparation which allowed us to record from MSO while electrically stimulating round windows on both cochleae. An improved method (Wiegner et al. 2016) could elucidate MSO ITD tuning differences between deafened and normal hearing animals and search for optimal stimulation parameters to recover ITD sensitivity.

References

- Adams JC.** Ascending projections to the inferior colliculus. *J Comp Neurol* 183: 519-538, 1979.
- Adams JC, and Mugnaini E.** Immunocytochemical evidence for inhibitory and disinhibitory circuits in the superior olive. *Hear Res* 49: 281-298, 1990.
- Agmon-Snir H, Carr CE, and Rinzel J.** The role of dendrites in auditory coincidence detection. *Nature* 393: 268-272, 1998.
- Ashida G, and Carr CE.** Sound localization: Jeffress and beyond. *Current Opinion in Neurobiology* 21: 745-751, 2011.
- Ashida G, Funabiki K, and Carr CE.** Theoretical foundations of the sound analog membrane potential that underlies coincidence detection in the barn owl. *Front Comput Neurosci* 7: 151, 2013.
- Barrett TW.** Superposition of binaural influences on single neuron activity in the medial superior olive elicited by electrical stimulation of the osseous spiral laminae. *Brain Res Bull* 1: 209-228, 1976.
- Batra R, Kuwada S, and Fitzpatrick DC.** 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, 1997a.
- Batra R, Kuwada S, and Fitzpatrick DC.** Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol* 78: 1237-1247, 1997b.
- Batra R, and Yin TCT.** Cross correlation by neurons of the medial superior olive: a reexamination. *J Assoc Res Otolaryngol* 5: 238-252, 2004.
- Baumann VJ, Lehnert S, Leibold C, and Koch U.** Tonotopic organization of the hyperpolarization-activated current (I_h) in the mammalian medial superior olive. *Front Neural Circuits* 7: 117, 2013.
- Beckius GE, Batra R, and Oliver DL.** Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci* 19: 3146-3161, 1999.
- Biedenbach MA, and Freeman WJ.** Click-evoked potential map from the superior olivary nucleus. *Am J Physiol* 206: 1408-1414, 1964.
- Bohne BA, and Carr CD.** Location of structurally similar areas in chinchilla cochleas of different lengths. *J Acoust Soc Am* 66: 411-414, 1979.
- Bonham BH, and Lewis ER.** Localization by interaural time difference (ITD): effects of interaural frequency mismatch. *J Acoust Soc Am* 106: 281-290, 1999.
- Brand A, Behrend O, Marquardt T, McAlpine D, and Grothe B.** Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417: 543-547, 2002.
- Brawer JR, and Morest DK.** Relations between auditory nerve endings and cell types in the cat's anteroventral cochlear nucleus seen with the Golgi method and Nomarski optics. *J Comp Neurol* 160: 491-506, 1975.
- Bremen P, and Joris PX.** Axonal recordings from medial superior olive neurons obtained from the lateral lemniscus of the chinchilla (*Chinchilla laniger*). *J Neurosci* 33: 17506-17518, 2013.

- Brill MH, Waxman SG, Moore JW, and Joyner RW.** Conduction velocity and spike configuration in myelinated fibres: computed dependence on internode distance. *J Neurol Neurosurg Psychiatry* 40: 769-774, 1977.
- Brown CH, and May BJ.** Comparative mammalian sound localization. In: *Sound source localization*, edited by Popper AN, and Fay RR. New York: Springer, 2005, p. 124-178.
- Buonaccorsi JP.** *Measurement error. Models, methods, and applications.* CRC press, 2010, p. 437.
- Caird D, and Klinke R.** Processing of binaural stimuli by cat superior olivary complex neurons. *Exp Brain Res* 52: 385-399, 1983.
- Callaway EM, and Luo L.** Monosynaptic circuit tracing with glycoprotein-deleted rabies viruses. *J Neurosci* 35: 8979-8985, 2015.
- Campbell N.** The study of discontinuous phenomena. *Proc Camb Phil Soc* 15: 117-136, 1909.
- Cant NB, and Casseday JH.** Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei. *J Comp Neurol* 247: 457-476, 1986.
- Cant NB, and Hyson RL.** Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil. *Hear Res* 58: 26-34, 1992.
- Cant NB, and Morest DK.** The structural basis for stimulus coding in the cochlear nucleus of the cat. In: *Hearing Science*, edited by Berlin CI. San Diego: College-Hill Press, 1984, p. 374-422.
- Carney LH, Sarkar S, Abrams KS, and Idrobo F.** Sound-localization ability of the Mongolian gerbil (*Meriones unguiculatus*) in a task with a simplified response map. *Hear Res* 275: 89-95, 2011.
- Carr CE, and Konishi M.** A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* 10: 3227-3246, 1990.
- Carr CE, Shah S, McColgan T, Ashida G, Kuokkanen PT, Brill S, Kempter R, and Wagner H.** Maps of interaural delay in the owl's nucleus laminaris. *J Neurophysiol* 114: 1862-1873, 2015.
- Carr CE, Soares D, Smolders J, and Simon JZ.** Detection of interaural time differences in the alligator. *J Neurosci* 29: 7978-7990, 2009.
- Cash S, and Yuste R.** Linear summation of excitatory inputs by CA1 pyramidal neurons. *Neuron* 22: 383-394, 1999.
- Caspary DM, Backoff PM, Finlayson PG, and Palombi PS.** Inhibitory inputs modulate discharge rate within frequency receptive fields of anteroventral cochlear nucleus neurons. *J Neurophysiol* 72: 2124-2133, 1994.
- Casseday JH, and Neff WD.** Auditory localization: role of auditory pathways in brain stem of the cat. *Journal of Neurophysiology* 38: 842-858, 1975.
- Chung Y, Hancock KE, and Delgutte B.** Neural Coding of Interaural Time Differences with Bilateral Cochlear Implants in Unanesthetized Rabbits. *J Neurosci* 36: 5520-5531, 2016.
- Clark GM.** The ultrastructure of nerve endings in the medial superior olive of the cat. *Brain Res* 14: 293-305, 1969a.
- Clark GM.** Vesicle shape versus type of synapse in the nerve endings of the cat medial superior olive. *Brain Res* 15: 548-551, 1969b.
- Clark GM, and Dunlop CW.** Field potentials in the cat medial superior olivary nucleus. *Exp Neurol* 20: 31-42, 1968.
- Colburn HS.** Computational models of binaural processing. In: *Auditory Computation*, edited by Hawkins HL, McMullen TA, Popper AN, and Fay RR. New York: Springer, 1996, p. 332-400.

- Colburn HS, Han Y-A, and Culotta CP.** Coincidence model of MSO responses. *Hear Res* 49: 335-346, 1990.
- Couchman K, Grothe B, and Felmy F.** Functional localization of neurotransmitter receptors and synaptic inputs to mature neurons of the medial superior olive. *J Neurophysiol* 107: 1186-1198, 2012.
- Couchman K, Grothe B, and Felmy F.** Medial superior olivary neurons receive surprisingly few excitatory and inhibitory inputs with balanced strength and short-term dynamics. *J Neurosci* 30: 17111-17121, 2010.
- Crow G, Rupert AL, and Moushegian G.** Phase locking in monaural and binaural medullary neurons: implications for binaural phenomena. *J Acoust Soc Am* 64: 493-501, 1978.
- Dasika VK, White JA, and Colburn HS.** Simple models show the general advantages of dendrites in coincidence detection. *J Neurophysiol* 97: 3449-3459, 2007.
- Day ML, and Semple MN.** Frequency-dependent interaural delays in the medial superior olive: implications for interaural cochlear delays. *Journal of Neurophysiology* 106: 1985-1999, 2011.
- Fesce R.** Stochastic approaches to the study of synaptic function. *Prog Neurobiol* 35: 85-133, 1990.
- Fischer BJ, Christianson GB, and Peña JL.** Cross-correlation in the auditory coincidence detectors of owls. *J Neurosci* 28: 8107-8115, 2008.
- Fischer BJ, and Peña JL.** Bilateral matching of frequency tuning in neural cross-correlators of the owl. *Biol Cybern* 100: 521-531, 2009.
- Fischer BJ, and Seidl AH.** Resolution of interaural time differences in the avian sound localization circuit—a modeling study. *Front Comput Neurosci* 8: 99, 2014.
- Fischer BJ, Steinberg LJ, Fontaine B, Brette R, and Pena JL.** Effect of instantaneous frequency glides on interaural time difference processing by auditory coincidence detectors. *Proc Natl Acad Sci U S A* 108: 18138-18143, 2011.
- Fischl MJ, Combs TD, Klug A, Grothe B, and Burger RM.** Modulation of synaptic input by GABA_B receptors improves coincidence detection for computation of sound location. *J Physiol* 590: 3047-3066, 2012.
- Ford MC, Alexandrova O, Cossell L, Stange-Marten A, Sinclair J, Kopp-Scheinflug C, Pecka M, Attwell D, and Grothe B.** Tuning of Ranvier node and internode properties in myelinated axons to adjust action potential timing. *Nat Commun* 6: 8073, 2015.
- Franken TP, Bremen P, and Joris PX.** Coincidence detection in the medial superior olive: mechanistic implications of an analysis of input spiking patterns. *Front Neural Circuits* 8: 42, 2014.
- Franken TP, Roberts MT, Wei L, Golding NL, and Joris PX.** *In vivo* coincidence detection in mammalian sound localization generates phase delays. *Nat Neurosci* 18: 444-452, 2015.
- Franken TP, Smith PH, and Joris PX.** *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, 2016.
- Freygang WH, Jr., and Frank K.** Extracellular potentials from single spinal motoneurons. *J Gen Physiol* 42: 749-760, 1959.
- Funabiki K, Ashida G, and Konishi M.** Computation of interaural time difference in the owl's coincidence detector neurons. *J Neurosci* 31: 15245-15256, 2011.
- Galambos R, Schwartzkopff J, and Rupert A.** Microelectrode study of superior olivary nuclei. *Am J Physiol* 197: 527-536, 1959.

- Gasparini S, and Magee JC.** State-dependent dendritic computation in hippocampal CA1 pyramidal neurons. *J Neurosci* 26: 2088-2100, 2006.
- Gerstner W, Kempter R, van Hemmen JL, and Wagner H.** A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383: 76-81, 1996.
- Ghanem A, and Conzelmann KK.** G gene-deficient single-round rabies viruses for neuronal circuit analysis. *Virus Res* 216: 41-54, 2015.
- Glendenning KK, Hutson KA, Nudo RJ, and Masterton RB.** Acoustic chiasm II: Anatomical basis of binaurality in lateral superior olive of cat. *Journal of Comparative Neurology* 232: 261-285, 1985.
- Glendenning KK, and Masterton RB.** Acoustic chiasm: efferent projections of the lateral superior olive. *J Neurosci* 3: 1521-1537, 1983.
- Glendenning KK, Masterton RB, Baker BN, and Wenthold RJ.** Acoustic chiasm. III: Nature, distribution, and sources of afferents to the lateral superior olive in the cat. *Journal of Comparative Neurology* 310: 377-400, 1991.
- Goblick TJ, Jr., and Pfeiffer RR.** Time-domain measurements of cochlear nonlinearities using combination click stimuli. *J Acoust Soc Am* 46: 924-938, 1969.
- Goldberg JM, and Brown PB.** Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study. *J Neurophysiol* 31: 639-656, 1968.
- Goldberg JM, and Brown PB.** Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32: 613-636, 1969.
- Goldwyn JH, Mc Laughlin M, Verschooten E, Joris PX, and Rinzel J.** A model of the medial superior olive explains spatiotemporal features of local field potentials. *J Neurosci* 34: 11705-11722, 2014.
- Goldwyn JH, McLaughlin M, Verschooten E, Joris PX, and Rinzel J.** Signatures of somatic inhibition and dendritic excitation in auditory brainstem field potentials. *J Neurosci* 37: 10451-10467, 2017.
- Grau-Serrat V, Carr CE, and Simon JZ.** Modeling coincidence detection in nucleus laminaris. *Biol Cybern* 89: 388-396, 2003.
- Greenwood DD.** A cochlear frequency-position function for several species--29 years later. *J Acoust Soc Am* 87: 2592-2605, 1990.
- Grothe B.** The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Prog Neurobiol* 61: 581-610, 2000.
- Grothe B, Pecka M, and McAlpine D.** Mechanisms of sound localization in mammals. *Physiol Rev* 90: 983-1012, 2010.
- Grothe B, and Sanes DH.** Bilateral inhibition by glycinergic afferents in the medial superior olive. *J Neurophysiol* 69: 1192-1196, 1993.
- Grothe B, and Sanes DH.** Synaptic inhibition influences the temporal coding properties of medial superior olivary neurons: an in vitro study. *J Neurosci* 14: 1701-1709, 1994.
- Guinan JJ, Jr., and Li RY-S.** 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, 1990.
- Guinan JJ, Jr., Norris BE, and Guinan SS.** Single auditory units in the superior olive complex II: Tonal organization and locations of unit categories. *Int J Neurosci* 4: 147-166, 1972.
- Hall JL, 2nd.** Binaural interaction in the accessory superior-olivary nucleus of the cat. *J Acoust Soc Am* 37: 814-823, 1965.

- Han Y, and Colburn HS.** Point-neuron model for binaural interaction in MSO. *Hear Res* 68: 115-130, 1993.
- Hartmann R, Topp G, and Klinke R.** Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. *Hear Res* 13: 47-62, 1984.
- Heffner RS, and Heffner HE.** Sound localization and use of binaural cues by the gerbil (*Meriones unguiculatus*). *Behav Neurosci* 102: 422-428, 1988.
- Hermann J, Pecka M, von Gersdorff H, Grothe B, and Klug A.** Synaptic transmission at the calyx of Held under in vivo-like activity levels. *J Neurophysiol* 98: 807-820, 2007.
- Herz AVM, Gollisch T, Machens CK, and Jaeger D.** Modeling single-neuron dynamics and computations: a balance of detail and abstraction. *Science* 314: 80-85, 2006.
- Holmes MH, and Cole JD.** Cochlear mechanics: analysis for a pure tone. *J Acoust Soc Am* 76: 767-778, 1984.
- Horikawa K, and Armstrong WE.** A versatile means of intracellular labeling: injection of biocytin and its detection with avidin conjugates. *J Neurosci Methods* 25: 1-11, 1988.
- Irving R, and Harrison JM.** The superior olivary complex and audition: a comparative study. *J Comp Neurol* 130: 77-86, 1967.
- Jagadeesh B, Wheat HS, and Ferster D.** Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science* 262: 1901-1904, 1993.
- Jeffress LA.** A place theory of sound localization. *J Comp Physiol Psychol* 41: 35-39, 1948.
- Jercog PE, Svirskis G, Kotak VC, Sanes DH, and Rinzel J.** Asymmetric excitatory synaptic dynamics underlie interaural time difference processing in the auditory system. *PLoS Biol* 8: e1000406, 2010.
- Joris P, and Yin TCT.** A matter of time: internal delays in binaural processing. *Trends Neurosci* 30: 70-78, 2007.
- Joris PX.** Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J Neurophysiol* 76: 2137-2156, 1996.
- Joris PX.** Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* 23: 6345-6350, 2003.
- Joris PX, Carney LH, Smith PH, and Yin TCT.** Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of Neurophysiology* 71: 1022-1036, 1994a.
- Joris PX, and Smith PH.** The volley theory and the spherical cell puzzle. *Neuroscience* 154: 65-76, 2008.
- Joris PX, Smith PH, and Yin TCT.** Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *Journal of Neurophysiology* 71: 1037-1051, 1994b.
- Joris PX, Van de Sande B, Louage DH, and van der Heijden M.** Binaural and cochlear disparities. *Proc Natl Acad Sci U S A* 103: 12917-12922, 2006.
- Joris PX, and van der Heijden M.** Early Binaural Hearing: The Comparison of Temporal Differences at the Two Ears. *Annu Rev Neurosci* 42: 433-457, 2019.
- Joris PX, and Yin TCT.** Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J Neurophysiol* 73: 1043-1062, 1995.
- Joris PX, and Yin TCT.** Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J Neurophysiol* 79: 253-269, 1998.

- Karino S, Smith PH, Yin TCT, and Joris PX.** Axonal branching patterns as sources of delay in the mammalian auditory brainstem: a re-examination. *J Neurosci* 31: 3016-3031, 2011.
- Khurana S, Liu Z, Lewis AS, Rosa K, Chetkovich D, and Golding NL.** An essential role for modulation of hyperpolarization-activated current in the development of binaural temporal precision. *J Neurosci* 32: 2814-2823, 2012.
- Khurana S, Remme MWH, Rinzel J, and Golding NL.** Dynamic interaction of I_h and I_{K-LVA} during trains of synaptic potentials in principal neurons of the medial superior olive. *J Neurosci* 31: 8936-8947, 2011.
- Kil J, Kageyama GH, Semple MN, and Kitzes LM.** Development of ventral cochlear nucleus projections to the superior olivary complex in gerbil. *J Comp Neurol* 353: 317-340, 1995.
- Kim EJ, Jacobs MW, Ito-Cole T, and Callaway EM.** Improved monosynaptic neural circuit tracing using engineered rabies virus glycoproteins. *Cell Rep* 15: 692-699, 2016.
- Kiss A, and Majorossy K.** Neuron morphology and synaptic architecture in the medial superior olivary nucleus. Light- and electron microscope studies in the cat. *Exp Brain Res* 52: 315-327, 1983.
- Klumpp RG, and Eady HR.** Some measurements of interaural time difference thresholds. *J Acoust Soc Am* 28: 859-860, 1956.
- Kopp-Scheinflug C, Dehmel S, Dörrscheidt GJ, and Rübsamen R.** Interaction of excitation and inhibition in anteroventral cochlear nucleus neurons that receive large endbulb synaptic endings. *J Neurosci* 22: 11004-11018, 2002.
- Köpl C.** Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis of the barn owl *Tyto alba*. *J Neurophysiol* 77: 364-377, 1997.
- Köpl C, and Carr CE.** Maps of interaural time difference in the chicken's brainstem nucleus laminaris. *Biol Cybern* 98: 541-559, 2008.
- Kuba H, Ishii TM, and Ohmori H.** Axonal site of spike initiation enhances auditory coincidence detection. *Nature* 444: 1069-1072, 2006.
- Kuba H, Yamada R, Fukui I, and Ohmori H.** Tonotopic specialization of auditory coincidence detection in nucleus laminaris of the chick. *J Neurosci* 25: 1924-1934, 2005.
- Kuenzel T, Borst JGG, and van der Heijden M.** Factors controlling the input-output relationship of spherical bushy cells in the gerbil cochlear nucleus. *J Neurosci* 31: 4260-4273, 2011.
- Kuwabara N, DiCaprio RA, and Zook JM.** Afferents to the medial nucleus of the trapezoid body and their collateral projections. *Journal of Comparative Neurology* 314: 684-706, 1991.
- Kuwabara N, and Zook JM.** Local collateral projections from the medial superior olive to the superior paraolivary nucleus in the gerbil. *Brain Res* 846: 59-71, 1999.
- Kuwabara N, and Zook JM.** Projections to the medial superior olive from the medial and lateral nuclei of the trapezoid body in rodents and bats. *Journal of Comparative Neurology* 324: 522-538, 1992.
- Kuwada S, Yin TCT, Syka J, Buunen TJF, and Wickesberg RE.** Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. *J Neurophysiol* 51: 1306-1325, 1984.
- Laback B, Egger K, and Majdak P.** Perception and coding of interaural time differences with bilateral cochlear implants. *Hear Res* 322: 138-150, 2015.
- Langford TL.** Responses elicited from medial superior olivary neurons by stimuli associated with binaural masking and unmasking. *Hear Res* 15: 39-50, 1984.
- Leakey DM, Sayers BM, and Cherry C.** Binaural Fusion of Low- and High-Frequency Sounds. *The Journal of the Acoustical Society of America* 30: 222-222, 1958.

- Leibold C.** Influence of inhibitory synaptic kinetics on the interaural time difference sensitivity in a linear model of binaural coincidence detection. *J Acoust Soc Am* 127: 931-942, 2010.
- Leibold C, and van Hemmen JL.** Spiking neurons learning phase delays: how mammals may develop auditory time-difference sensitivity. *Phys Rev Lett* 94: 168102, 2005.
- Liberman MC.** Central projections of auditory-nerve fibers of differing spontaneous rate. I. Anteroventral cochlear nucleus. *J Comp Neurol* 313: 240-258, 1991.
- Liberman MC.** The cochlear frequency map for the cat: labeling auditory-nerve fibers of known characteristic frequency. *J Acoust Soc Am* 72: 1441-1449, 1982.
- Licklider JCR, Webster JC, and Hedlund JM.** On the Frequency Limits of Binaural Beats. *The Journal of the Acoustical Society of America* 22: 468-473, 1950.
- Lorteije JAM, Rusu SI, Kushmerick C, and Borst JGG.** Reliability and precision of the mouse calyx of Held synapse. *J Neurosci* 29: 13770-13784, 2009.
- Louage DHG, Joris PX, and van der Heijden M.** Decorrelation sensitivity of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise. *J Neurosci* 26: 96-108, 2006.
- Louage DHG, van der Heijden M, and Joris PX.** Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. *J Neurosci* 25: 1560-1570, 2005.
- Macleod KM, and Carr CE.** Synaptic mechanisms of coincidence detection. In: *Synaptic mechanisms in the auditory system*, edited by Trussell LO, Popper AN, and Fay RR. New York: Springer, 2011, p. 135-164.
- Magnusson AK, Kapfer C, Grothe B, and Koch U.** Maturation of glycinergic inhibition in the gerbil medial superior olive after hearing onset. *J Physiol* 568: 497-512, 2005.
- Maier JK, and Klump GM.** Resolution in azimuth sound localization in the Mongolian gerbil (*Meriones unguiculatus*). *J Acoust Soc Am* 119: 1029-1036, 2006.
- Maki K, and Furukawa S.** Acoustical cues for sound localization by the Mongolian gerbil, *Meriones unguiculatus*. *J Acoust Soc Am* 118: 872-886, 2005.
- Mardia KV, and Jupp PE.** *Directional statistics*. Chichester, UK: Wiley, 2000.
- Mathews PJ, Jercog PE, Rinzel J, Scott LL, and Golding NL.** Control of submillisecond synaptic timing in binaural coincidence detectors by K_v1 channels. *Nat Neurosci* 13: 601-609, 2010.
- Mc Laughlin M, Chabwine JN, van der Heijden M, and Joris PX.** Comparison of bandwidths in the inferior colliculus and the auditory nerve. II: Measurement using a temporally manipulated stimulus. *J Neurophysiol* 100: 2312-2327, 2008.
- Mc Laughlin M, Franken TP, van der Heijden M, and Joris PX.** The interaural time difference pathway: a comparison of spectral bandwidth and correlation sensitivity at three anatomical levels. *J Assoc Res Otolaryngol* 15: 203-218, 2014.
- Mc Laughlin M, Verschooten E, and Joris PX.** Oscillatory dipoles as a source of phase shifts in field potentials in the mammalian auditory brainstem. *J Neurosci* 30: 13472-13487, 2010.
- McAnally KI, Brown M, and Clark GM.** Acoustic and electric forward-masking of the auditory nerve compound action potential: evidence for linearity of electro-mechanical transduction. *Hear Res* 106: 137-145, 1997.
- Middlebrooks JC, and Green DM.** Sound localization by human listeners. *Annu Rev Psychol* 42: 135-159, 1991.

- Miller CA, Abbas PJ, Robinson BK, Nourski KV, Zhang F, and Jeng FC.** Electrical excitation of the acoustically sensitive auditory nerve: single-fiber responses to electric pulse trains. *J Assoc Res Otolaryngol* 7: 195-210, 2006.
- Mills AW.** Lateralization of High-Frequency Tones. *The Journal of the Acoustical Society of America* 32: 132-134, 1960.
- Mills AW.** On the Minimum Audible Angle. *The Journal of the Acoustical Society of America* 30: 237-246, 1958.
- Moushegian G, Rupert A, and Whitcomb MA.** Brain-stem neuronal response patterns to monaural and binaural tones. *J Neurophysiol* 27: 1174-1191, 1964a.
- Moushegian G, Rupert A, and Whitcomb MA.** Medial superior-olivary-unit response patterns to monaural and binaural clicks. *Journal of the Acoustical Society of America* 36: 196-202, 1964b.
- Moushegian G, Rupert AL, and Gidda JS.** Functional characteristics of superior olivary neurons to binaural stimuli. *J Neurophysiol* 38: 1037-1048, 1975.
- Moxon EC.** Neural and mechanical responses to electric stimulation of the cat's inner ear. Massachusetts Institute of Technology, Cambridge, Mass., 1971, p. 265.
- Müller M.** The cochlear place-frequency map of the adult and developing mongolian gerbil. *Hear Res* 94: 148-156, 1996.
- Myoga MH, Lehnert S, Leibold C, Felmy F, and Grothe B.** Glycinergic inhibition tunes coincidence detection in the auditory brainstem. *Nat Commun* 5: 3790, 2014.
- Nabel AL, Callan AR, Gleiss SA, Kladisios N, Leibold C, and Felmy F.** Distinct distribution patterns of potassium channel sub-units in somato-dendritic compartments of neurons of the medial superior olive. *Front Cell Neurosci* 13: 38, 2019.
- Navawongse R, and Voigt HF.** Single neuron recordings in dorsal cochlear nucleus (DCN) of awake gerbil. *Hear Res* 255: 44-57, 2009.
- Ohlemiller KK, and Echter SM.** Functional correlates of characteristic frequency in single cochlear nerve fibers of the Mongolian gerbil. *J Comp Physiol A* 167: 329-338, 1990.
- Overholt EM, Rubel EW, and Hyson RL.** A circuit for coding interaural time differences in the chick brainstem. *J Neurosci* 12: 1698-1708, 1992.
- Palmer AR, and Shackleton TM.** Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. *J Assoc Res Otolaryngol* 10: 233-250, 2009.
- Parameshwaran S, Carr CE, and Perney TM.** Expression of the Kv3.1 potassium channel in the avian auditory brainstem. *J Neurosci* 21: 485-494, 2001.
- Pecka M, Brand A, Behrend O, and Grothe B.** Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *J Neurosci* 28: 6914-6925, 2008.
- Peña JL, Viète S, Funabiki K, Saberi K, and Konishi M.** Cochlear and neural delays for coincidence detection in owls. *J Neurosci* 21: 9455-9459, 2001.
- Perkins RE.** An electron microscopic study of synaptic organization in the medial superior olive of normal and experimental chinchillas. *J Comp Neurol* 148: 387-415, 1973.
- Plassmann W, Peetz W, and Schmidt M.** The cochlea in gerbilline rodents. *Brain Behav Evol* 30: 82-101, 1987.
- Plauška A, Borst JGG, and van der Heijden M.** Predicting binaural responses from monaural responses in the gerbil medial superior olive. *J Neurophysiol* 115: 2950-2963, 2016.
- Polsky A, Mel BW, and Schiller J.** Computational subunits in thin dendrites of pyramidal cells. *Nat Neurosci* 7: 621-627, 2004.

- Rautenberg PL, Grothe B, and Felmy F.** 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: 385-396, 2009.
- Ravicz ME, Rosowski JJ, and Voigt HF.** Sound-power collection by the auditory periphery of the Mongolian gerbil *Meriones unguiculatus*. I: Middle-ear input impedance. *J Acoust Soc Am* 92: 157-177, 1992.
- Rayleigh L.** XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 13: 214-232, 1907.
- Reardon TR, Murray AJ, Turi GF, Wirblich C, Croce KR, Schnell MJ, Jessell TM, and Losonczy A.** Rabies virus CVS-N2c^{ΔG} strain enhances retrograde synaptic transfer and neuronal viability. *Neuron* 89: 711-724, 2016.
- Recio A, and Rhode WS.** Basilar membrane responses to broadband stimuli. *J Acoust Soc Am* 108: 2281-2298, 2000.
- Roberts MT, Seeman SC, and Golding NL.** A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry. *Neuron* 78: 923-935, 2013.
- Roberts MT, Seeman SC, and Golding NL.** 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, 2014.
- Rose JE, Brugge JF, Anderson DJ, and Hind JE.** Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* 30: 769-793, 1967.
- Rose JE, Gross NB, Geisler CD, and Hind JE.** 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, 1966.
- Roth GL, Aitkin LM, Andersen RA, and Merzenich MM.** Some features of the spatial organization of the central nucleus of the inferior colliculus of the cat. *J Comp Neurol* 182: 661-680, 1978.
- Rubel EW, and Parks TN.** Organization and development of brain stem auditory nuclei of the chicken: tonotopic organization of n. magnocellularis and n. laminaris. *J Comp Neurol* 164: 411-433, 1975.
- Ryan A.** Hearing sensitivity of the mongolian gerbil, *Meriones unguiculatus*. *J Acoust Soc Am* 59: 1222-1226, 1976.
- Sato M, Baumhoff P, and Kral A.** Cochlear implant stimulation of a hearing ear generates separate electrophonic and electroneural responses. *J Neurosci* 36: 54-64, 2016.
- Sayers BM, and Cherry EC.** Mechanism of binaural fusion in the hearing of speech. *Journal of the Acoustical Society of America* 29: 973-987, 1957.
- Schroeder MR.** New viewpoints in binaural interactions. In: *Psychophysics and Physiology of Hearing*, edited by Evans EF, and Wilson JPAcademic Press, 1977, p. 455-467.
- Scott LL, Hage TA, and Golding NL.** Weak action potential backpropagation is associated with high-frequency axonal firing capability in principal neurons of the gerbil medial superior olive. *J Physiol* 583: 647-661, 2007.
- Scott LL, Mathews PJ, and Golding NL.** Perisomatic voltage-gated sodium channels actively maintain linear synaptic integration in principal neurons of the medial superior olive. *J Neurosci* 30: 2039-2050, 2010.
- Scott LL, Mathews PJ, and Golding NL.** Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive. *J Neurosci* 25: 7887-7895, 2005.

- Seidl AH, and Rubel EW.** Systematic and differential myelination of axon collaterals in the mammalian auditory brainstem. *Glia* 64: 487-494, 2016.
- Seidl AH, Rubel EW, and Barría A.** Differential conduction velocity regulation in ipsilateral and contralateral collaterals innervating brainstem coincidence detector neurons. *J Neurosci* 34: 4914-4919, 2014.
- Seidl AH, Rubel EW, and Harris DM.** Mechanisms for adjusting interaural time differences to achieve binaural coincidence detection. *J Neurosci* 30: 70-80, 2010.
- Semple MN, and Kitzes LM.** Single-unit responses in the inferior colliculus: different consequences of contralateral and ipsilateral auditory stimulation. *J Neurophysiol* 53: 1467-1482, 1985.
- Shackleton TM, Arnott RH, and Palmer AR.** Sensitivity to interaural correlation of single neurons in the inferior colliculus of guinea pigs. *J Assoc Res Otolaryngol* 6: 244-259, 2005.
- Shamma SA, Shen NM, and Gopalaswamy P.** Stereausis: binaural processing without neural delays. *J Acoust Soc Am* 86: 989-1006, 1989.
- Silver RA.** Neuronal arithmetic. *Nat Rev Neurosci* 11: 474-489, 2010.
- Singheiser M, Fischer BJ, and Wagner H.** Estimated cochlear delays in low best-frequency neurons in the barn owl cannot explain coding of interaural time difference. *J Neurophysiol* 104: 1946-1954, 2010.
- Smith AJ, Owens S, and Forsythe ID.** Characterisation of inhibitory and excitatory postsynaptic currents of the rat medial superior olive. *J Physiol* 529: 681-698, 2000.
- Smith DI, and Mills JH.** Anesthesia effects: auditory brain-stem response. *Electroencephalogr Clin Neurophysiol* 72: 422-428, 1989.
- Smith PH.** Structural and functional differences distinguish principal from nonprincipal cells in the guinea pig MSO slice. *J Neurophysiol* 73: 1653-1667, 1995.
- Smith PH, Joris PX, Carney LH, and Yin TC.** Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *Journal of Comparative Neurology* 304: 387-407, 1991.
- Smith PH, Joris PX, and Yin TCT.** Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. *J Neurophysiol* 79: 3127-3142, 1998.
- Smith PH, Joris PX, and Yin TCT.** Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J Comp Neurol* 331: 245-260, 1993.
- Spangler KM, Warr WB, and Henkel CK.** The projections of principal cells of the medial nucleus of the trapezoid body in the cat. *Journal of Comparative Neurology* 238: 249-262, 1985.
- Spirou GA, and Berrebi AS.** Glycine immunoreactivity in the lateral nucleus of the trapezoid body of the cat. *J Comp Neurol* 383: 473-488, 1997.
- Spirou GA, Brownell WE, and Zidanic M.** Recordings from cat trapezoid body and HRP labeling of globular bushy cell axons. *Journal of Neurophysiology* 63: 1169-1190, 1990.
- Spitzer MW, and Semple MN.** Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *J Neurophysiol* 73: 1668-1690, 1995.
- Stevens SS, and Newman EB.** The Localization of Actual Sources of Sound. *The American Journal of Psychology* 48: 297-306, 1936.
- Stotler WA.** An experimental study of the fiber connections of the cells of the medial trapezoid nucleus of the cat. *Anatomical Record* 115: 414-415, 1953.

- Svirskis G, Dodla R, and Rinzel J.** Subthreshold outward currents enhance temporal integration in auditory neurons. *Biol Cybern* 89: 333-340, 2003.
- Svirskis G, Kotak V, Sanes DH, and Rinzel J.** Enhancement of signal-to-noise ratio and phase locking for small inputs by a low-threshold outward current in auditory neurons. *J Neurosci* 22: 11019-11025, 2002.
- Tamás G, Szabadics J, and Somogyi P.** Cell type- and subcellular position-dependent summation of unitary postsynaptic potentials in neocortical neurons. *J Neurosci* 22: 740-747, 2002.
- Tan Q, and Carney LH.** A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. *J Acoust Soc Am* 114: 2007-2020, 2003.
- Tasaki I.** Nerve impulses in individual auditory nerve fibers of guinea pig. *J Neurophysiol* 17: 97-122, 1954.
- Temchin AN, Recio-Spinoso A, Cai H, and Ruggero MA.** Traveling waves on the organ of Corti of the chinchilla cochlea: spatial trajectories of inner hair cell depolarization inferred from responses of auditory-nerve fibers. *J Neurosci* 32: 10522-10529, 2012.
- Ter-Mikaelian M, Sanes DH, and Semple MN.** Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. *J Neurosci* 27: 6091-6102, 2007.
- Thompson AM, and Schofield BR.** Afferent projections of the superior olivary complex. *Microsc Res Tech* 51: 330-354, 2000.
- Thompson SP.** LI. On the function of the two ears in the perception of space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 13: 406-416, 1882.
- Tollin DJ, and Yin TCT.** Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *J Neurosci* 25: 10648-10657, 2005.
- Tsuchitani C.** Functional organization of lateral cell groups of cat superior olivary complex. *J Neurophysiol* 40: 296-318, 1977.
- van den Honert C, and Stypulkowski PH.** Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. *Hear Res* 14: 225-243, 1984.
- van den Honert C, and Stypulkowski PH.** Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hear Res* 29: 207-222, 1987.
- van der Heijden M.** A synaptic theory of internal delays. *J Acoust Soc Am* 144: 2967-2970, 2018.
- van der Heijden M, and Joris PX.** Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies. *J Neurosci* 23: 9194-9198, 2003.
- van der Heijden M, and Joris PX.** Panoramic measurements of the apex of the cochlea. *J Neurosci* 26: 11462-11473, 2006.
- van der Heijden M, Lorteije JAM, Plauška A, Roberts MT, Golding NL, and Borst JGG.** Directional hearing by linear summation of binaural inputs at the medial superior olive. *Neuron* 78: 936-948, 2013.
- Versteegh CPC, and van der Heijden M.** Basilar membrane responses to tones and tone complexes: nonlinear effects of stimulus intensity. *J Assoc Res Otolaryngol* 13: 785-798, 2012.
- Vonderschen K, and Wagner H.** Detecting interaural time differences and remodeling their representation. *Trends Neurosci* 37: 289-300, 2014.
- Wagner H, Asadollahi A, Bremen P, Endler F, Vonderschen K, and von Campenhausen M.** Distribution of interaural time difference in the barn owl's inferior colliculus in the low- and high-frequency ranges. *J Neurosci* 27: 4191-4200, 2007.

- Wiegner A, Wright CG, and Vollmer M.** Multichannel cochlear implant for selective neuronal activation and chronic use in the free-moving Mongolian gerbil. *J Neurosci Methods* 273: 40-54, 2016.
- Wightman FL, and Kistler DJ.** The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91: 1648-1661, 1992.
- Winters BD, Jin SX, Ledford KR, and Golding NL.** Amplitude normalization of dendritic EPSPs at the soma of binaural coincidence detector neurons of the medial superior olive. *J Neurosci* 37: 3138-3149, 2017.
- Yamada R, Okuda H, Kuba H, Nishino E, Ishii TM, and Ohmori H.** The cooperation of sustained and phasic inhibitions increases the contrast of ITD-tuning in low-frequency neurons of the chick nucleus laminaris. *J Neurosci* 33: 3927-3938, 2013.
- Yin TCT, and Chan JCK.** Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* 64: 465-488, 1990.
- Yin TCT, and Kuwada S.** Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J Neurophysiol* 50: 1020-1042, 1983.
- Young SR, and Rubel EW.** Frequency-specific projections of individual neurons in chick brainstem auditory nuclei. *J Neurosci* 3: 1373-1378, 1983.
- Zhou Y, Carney LH, and Colburn HS.** 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: 3046-3058, 2005.

Appendix

Summary

Samenvatting

Curriculum Vitae

PhD-portfolio

Acknowledgements

Summary

The ability to locate sound sources is essential for both predatory and prey species, especially for tetrapods. Different species of this class employ varying neural mechanisms for sound localization. For some species in order to locate sound sources, the ability to detect subtle differences between sounds arriving at both ears constitutes a major evolutionary advantage. For mammals with good low-frequency hearing (including humans) the first sound localization step happens in the brainstem – this is where the inputs from both ears meet for the first time. The auditory brainstem consists of several nuclei that contribute to different aspects of sound processing. One of those nuclei – the medial superior olive (MSO) - is well known to play key role in horizontal sound localization, especially for low-frequency sounds. An important feature of MSO neurons is their response sensitivity to time delays (or ITD) between the ears.

With virtually no experimental evidence back in 1948, Lloyd Jeffress envisioned a neural mechanism through which a hypothetical nucleus could operate as a sound localizer. He described gradually shorter or longer inputs innervating different neurons from both ears. This gradient of inputs – delay lines – would be responsible for different location sound sources preferentially activating different cells, which in turn would convey this sound location information towards higher-level brain areas. Since then experimental studies have shown that this mechanism is present in birds, however, discrepancies with this theory were found for mammals. To explain these discrepancies several alternative theories have been proposed, the two most prominent ones being a) local inhibition adjusting internal delays (Brand et al. 2002; Pecka et al. 2008) b) time mismatches between ears stem from sound perception on different physical locations in each cochlea (stereausis) (Day and Semple 2011; Shamma et al. 1989).

The main goal of our studies was to elucidate which of these different proposed mechanisms could account for sound localization in the mammalian brainstem. The experimental data so far was mostly gathered from recordings done from individual MSO cells, but no subthreshold information could be recorded. We performed not only cell-attached recordings but also whole-cell recordings *in vivo* while presenting novel sound stimuli. These experiments were

the first of its kind and allowed us to capture the inputs from both ears – subthreshold responses – and have a better insight in the way in which MSO neurons integrate dendritic inputs at their soma.

In **Chapter 2** we showed that juxtacellular recordings can be used to describe inputs to MSO in a quantitative manner. This was done by comparing *in vivo* whole-cell, juxtacellular and *in vitro* recordings. First of all, we observed that the measured best delay of the MSO neuron could be estimated from the timing of monaural subthreshold responses. Furthermore, linear summation of the two inputs predicted binaural features of the MSO remarkably well. We found no significant differences between left and right subthreshold input characteristics, suggesting that MSO itself does not contribute to internal delay generation. Absence of responses during unfavorable time delays seemed to come from the lack of timed inputs from both ears rather than inhibitory effects. Finally, we found that MSO neurons respond in a supralinear fashion when inputs from both ears arrive within a very short time window, making them very sensitive coincidence detectors.

Further on, we looked into the linear summation importance for the MSO function. In **Chapter 3** we used a multiple-frequency (wideband) stimulus to evoke monaural and binaural responses in the MSO. Monaural responses were cross-correlated (linear summation) and the final result yielded an expected ITD sensitivity of that neuron. Binaural stimulation allowed us to record the true ITD sensitivity of the same neuron. We then compared the two ITD curves by fitting them with standard fitting function and comparing their parameter values. Best ITDs were predicted quite accurately, just as in Chapter 2. Additionally, we found a high correlation between measured and predicted ITD curve phases and frequencies – the temporal parameters. Overall, linear input summation reproduced the measured ITD responsiveness (not only best ITD) quite accurately. We did not observe any differential filtering of inputs during their interaction, thus, together with the results of Chapter 2, we could not find evidence supporting a major role for well-timed inhibition in ITD regulation.

Finally, we aimed to test the stereausis hypothesis in **Chapter 4**. The idea behind this hypothesis is that internal delay is arising due to sounds activating different locations on different cochleae. Three important implications arise from this theory: a) MSO neurons should be sensitive to different sound frequencies for both ears, b) electrical stimulation of the cochlea should circumvent this time delay mechanism and c) positive best delays should

correspond to lower frequencies on the contralateral side and vice versa. In agreement with other studies we found that MSO neurons are sensitive to slightly different frequencies for both ears; though the difference was small, it was enough to generate relevant time delays. By implanting electrodes close to each ear, we measured electrically evoked ITD responses. They indeed did not match those evoked by sound. Despite this, we did not find a correlation between frequency mismatches between the ears and best delay of MSO neurons, making stereausis unlikely to be a dominant mechanism underlying sound localization in the MSO.

In short, being able to record subthreshold responses in the MSO, we could investigate its function in greater detail. MSO cells sum their inputs from both ears in a linear manner and produce supralinear outputs when these inputs arrive within a short time window. Linear input summation is enough to reproduce the neuronal ITD sensitivity, without a need for inhibitory or local modulation effects. Despite the fact that MSO neurons receive two inhibitory inputs and two excitatory inputs from both ears, often tuned to different frequencies, we found no evidence for local, well-timed inhibition or stereausis to affect ITD sensitivity in the MSO. This leaves, by exclusion, Jeffress' model as the most plausible internal delay source. However, some of the predictions of a pure delay line model were not borne out by the data. Recent studies found evidence for differential axonal conductance velocities between inputs from both ears; more detailed experiments are needed for conclusive results. The role of inhibition at the MSO is still largely undetermined and hopefully can be understood by a pharmacological approach or by performing developmental MSO research. A further understanding of the cellular mechanisms of sound localization may help in improving cochlear implant performance.

Samenvatting

Weten waar geluiden vandaan komen is belangrijk voor zowel roof- als prooidieren, met name viervoeters (tetrapoden). Viervoeters gebruiken verschillende neurale mechanismen om geluiden te lokaliseren. Hierbij biedt het vermogen om subtiele verschillen in de aankomsttijden van geluiden tussen beide oren te detecteren een belangrijk evolutionair voordeel. Voor zoogdieren met een goed laagfrequent gehoor (waaronder mensen) vindt de eerste geluidslokalisatiestap plaats in de hersenstam, waar de input van beide oren voor het eerst samenkomt. De auditieve hersenstam bestaat uit verschillende kernen die bijdragen aan verschillende aspecten van geluidsverwerking. Eén van die kernen - de mediale bovenste olijkern (Medial Superior Olive, MSO) - speelt een sleutelrol bij geluidslokalisatie in het horizontale vlak, met name voor laagfrequente geluiden. Een belangrijk kenmerk van MSO neuronen is hun gevoeligheid voor de relatieve aankomsttijd van geluiden bij beide oren (Interaural Time Difference, ITD).

Lloyd Jeffress stelde in 1948 een neuraal mechanisme voor geluidslokalisatie voor. Hij beschreef een hersenkern die geïnnerveerd werd door inputs vanuit beide oren waarvan de lengte systematisch varieerde. Omdat de tijd die een signaal erover doet afhangt van de lengte van een axon, zou voor elke cel een ander verschil in reistijd ('internal delay') gelden voor de signalen uit beide oren. Jeffress postuleerde verder dat deze cellen alleen actief zouden worden als de inputs uit beide oren tegelijk zouden aankomen, met andere woorden als het verschil in reistijd juist het verschil in aankomsttijd van het geluid bij beide oren compenseert ('best delay'). Deze gradiënt van vertragingslijnen zou zó verantwoordelijk zijn voor een systematische map van geluidsbronnen binnen deze hersenkern. Sindsdien hebben experimentele studies aangetoond dat dit mechanisme aanwezig is bij vogels, maar er werden discrepanties met deze theorie gevonden voor zoogdieren. Om deze discrepanties te verklaren zijn verschillende alternatieve theorieën voorgesteld. De twee meest prominente zijn: a) lokale synaptische inhibitie, die zorgt voor een verandering van de effectieve reistijd (Brand et al. 2002; Pecka et al. 2008) en b) cellen in de MSO worden geïnnerveerd vanuit verschillende plekken binnen het slakkenhuis, waardoor verschillen in reistijd binnen het

slakkenhuis bijdragen aan de reistijdverschillen ('stereausis') (Day and Semple 2011; Shamma et al. 1989).

Het belangrijkste doel van dit proefschrift was om op te helderen welke van deze verschillende voorgestelde mechanismen het meest toepasbaar is bij zoogdieren. De experimentele gegevens tot dusver waren weliswaar grotendeels afkomstig van opnames van individuele MSO cellen, maar er waren geen metingen van hun inputs *in vivo* gemaakt. Wij hebben zowel juxtacellulaire als zgn. whole-cell afleidingen *in vivo* in gerbils gemaakt terwijl we (verschillende) geluiden bij beide oren afspeelden. Hiermee konden we een beter inzicht te krijgen in de manier waarop MSO-neuronen dendritische inputs in hun cellichaam integreren.

In **Hoofdstuk 2** hebben we laten zien dat juxtacellulaire opnames gebruikt kunnen worden om inputs in een MSO neuron op een kwantitatieve manier te meten. Dit werd aangetoond door *in vivo* whole-cell en juxtacellulaire opnames te vergelijken met opnames in hersenplakjes. We hebben dit type metingen gebruikt om te laten zien dat de gemeten 'best delay' van het MSO neuron kon worden voorspeld op basis van de timing van synaptische potentialen in antwoord op stimulatie van één oor. Verder vonden we dat een simpele optelling ($1+1=2$) van de responsen op stimulatie van het linker- en het rechteroor de synaptische potentialen in respons op stimulatie van beide oren opmerkelijk goed. We vonden geen significante verschillen tussen de temporele eigenschappen van de linker en rechter synaptische potentialen, wat suggereert dat de MSO zelf niet bijdraagt aan het genereren van de internal delay. We vonden verder aanwijzingen dat de afwezigheid van activiteit tijdens stimulatie bij de 'worst delay' eerder leek te komen door afwezigheid van input dan door synaptische inhibitie. Ten slotte ontdekten we dat de kans op een actiepotentiaal in MSO-neuronen een supralineaire functie is van de grootte van de inputs, wat een verklaring biedt voor hun functie als coïncidentiedetectoren.

In **Hoofdstuk 3** hebben we een stimulus met meerdere toonfrequenties (breedband stimulus) gebruikt om monaurale en binaurale reacties op te wekken in de MSO. Door kruiscorrelatie van de monaurale responsen kon de ITD gevoeligheid van het neuron adequaat voorspeld worden, niet alleen de waarde van de 'best ITD', maar ook het gedrag van de cel daarbuiten. Net als in hoofdstuk 2 vormde lineaire sommatie van de synaptische inputs een goede verklaring voor de metingen. Ten slotte wilden we de stereausis hypothese testen in

hoofdstuk 4. Het idee achter deze hypothese is dat geluiden voor het linker en het rechteroor verschillende locaties op de cochleae activeren, waardoor de looptijd van die signalen verschillend wordt voor de inputs in de MSO vanuit beide oren. Deze theorie doet drie voorspellingen: a) MSO neuronen moeten gevoelig zijn voor verschillende geluidsfrequenties voor beide oren, b) elektrische stimulatie van de auditieve zenuw moet dit tijdsvertragingmechanisme omzeilen en c) positieve 'best delays' moeten overeenkomen met lagere frequenties op de contralaterale kant en *vice versa*. In overeenstemming met andere studies vonden we dat MSO-neuronen gevoelig zijn voor verschillende frequenties voor beide oren; hoewel het verschil klein was, was het voldoende om relevante tijdsvertragingen te genereren. Door elektroden dicht bij elk oor te implanteren konden we elektrisch opgewekte best ITDs meten. Deze kwamen inderdaad niet overeen met de auditieve 'best ITD' in hetzelfde neuron. Desondanks hebben we geen correlatie gevonden tussen frequentie verschillen tussen de oren en de best delay van MSO neuronen, waardoor het onwaarschijnlijk is dat stereausis een dominant mechanisme kan zijn, wat een belangrijke rol speelt bij geluidslokalisatie in de MSO.

Kortom, doordat we de inputs in MSO neuronen *in vivo* konden meten, konden we hun functie onderzoeken. MSO cellen tellen hun inputs van beide oren op een lineaire manier op en produceren supralineaire output wanneer deze inputs binnen een kort tijdsbestek arriveren. Sommatie van lineaire inputs is voldoende om de neuronale ITD-gevoeligheid te reproduceren, zonder dat er remmende of lokale modulatie-effecten nodig zijn. Ondanks het feit dat MSO-neuronen twee remmende inputs en twee exciterende inputs vanuit beide oren ontvangen, vaak gevoelig voor verschillende geluidsfrequenties, vonden we geen bewijs voor lokale, goed getimede synaptische inhibitie dan wel stereausis om de ITD-gevoeligheid in de MSO te verklaren. Hierdoor blijven de axonale vertragingen uit het model van Jeffress als de meest plausibele bron voor 'internal delay' over. Sommige voorspellingen van een strikt vertraginglijnmodel werden echter niet bevestigd door de metingen. Recente studies hebben bewijs gevonden voor differentiële axonale geleidingssnelheden tussen inputs van beide oren; meer gedetailleerde experimenten zijn echter nodig. De rol van synaptische inhibitie binnen de MSO is nog grotendeels onbekend en kan hopelijk worden begrepen door een farmacologische benadering of door het bestuderen van de ontwikkeling van de MSO.

Een verder begrip van de cellulaire mechanismen van geluidslokalisatie kan helpen bij het verbeteren van de prestaties van cochleaire implantaten.

Curriculum Vitae

Personal information

Name: Andrius Plauška

Date of birth: January 24th 1988

Place of Birth: Šilutė, Lithuania

Education/Work

- | | |
|-----------------------|---|
| 2016 – present | Junior researcher
Femtonics Kft.
Budapest, Hungary |
| 2011 – 2016 | PhD student in Neuroscience
Erasmus MC
Rotterdam, the Netherlands |
| 2009 – 2011 | Master student in Biophysics (<i>cum laude</i>)
Vrije Universiteit
Amsterdam, the Netherlands |
| 2005 – 2009 | Bachelor Student in Theoretical Physics
Vilnius University
Vilnius, Lithuania |

Publications

Stefan Witte, **Andrius Plauška**, Margreet C. Ridder, Laura van Berge, Huibert D. Mansvelder, and Marie Louise Groot, Short-coherence off-axis holographic phase microscopy of live cell dynamics, **Biomed. Opt. Express** Volume 3, Issue 9, 22 August 2012, Pages 2184-2189

Marcel van der Heijden, Jeannette A.M. Lorteije, **Andrius Plauška**, Michael T. Roberts, Nace L. Golding, J. Gerard G. Borst, Directional Hearing by Linear Summation of Binaural Inputs at the Medial Superior Olive, **Neuron**, Volume 78, Issue 5, 5 June 2013, Pages 936-948

Andrius Plauška, J. Gerard G. Borst, Marcel van der Heijden, Predicting binaural responses from monaural responses in the gerbil medial superior olive, **Journal of Neurophysiology**, Volume 115, Issue 6, 21 March 2016, Pages 2950-2963

Andrius Plauška, Marcel van der Heijden J. Gerard G. Borst, A test of the stereausis hypothesis for sound localization in mammals, **the Journal of Neuroscience**, Volume 37, Issue 30, 26 July 2017, Pages 7278-7289

PhD-Portfolio

Courses

Laboratory Animal Science course

Cognitive neuroscience

Experimental neurophysiology: theory and practice

Imaging life at the molecular level

Teaching

VO Hearing: From research to diagnosis (2013 - 2015)

VO Action potential (2014 - 2015)

VO Heart potential (2015)

Attended conferences

- | | |
|------|---|
| 2015 | WASdag , Rotterdam, December, talk |
| | Binaural Pentagon , Paris, October, talk |
| | Dutch Neuroscience Meeting , Lunteren, June, poster |
| | WASdag , Groningen, May |
| 2014 | ARO Midwinter meeting , Baltimore, Maryland, February, poster |
| | WASdag , Leiden, December |
| | ONWAR meeting , Woudschoten, November, talk |
| 2013 | WASdag , Leuven, May |
| | ONWAR meeting , Woudschoten, November, poster |
| 2012 | ONWAR meeting , Woudschoten, November, poster |
| | Dutch Neuroscience Meeting , Lunteren, May |
| 2011 | ONWAR meeting , Woudschoten, November |

Acknowledgements

First and foremost, I want to thank my promotor Gerard for all the help and guidance throughout and after my PhD years. Without your continuous support and belief in me, this thesis would not have seen the light of day. You gave me this great opportunity to join your lab and perform intricate experiments and data interpretation that yielded three great publications. I have thoroughly enjoyed every conversation and debate I had with you. And I still remember those late nights in the lab doing experiments together. To this date I consider you to be the most knowledgeable person I have ever known. Thank you for everything.

I am also very grateful to my co-promotor Marcel for making me feel welcome in his lab. Your amazing knowledge of MatLab and programming, in general, was so infectious that I became a full-time programmer and data analyst in my current lab. Much of the knowledge that you have shared with me is now a foundation for my current work. I will always remember you as an inventive and driven person with a generous smile.

Both Marcel and Gerard's labs were full of lively and happy researchers, from bachelor students to postdocs and I was very happy to feel like a part of this family. Corstiaen, you were my lab mate and we shared a lot of laughter and musing discussions. During my first year I was quite lost (and I know you were aware of that) when you were explaining me how things work in the ear, but as time went by, I started understanding you more and more. I also remember you as one of the most stylish people in the department. Kudos for keeping up appearances. Wherever you are now, I hope you found a way to quench your scientific thirst.

To all the people that I met in Gerard's lab. Thank you all for being so friendly and helpful. TianTian, you always said 'yes' to all my requests to help, be it pipette pulling or solution making. Thank you for being a big mood lifter no matter the day. Rudiger, the modest and friendly German, who spent so much time with me downstairs, thank you for being there when I had emergency situations and also for introducing me to people in the department during my first months. You helped me overcome some of the hardest social barriers. Milly, you were always kind to me and we had quite some emotional moments for which I'm very grateful. Also, you were my number one source for gossip. Bonus for that. Tom, I'm sure I

heard my first 'dad jokes' from you, you also have unforgettable smile and laughter, never forget that. Laura, we spent the most time together outside the lab and I remember our late nights in pubs and long-distance bike trips together, especially the one when I broke my arm. Thank you for being such a good-hearted friend and listener. Gheis, you boisterous man. When you came to the lab you brought so much vitality and good spirit. You were always so pro-active and it was a pleasure to know you. Also thank you for not making me go to the gym even though you hinted at it up several times. Martijn, despite being very serious and critical-thinking during our lab meetings and journal clubs you showed your laid-back side during our social gatherings and it was always a pleasure to be in your company. And finally, last but not least, Aaron. You are a true friend, whom I will always treasure and remember. It was always nice to see you coming by our lab to discuss some science or just for a quick chat or a joke. Thank you for all the times you came over to gather us for lunch and even wait when we were still in the middle of something. Whenever I needed some help or just to talk, you were the first one I would go to for that. I am very happy to still be in touch with you after all these years.

Though I am not very social person, making friends or just having a nice conversation was always easy in our department. That is due to all the great and kind people from the labs of Freek, Ype, Dick and Chris. Thank you to all the amazing people that made our Friday evening borrels so lively and fun. Sander, for all the jokes – good and bad ones – and inviting us to your band performances. Nouk, you were one of the first people I got to know outside my lab and it was always such a pleasure to see you laugh after you have been working countless hours with mouse whiskers in your dark lab. I still look at your thesis (cover only) once in a while and hope you are doing great.

Freek, you had the most Italians in your squad than any other lab leader I have ever known. I don't know how you managed to handle all of them and whether this was intentional, but I'm sure that they left you some really great moments to remember. Despite spending 7 years in the Netherlands, I believe I learned more about Italian lifestyle than the Dutch one, thanks to the "Italian gang" that I got involved with throughout all 5 years of my PhD – Simo, Costanza, Tommaso, Licia, Negah, Farnaz, Sasha, Mohamad and Malik.

Simo(na), my favorite Italian! Getting to know you was such an amazing life experience for me. Everything about you is just so special and unique, your contagious laughter, your

spontaneity, Italian gesticulation, food regulations, the list goes on. There was not a single moment that I did not enjoy spending with you and to this day it is always an excitement and pleasure to meet you. Thank you to your amazing housemates, Costanza and Tommy, who always made me feel welcome at your home, teaching me Italian manners and introducing me to some of the best foods I have ever had. My Negah, you were always so kind, loving and inviting. You have the most captivating smile and kindness just radiates from you. I feel privileged to call you my friend. And don't forget that with or without Ace I love you the same. Sasha, you crazy Croatian, you were always so expressive, and I'm sure it was not an Italian influence, you were just destined to make people happy. Speaking of happy people, Farnaz, you were truly special and I always enjoyed your company through and through. Malik, how many French jokes you threw at me, and boy did I not get most of them. But those that hit me, always made me laugh and I always appreciated those moments. Mohamad, your medical stories always livened up our conversations. I hope you found your favorite field in medicine. For such a smart guy like you the sky is the limit. Licia, for such a lightweight girl you surely had a surprisingly strong grip when hugging, this was the first thing I noticed about you and through the years I got to enjoy your company and your laughter as well as your hugs. Together with Aaron and the nine of you I experienced the most memorable moments in the Netherlands. From Halloween film festival throughout the night, to Queen's Days in Amsterdam, posh drinks on top of Rotterdam, Christmas dinners and so much more. I truly love you all. And finally, Carmen, though I got to know you for a shorter period of time than the rest, your humble and compassionate personality was always a great company wherever we went.

A big thank you also goes to the staff of the department. Kees, for saving us when we were struck by any technical issues and also having an amazing workshop. It was always a pleasure to see you tinkering with old gadgets. Elize and Erika, thank you for teaching me slicing, histology and lab etiquette so patiently and thoroughly. Mandy, for keeping the conversations going in the histology lab during long hours of waiting. An extended thank you goes to Elize for always making time to keep the borrel list populated and providing us with all the supplies for it.

Elise and Loes, a heartfelt gratitude from me to you for always making sure that our paperwork is done, I never felt lost in bureaucracy thanks to you. The staff and students are

lucky to have you as coordinators. Chris, thank you for keeping the department running like clockwork and for inviting and approving so many great guest speakers that I had honor to hear from throughout the years.

Tom, you were the first one who taught me how to use a light microscope and I will not forget when you told me that it is not neurons I am seeing, but my eyelashes. I swear I got better at this since then. Linda, you were there for me when I needed a place to crash in-between my living places. I will never forget that, thank you from the bottom of my heart.

There are many other people whom I met daily in the hallways or the labs, and I do remember you all, Yarmo, Vincenzo, Jan Willem, Cullen, Oscar, Paolo, Ype, Maarten, Steven, Hans, Marcel, Gao, Geeske, Laurens, Mario and many more. It was always a pleasure to spend time in your company.

My final thank you goes to my mother. You raised me alone, and I know it was not easy. Yet you did everything you could to provide me the best education possible and I would not have reached these heights without your unconditional love and support. Ačiū už viską, Mamyte.