Temporal coding of the periodicity of monaural and binaural complex tones in the guinea pig auditory brainstem

Sami Alsindi

Centre for the Neural Basis of Hearing,

The Physiological Laboratory,

University of Cambridge

This dissertation is submitted to the University of Cambridge for

the degree of Doctor of Philosophy, July 2017

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

It does not exceed the 60000 word limit prescribed by the Degree Committee of Biology.

Sami Alsindi

Corpus Christi College, Cambridge

July 2017

Summary

Sami Alsindi

Temporal coding of the periodicity of monaural and binaural complex tones in the guinea pig auditory brainstem

Humans report a strong pitch percept in response to a complex tone – the sum of a series of harmonics – presented to either a single ear (‘monaurally’) or both ears (‘diotically’). Interspike interval histograms of responses of neurons in the auditory system to monaural complex tones show a peak at the period of the pitch reported by humans – a ‘neural correlate of pitch’. However, the same pitch percept can be generated by presenting complexes with harmonics distributed across both ears (‘dichotically’). This requires combination of the neural signals underlying pitch from both sides of the auditory system, termed ‘binaural fusion’. Temporal coding generally deteriorates along the auditory pathway; binaural fusion should occur at a relatively early stage. One of the prime candidates is in the superior olivary complex (SOC).

Although the guinea pig auditory system has been extensively studied, this work is the first *in vivo* investigation of the guinea pig SOC. Cells of the lateral superior olive (LSO) show sensitivity to interaural level differences; medial superior olive (MSO) cells show sensitivity to interaural time differences. Additionally, cells with responses similar to the medial nucleus of the trapezoid body (MNTB) and superior paraolivary nucleus (SPN) of other species were found in the guinea pig SOC. Presumed MNTB cells showed a three-component spike waveform shape; presumed SPN cells responded at the offset of contralaterally-presented stimuli.

MSO and LSO cells respond to the overall pitch of complex tones, even if the monaural waveforms presented to each ear differ; this is consistent with the perception of humans. In contrast, cells of the ventral cochlear nucleus, which provide the main input to MSO and LSO cells, do not show evidence of a binaural pitch response. In conclusion, SOC cells are able to encode the pitch of binaural complex tones in their spike timing patterns.

Acknowledgements

I originally applied to Cambridge to be a Biochemist, of all things! After the best three lectures I have ever attended, I soon realised the error of my ways. In one of those lectures, I was introduced to the tonotopy of the cochlea and was riveted (I may have kept the lecturer behind for an hour after a Saturday lecture asking questions!). Thank you, David Tolhurst, not only for this but also for being a very supportive Advisor during my PhD.

For a 2nd year summer project, I looked to Cambridge-based Principal Investigators. Thankfully, the PI I was most interested in working with (Ian) responded and invited me to his lab, where he introduced me to the neurophysiological study of pitch in the auditory system. It seemed like the perfect confluence of my two passions: neuroscience and music. There, he introduced me to a type of pitch stimulus that was generated by what sounded like mere white noise in each headphone on its own but, when listened together, produced a clear pitch percept. I was baffled, and knew that binaural pitch was what I wanted to investigate. Of course it should go without saying that I thank my supervisor, Ian Winter, for not only sparking my interest in this field, but also providing a great environment in which to carry out such an interesting PhD project, and his invaluable contribution to my scientific understanding and training. It's been a fantastic three-and-a-half years.

Throughout my PhD, Arkadiusz 'Arek' Stasiak, a postdoc in our lab, offered tremendous scientific and personal support. He helped immensely with development of my computer programming skills, for which I cannot ever thank him enough. We went on to become good friends, both in and out of the lab, although I can't rule out that this was due to the excellent electrodes he manufactured, which facilitated the best experiments of my PhD!

I owe an enormous debt of gratitude to Anita Shelley and Mel Quy, who provided excellent training for the histology crucial to this Thesis. Thanks also goes to Mark Sayles, Roy Patterson, Ray Meddis, Hedwig Gockel, Jan Diepenbrock, Nihaad Paraouty, Alex Billig, Alan Archer-Boyd and many, many others for stimulating discussions and for making my PhD such an enjoyable one.

An endless thanks goes to my father Zuhair and my mother Sahar, who were instrumental in instilling an early love of science, and, along with my brothers Wassim and Nabeel, somehow managed to put up with my antics over the years. However, above all, I need to express my gratitude for the one thing that was able to pull me through the long nights of searching for MSO units: the DeLonghi MAGNIFICA coffee machine. I couldn't have done it without you…

Sami Alsindi, July 2017

Table of Contents

List of Abbreviations

Φ_b	Mean interaural phase
AM	Amplitude modulation
AN	Auditory nerve
ANF	Auditory nerve fibre
AOISI	All-order interspike interval histogram
AVCN	Anteroventral cochlear nucleus
bITD	Best interaural time difference
BBN	Broadband noise
BF	Best frequency
BM	Basilar membrane
BMLD	Binaural masking level difference
BP	Best phase
CAP	Compound action potential
CD	Characteristic delay
CF	Characteristic frequency
CN	Cochlear nucleus
CNIC	Central nucleus of the inferior colliculus
CP	Characteristic phase
cpHTC	Cosine-phase harmonic tone complex
CS	Sustained chopper
CT	Transient chopper
CV	Coefficient of variation
dB	Decibel
DCN	Dorsal cochlear nucleus
DCT	Dichotic complex tone
DF	Dominant frequency
DMPO	Dorsomedial periolivary nucleus
DNLL	Dorsal nucleus of the lateral lemniscus
DRP	Dichotic repetition pitch
EC	Equalisation-cancellation model
EPSP	Excitatory postsynaptic potential
ENV	Temporal envelope
F0	Fundamental frequency
F0DL	F0-difference limen

FI	Facilitation index
FOISI	First-order interspike interval histogram
frzHTC	Frozen random-phase harmonic tone complex
FSL	First spike latency
GABA	γ-amino butyric acid
GBC	Globular bushy cell
GDT	Gap detection threshold
HCN	Hyperpolarisation-activated cyclic nucleotide-gated ion channel
HP	Huggins pitch
HTC	Harmonic tone complex
IC	Inferior colliculus
IHC	Inner hair cell
I_h	Hyperpolarisation-activated current
I_K	Low-voltage-activated potassium current
ILD	Interaural level difference
IPD	Interaural phase difference
IPSP	Inhibitory postsynaptic potential
IRN	Iterated rippled noise
ISI	Interspike interval
I_{TCa}	Hyperpolarisation-activated calcium ion channel-mediated current
ITD	Interaural time difference
IXT	Interaural crosstalk
LF	Low-frequency
LL	Lateral lemniscus
LNTB	Lateral nucleus of the trapezoid body
LOCS	Lateral olivocochlear system
LSO	Lateral superior olive
m-EC	Modified equalisation-cancellation model
MGB	Medial geniculate body
MNTB	Medial nucleus of the trapezoid body
MOCS	Medial olivocochlear system
MSO	Medial superior olive
NBN	Narrowband noise
NL	Nucleus laminaris
NLL	Nuclei of the lateral lemniscus
NM	Nucleus magnocellularis
OC	Onset chopper

OHC	Outer hair cell
OI	Onset ideal
OL	Onset locker
PDI	Pitch discrimination interference
PL	Primary-like
PL(PP)	Primary-like with prepotentials
PN	Primary-like with notch
PP	Prepotential
PSTH	Peri-stimulus time histogram
PVCN	Posteroventral cochlear nucleus
$Q_{10\text{-dB}}$	Filter 10-dB quality factor
rpHTC	Random-phase harmonic tone complex
SAC	Shuffled autocorrelogram
SAM	Sinusoidally amplitude-modulated
SBC	Spherical bushy cell
SD	Standard deviation
SL	Sensation level
SNR	Signal-to-noise ratio
SPL	Sound pressure level
SPN	Superior paraolivary nucleus
SR	Spontaneous rate
SOC	Superior olivary complex
TFS	Temporal fine structure
VCN	Ventral cochlear nucleus
VNTB	Ventral nucleus of the trapezoid body
VS	Vector strength

Chapter 1

General introduction

1.1 OVERVIEW

Human psychophysical studies have shown that stimuli with very different properties can produce the same pitch. This includes stimuli that differ between ears, or 'dichotic' stimuli. Although physiological studies have found neural correlates of monaural pitch perception, to date, none has shown evidence of binaural integration of pitch consistent with human psychophysics.

This study recorded responses of single neurons at early stages of the ascending auditory pathway, with a focus on the medial superior olive (MSO) in the superior olivary complex (SOC). MSO neurons receive excitatory inputs from both sides of the auditory system. This gives rise to interaural time difference (ITD)-sensitivity. Their responses are hypothesised to underpin sound source localisation. Neural correlates of binaural integration of pitch were found in the responses of these neurons to the same stimuli that evoke binaural pitch percepts in humans.

In this Chapter, I will briefly outline the path by which neural signals reach the SOC, highlighting evidence from the guinea pig where possible. Then, I will review the basics of pitch perception and models of how pitch is thought to be encoded. Finally, I will outline the benefits of binaural listening – using both ears – and discuss binaural pitch.

1.2 BASIC AUDITORY PATHWAY

1.2.1 External and middle ear

Sound waves emanating from a sound source enter the external ear, consisting of the pinna, concha and external auditory meatus. The folds of the pinna and properties of the concha and external auditory meatus lead to an initial filtering of the sound spectrum. Interference of reflections of sound waves leads to frequency-dependent reinforcement or cancellation,

forming the basis of the head-related transfer function, a pattern of spectral peaks and troughs that varies with the direction of the sound source relative to the head (cat – Rice *et al.*, 1992; human – Pralong & Carlile, 1994); this is thought to be a cue for (monaural) sound source localisation.

The filtered sound waves then impart vibration onto the tympanic membrane. This is connected to the cochlea via the ossicular chain of bones: tympanic membrane movement causes the malleus to move, which in turn causes the incus to move, which is in turn connected to the stapes. The stapes contacts the oval window, a membranous structure on the wall of the cochlea. These elements comprise the middle ear, and are shown schematically in **Figure 1A (page 5)**. Together, they act to reduce the impedance mismatch between the air of the external and middle ear and the endolymph fluid in the cochlea. Motion of the stapes and the tympanic membrane can be reduced via the middle-ear reflex activation of stapedius and tensor tympani muscles, which offers some protection to the cochlea from loud sounds (rabbit – Counter & Borg, 1993).

1.2.2 Cochlea

Running almost the entire extent of the spiral-shaped cochlea are three fluid-filled spaces: the scala vestibuli, the scala media and the scala tympani. Partially dividing the latter two compartments is the basilar membrane (BM). Vibrations of the oval window cause the formation of a travelling wave on the BM, moving from the base of the cochlea to the apex (human – Békésy, 1953). Different regions of the BM resonate at different frequencies. At the base of the cochlea, the BM is stiffer and narrower: this leads to a maximal response to higher frequencies. Towards the apex, the BM becomes wider and less stiff, leading to a greater

response to lower frequencies. The frequency response of a particular BM region is described as a cochlear filter, with centre frequency corresponding to the cochlear place.

The organ of Corti contains around 3500 inner hair cells (IHCs) (human – Wright *et al.*, 1987). Hair cells are so-called due to hair-like stereocilia protruding from their top surface. Vibration of the BM causes displacement of the stereocilia; displacement in the direction of the largest stereocilium causes tension to develop in tip links, opening mechanically-gated cation channels (guinea pig – Pickles *et al.*, 1984). Potassium ions in the endolymph then flow into the IHCs, depolarising them and triggering neurotransmitter release via ribbon synapses (cat – Liberman *et al.*, 1990), activating auditory nerve fibres (ANFs).

Figure 1B (page 5) shows a cross-section through the cochlea. An IHC atop a given region of the BM shows a characteristic frequency (CF) equivalent to the cochlear filter centre frequency. ANFs connected to these IHCs show the same CF, inheriting their receptive field from the IHC (guinea pig – Evans, 1972; guinea pig – Russell & Sellick, 1978; guinea pig – Dallos, 1986). The frequency-to-place map of the cochlea is preserved in tonotopic maps throughout the auditory pathway.

The organ of Corti also contains three to five rows of outer hair cells (OHCs) (human – Glueckert *et al.*, 2005). These are contractile (guinea pig – Brownell *et al.*, 1985) and act to decrease hearing thresholds (guinea pig – Murugasu & Russell, 1995); the exact mechanism underpinning this remains uncertain (reviewed in Fettiplace & Hackney, 2006).

Signals arising from sound transduction in the cochlea pass through the auditory nerve (AN) to the cochlear nucleus (CN), thereafter branching extensively into multiple pathways. Cells in the ventral CN (VCN) provide the inputs to the superior olivary complex (SOC). Ascending projections from the majority of cells in the auditory brainstem converge at the inferior colliculus (IC). **Figure 2 (page 5)** shows a simplified diagram of this ascending auditory pathway.

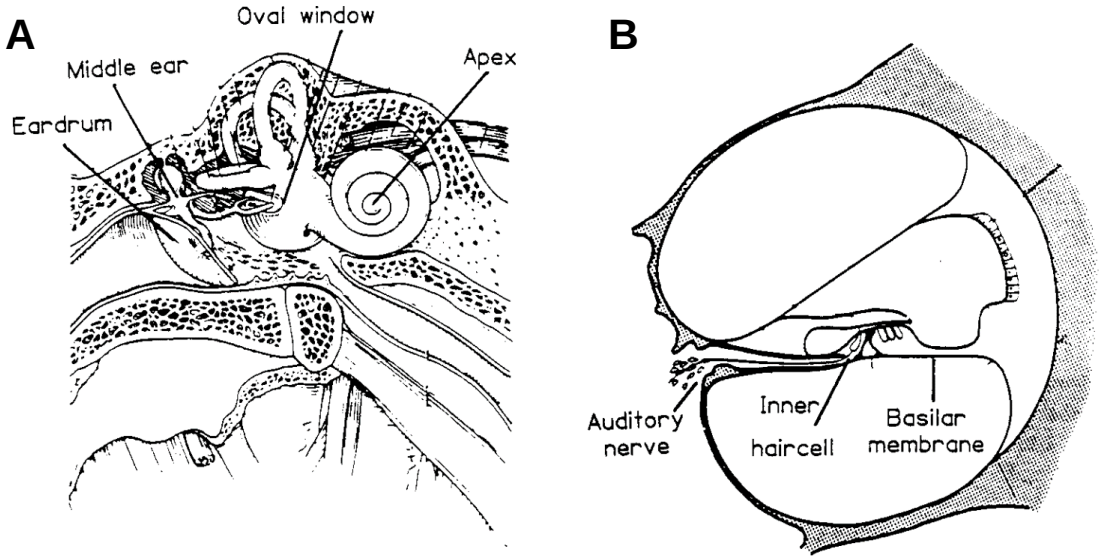

Figure 1 **Schematic cross-sections through the temporal bone (A) and the cochlea (B).** Figure excerpted from Hartmann, 1996.

A: Incident sound waves cause the tympanic membrane ('eardrum') to vibrate; this vibration is passed through the ossicular chain to the cochlea, with the stapes contacting the oval window of the cochlea.

B: Vibration of the BM causes displacement of IHCs, opening mechanically-gated cation channels; this allows a potassium ion influx, depolarising IHCs and activating connected ANFs.

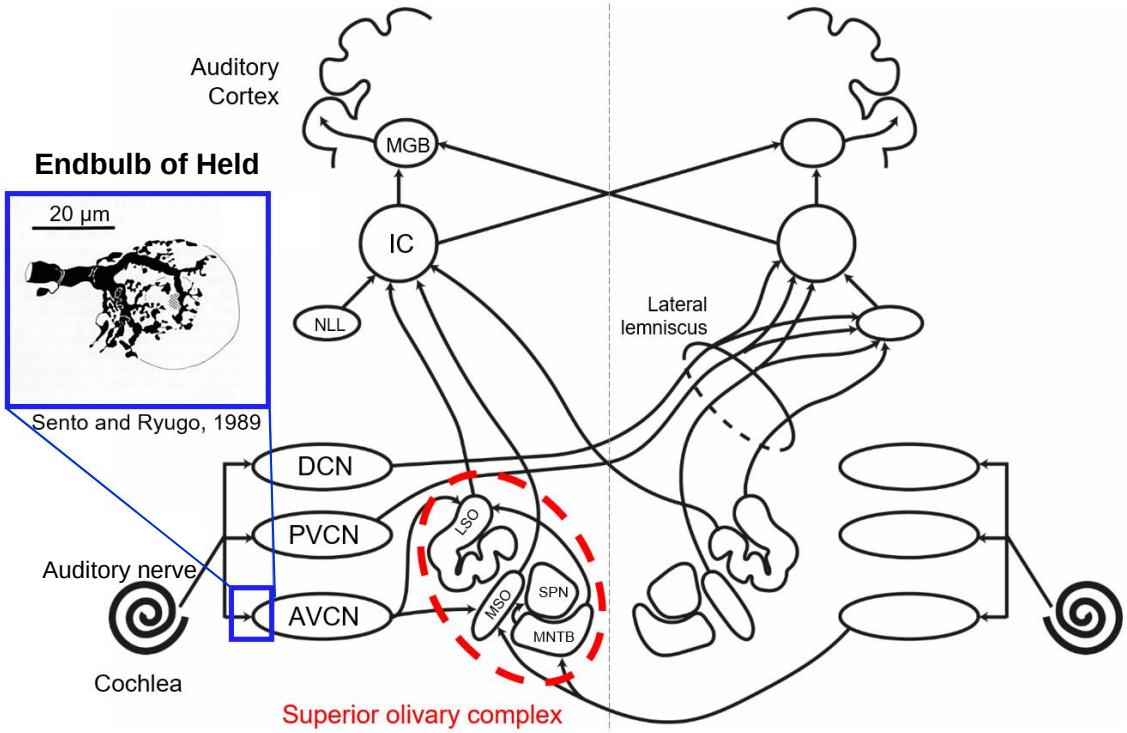

Figure 2 **Simplified diagram of the ascending auditory system**. A synaptic specialisation, the endbulb of Held, is shown. MGB: medial geniculate body; NLL: nuclei of the lateral lemniscus.

1.2.3 Auditory nerve

Up to 30 ANFs are connected to a given IHC (cat – Liberman *et al.*, 1990). The cochlea gives rise to around 31000 ANFs in humans (Spoendlin & Schrott, 1989).

ANFs are divided into three classes based on their spontaneous rate (SR), the spike rate obtained in the absence of stimulation: low-, medium- and high-SR. The three categories show different auditory thresholds: high-SR fibres have the lowest thresholds and saturate at high sound levels; low-SR units have the highest thresholds and do not saturate at high sound levels (guinea pig – Evans, 1972; cat – Liberman, 1978). The threshold occurs at the CF of an ANF. All ANFs arising from the cochlea terminate onto cells of the ipsilateral CN (cat – Ramón y Cajal, 1909). This is shown schematically in **Figure 2 (page 5)**.

1.2.4 Cochlear nucleus

The CN has three subdivisions: the anteroventral CN (AVCN), which contains bushy cells that show similar response profiles to ANFs, the posteroventral CN (PVCN), which contains cells that provide inhibitory inputs to the rest of the CN and cells that detect the onset of sounds, and the dorsal CN (DCN), a region with highly non-linear responses (cat – Bandyopadhyay & Young, 2013) that also receives somatosensory inputs (guinea pig – Shore, 2005).

Cells of the DCN and PVCN generally project through the dorsal and intermediate acoustic stria respectively, passing through the lateral lemniscus (LL), directly to the inferior colliculus (IC). In contrast, cells of the AVCN project to the ipsilateral or contralateral SOC via the ventral acoustic stria. The synapse between ANFs and bushy cells is particularly large and shows adaptations called ‘endbulbs of Held’ (see **Figure 2, page 5** and **Section 5.1.1.1, page 140**). The cell types and response properties of the ventral CN (VCN) are discussed further in

Section 5.1 (page 139). In **Figure 2 (page 5)**, a schematic of the ANF inputs to the CN subdivisions, and the ascending output pathways of the CN subdivisions, can be seen in the context of the wider auditory pathway.

1.2.5 Superior olivary complex

The SOC consists of a group of auditory nuclei in the mammalian brainstem. There are 13 anatomically-defined nuclei in the guinea pig SOC (Schofield & Cant, 1991, 1992). The largest nuclei, the MSO and LSO, form the principal ascending output of the SOC. The other 11 nuclei are termed 'periolivary nuclei'. There is substantial variation in the morphology of the nuclei between mammals. In animals with larger head sizes, there is an enlargement of low-frequency regions in nuclei such as the lateral superior olive (LSO), and low-frequency nuclei such as the MSO are larger (Kuwada & Yin, 2012). The MSO and LSO receive inputs originating from bushy cells of the VCN of both sides of the auditory system. As such, the SOC is a major site of binaural integration in the ascending auditory pathway.

Cells of the MSO receive excitatory inputs from spherical bushy cells (SBCs) in the ipsilateral and contralateral AVCN. The time-course of these SBC inputs differs on each side. Stimuli must be presented at an ITD for neural signals to coincide at the cell; the ITD required to do this is termed the best ITD (bITD). The bITD distribution is generally biased to the contralateral hemifield[1] (dog – Goldberg & Brown, 1969; cat – Yin & Chan, 1990; guinea pig IC – McAlpine *et al.*, 2001; gerbil – Day & Semple, 2011; chinchilla LL – Bremen & Joris, 2013)[2]. ITDs

[1] By convention, this is positive ITD.

[2] Studies labelled 'LL' purport to record from MSO cell axons in the LL, however anatomical verification of this paradigm has yet to be published. A particular issue is that unit first spike latency is used to distinguish MSO units from others, however MSO and LSO units are not separated using this criterion, thus it is possible that there is LSO 'contamination' of the data.

manifest when a sound source is closer to one ear; the ITD sensitivity of the MSO is thought to underpin low-frequency sound source localisation (Lord Rayleigh, 1907; Jeffress, 1948; Goldberg & Brown, 1969). Cells of the LSO are sensitive to interaural level differences (ILDs), arising from head-shadowing of sound energy at the ear farther from the sound source. The response properties of the MSO and LSO are discussed in greater depth in **Sections 1.2.5.1 (page 9)**, **3.1.3 (page 53)** and **3.1.4 (page 61)**.

Cells of the medial nucleus of the trapezoid body (MNTB) form inhibitory inputs to the LSO (see **Section 3.1.4.1, page 61**), giving rise to its ILD-sensitivity, and the MSO (see **Section 3.1.3.1, page 54**), which may alter its ITD-sensitivity (see **Section 1.2.5.1, page 9**). Cells of the superior paraolivary nucleus (SPN)[3] respond at the offset of sounds and to gaps within sounds. Their response is thought to underpin gap detection threshold. The MNTB and SPN are discussed in greater depth in **Sections 3.1.1 (page 48)** and **3.1.2 (page 51)** respectively.

The functions of the remaining periolivary nuclei are not as clear. Some of the cells within these nuclei give rise to the olivocochlear bundle, a descending projection to the cochleae and both CN (guinea pig – Winter *et al.*, 1989). Cells lateral of the MSO, including some located inside the LSO, give rise to the lateral olivocochlear system (LOCS), which contacts auditory nerve dendrites below IHCs; cells medial of the MSO form the medial olivocochlear system (MOCS), which contacts OHCs (cat – Guinan *et al.*, 1983; reviewed in Brown, 2011). The MOCS reduces the vibration of the BM in response to sound (guinea pig – Cooper & Guinan, 2006), which is thought to protect the cochlea from loud sounds (Maison & Liberman, 2000). The functions of the LOCS are still debated. There is evidence that it may also limit damage from sound overexposure by modulating the responses of cochlear afferents (guinea pig – Groff &

[3] This is present in the SOC of rodents; in other mammals, a homologous nucleus, the dorsomedial periolivary nucleus, is instead present.

Liberman, 2003; mouse – Darrow *et al.*, 2007). Additionally, it has been shown that LOCS balances the neural excitability underpinning the ILD-sensitivity of the LSO, allowing for accurate sound source localisation (mouse – Darrow *et al.*, 2006).

The focus of experiments conducted in the SOC for my Thesis was to record from low-BF cells of the MSO. These cells act as coincidence detectors of their SBC inputs (see **Section 3.1.3.2, page 56**). As such, their responses could underlie binaural pitch perception (see **Section 1.4.3, page 25**). The vast majority of studies into MSO have focussed on cell responses to ITDs and the mechanisms underpinning the distribution of bITDs. The following section will describe evidence for and against these mechanisms. The specific ITD-sensitivity properties of the guinea pig MSO were not a focus of investigation in this Thesis. All binaural signals were presented approximately matched in level and close to zero ITD.

1.2.5.1 Mechanisms affecting MSO ITD-sensitivity

The first neural circuit model of ITD sensitivity was published almost 70 years ago (Jeffress, 1948); Jeffress later constrained the model to the MSO (Jeffress, 1958). The model proposes two features that account for ITD-sensitivity and variable unit bITD. Firstly, MSO cells act as coincidence detectors (see **Section 3.1.3.2, page 56**). Secondly, there exists a network of delay lines in the inputs to the MSO. If a sound source is located close to one ear, coincidence at a particular MSO cell requires that neural signals are delayed for longer before reaching the ipsilateral side, but only slightly delayed at the contralateral. The delay lines were hypothesised to be axonal path length variations. Jeffress' original depiction is shown in **Figure 3** below.

While the existence of delay lines has been shown in the barn owl (Carr & Konishi, 1988, 1990), strong evidence of delay lines (cat – Smith *et al.*, 1993; cat – Beckius *et al.*, 1999) or a place code for ITDs (cat – Yin & Chan, 1990) has not been found in mammals.

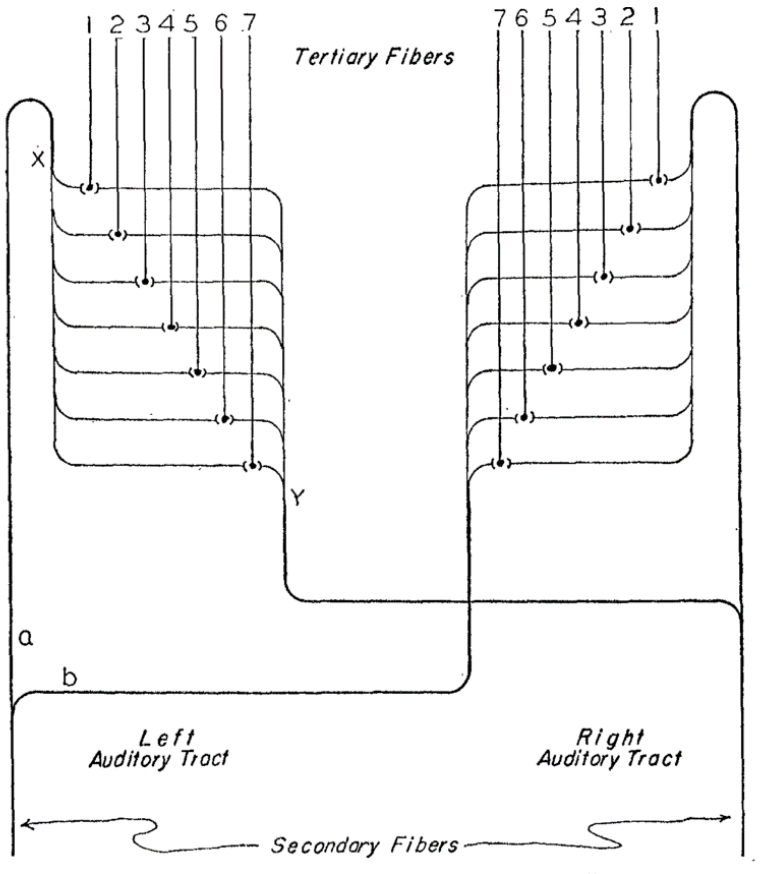

Figure 3 **The Jeffress delay-line model of ITD sensitivity**. Signals resulting from stimulation of both ears travel through a series of delay lines; cells respond maximally when they receive coincident inputs. For a given cell, the difference in physical path length of the inputs arising from the left and right auditory tracts will be offset at that unit's best ITD. Excerpted from Jeffress, 1948.

Another mechanism proposed to underpin MSO unit ITD-sensitivity is the cochlear delay hypothesis (stereausis). A delay in the input signals to an MSO cell of one side relative to the other could be introduced by a mismatch in the average BFs of the inputs to each side (Schroeder, 1977; Shamma *et al.*, 1989; Joris *et al.*, 2006). Evidence of such mismatches has

been found in recordings made from MSO output axons in chinchilla LL (Sayles *et al.*, 2016) using the approach delineated in Bremen & Joris, 2013 (see **Section 3.1.3.4, page 59**), however a recent study in gerbil found that while mismatches in MSO monaural receptive fields were present, they did not account for unit bITDs (Plauška *et al.*, 2017).

The MSO also receives inhibitory inputs, contralaterally-derived inhibition from cells of the ipsilateral MNTB and inhibition from cells of the ipsilateral lateral nucleus of the trapezoid body (LNTB) (see **Section 3.1.3.1, page 54**). These have been suggested as one of the major factors affecting MSO ITD-sensitivity, the so-called 'well-timed inhibition' hypothesis. Strychnine blockade of these inputs caused gerbil MSO bITDs to shift from contralateral-leading toward zero (Brand *et al.*, 2002). However, subsequent gerbil MSO studies either did not find evidence of IPSPs (van der Heijden *et al.*, 2013) or found that application of strychnine did not affect MSO cell bITDs, only altering the firing rates (Franken *et al.*, 2015). Additionally, no evidence was found of well-timed IPSPs in guinea pig MSO cells (Babalian, 2008).

Finally, a recent study has found that subthreshold EPSPs evoked from stimulation of the ipsilateral and/or contralateral sides can shift unit bITD (gerbil – Franken *et al.*, 2015).

1.2.6 Inferior colliculus

The output of the majority of cells from the CN and SOC converges at cells of the central nucleus of the IC (CNIC), as can be seen in **Figure 2 (page 5)** within the context of the ascending auditory pathway. However, some cells project directly to latter stages of the auditory pathway (guinea pig – Schofield *et al.*, 2014).

MSO cells project glutamatergically to the ipsilateral IC (cat – Oliver *et al.*, 1995). LSO cells project glycinergically to the ipsilateral IC and glutamatergically to contralateral IC (cat –

Glendenning & Masterton, 1983; cat – Glendenning *et al.*, 1992). Cells from the (low-frequency) lateral limb of the LSO project to the ipsilateral IC; cells from the medial limb project to the contralateral IC (cat – Glendenning *et al.*, 1985).

Cells of the IC show ITD- and ILD-sensitivity (guinea pig – Palmer *et al.*, 2007), thought to originate from MSO and LSO cell inputs. However, ILD-sensitivity also arises *de novo* at cells of the IC. ILD-sensitivity resulting from contralateral excitation and ipsilateral inhibition is reversibly blocked in some, but not all, IC cells by application of γ-amino butyric acid (GABA) antagonists; the inhibitory inputs to these cells have been traced to the dorsal nucleus of the lateral lemniscus (DNLL) (bat – Pollak *et al.*, 2002, 2003). ILD-sensitivity is maintained in the IC even after bilateral kainic acid lesions (destruction) of the SOC (rat – Li & Kelly, 1992).

In the next section, I discuss the encoding of monaural and binaural pitch, focussing on what is known from human psychophysical studies and neurophysiological correlates that may underlie these phenomena.

1.3 PITCH

Pitch is a subjective quality of sound – a perception related to the periodicity of soundwaves. The official 1990 ANSI definition is ‘a property of sound that allows it to be ordered on a scale from low to high’. Some definitions require that pitch-evoking stimuli are able to create melodies from sequences, *i.e.* that stimuli have similar representation to stimuli used in music.

Many natural stimuli evoke sensations of pitch. In Western languages, emphasis and prosody are conveyed by the contour of the pitch of voices; in some languages, *e.g.* Mandarin, the pitch contour plays a role in the meaning of otherwise identical phonemes within words. Although

phonemes themselves are conveyed by vocal tract and articulator movement that create different mid- and high-frequency spectra in the harmonics of the human voice, the fundamental frequency (F0) underlying them, originating from the periodic opening and closing of vocal folds, produces a profound pitch percept at F0. Pitch perception is also fundamental to the enjoyment of music.

For some time, it was thought that a vibration at the place of the F0 in the cochlea generated the sensation of pitch. However, one can generate a harmonic tone complex (HTC) – the sum of sine waves with frequencies that are integer multiples of an F0 ('harmonics') – missing the F0 itself; the pitch percept still occurs at the F0 (Licklider, 1954). This is called the 'missing fundamental'. Cochlear nonlinearity results in additional frequency components other than those at the places of individual carriers (reviewed in Robles & Ruggero, 2001), termed 'intermodulation distortion products'. In the case of HTCs, these are introduced at the place of the F0 (Pressnitzer & Patterson, 2001). However, such vibrations can be masked by noise (Patterson, 1969), preventing meaningful use of this cue for pitch. Subjects still perceive pitch at F0. Additionally, presenting a pair of harmonics dichotically, with one harmonic going to each ear (see **Section 4.1.1**, **page 103**), can still generate a pitch percept at F0 (Houtsma & Goldstein, 1971). Therefore, a vibration at the place of F0 cannot alone account for the pitch of harmonic sounds.

1.3.1 Periodicity

Harmonic sounds evoking the same residue pitch, equal to the F0, have common periodicities, even with very different frequency spectra (see **Figure 8, page 21**). In an HTC, the residue is at the F0: each harmonic within the complex repeats both at its own frequency and at the F0. Additionally, periodic stimuli not composed of harmonics can evoke strong pitch percepts.

Iterated rippled noise (IRN) is generated using a white noise token which is delayed and added back to itself repeatedly – the delay introduces periodicity into an otherwise aperiodic signal (Bilsen & Ritsma, 1969). The associated pitch is the inverse of the delay, *i.e.* the periodicity (Yost, 1996).

In harmonic sounds, the periodicity of the temporal envelope (ENV), the slower changes in amplitude of the signal, is equal to the periodicity of temporal fine structure (TFS), the more rapid amplitude fluctuations resulting from interactions of multiple harmonics. The relative roles of ENV and TFS on perception can be investigated by using inharmonic sounds that have differing ENV and TFS periodicities. The simplest type of inharmonic signal is an HTC where a linear frequency shift has been applied to the whole signal. Harmonic spacing is maintained at the F0 of the original complex, meaning that ENV is maintained, however, the underlying TFS is aperioidic, varying with frequency shift (Schouten *et al.*, 1962; Wightman, 1973; Evans, 1978). Human subjects perceive a change from the original pitch (de Boer, 1956a; Schouten *et al.*, 1962; Smoorenburg, 1970; Patterson & Wightman, 1976; Beerends, 1989), approximated by de Boer's rule: $\boldsymbol{p} = \frac{\Delta \boldsymbol{f}}{\boldsymbol{N}}$, where $\boldsymbol{p}$ is the perceived pitch, $\boldsymbol{\Delta f}$ is the frequency shift applied and $\boldsymbol{N}$ is the lower harmonic rank (de Boer, 1956b, 1976); this is termed the 'first effect of pitch shift'. Inharmonic stimuli have ambiguous pitches (Ritsma & Engel, 1964) and the pitch percept can vary between trials (Wiersinga-Post & Duifhuis, 1994). Pitch percepts reflect pseudoperiods of maxima in the TFS of the signals (de Boer, 1956b). That differences in TFS lead to markedly different pitch percepts points to a crucial role of TFS in pitch perception. **Figure 4** below shows how the TFS of a complex tone changes when frequency-shifted and the resulting pseudoperiods that correlate with reported pitch percepts.

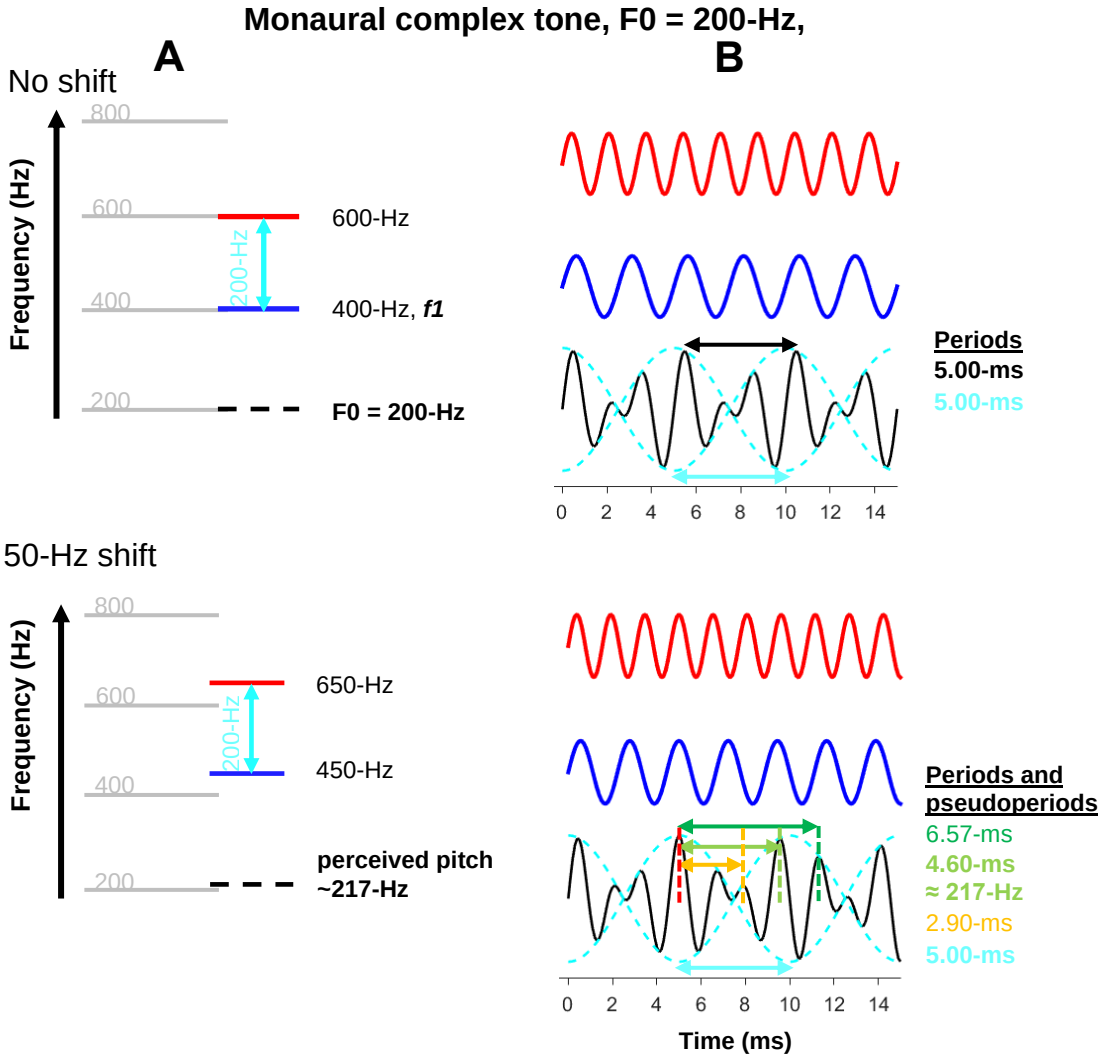

Figure 4 **Applying a frequency shift to a complex tone changes the TFS (black) while maintaining the ENV (turquoise); the pitch percept also changes, reflecting pseudoperiods in the TFS**.

A: Schematic harmonic ladders, showing pitch percepts.

B: Waveforms of each harmonic (red and blue) and the waveform of the sum (black, with turquoise envelope). Periods and pseudoperiods are represented by arrows.

Top: humans perceive the pitch of consecutive harmonics of F0 = 200-Hz to be 200-Hz. In the waveform created by adding the harmonics together, both the ENV (turquoise) and the TFS (black) repeat with the same periodicity, 5.00-ms = 200-Hz.

Bottom: Applying a 50-Hz shift results in a pitch perception of 217-Hz, rather than the envelope spacing of 200-Hz. Although the ENV is still 5-ms, pseudoperiods arise in the TFS; a neuron phase-locking (see **Section 1.3.3, page 19**) to this would show (at least a) trimodal distribution of interspike intervals. One of these pseudoperiods, 4.60-ms, is close to that of the period of the perceived pitch, 217-Hz. This means that the TFS information is more critical in pitch perception of this stimulus than ENV information.

Additional support for the role of TFS in pitch perception comes from psychophysical experiments using IRN. The pitch percept of IRN corresponded to its TFS rather than its ENV (Yost *et al.*, 1996, 1998).

Pitch perception arising from the detection of periodicity has been proposed, usually in the form of an autocorrelation operation (Licklider, 1951). Models of pitch invoking this operation are termed 'temporal models'. An influential model of this type is the summary autocorrelation function (SACF) (Meddis & Hewitt, 1991a, 1991b). Stimuli are passed through a simulated cochlear filterbank, autocorrelated, and the activity pattern generated is summed across all filter centre frequencies. The first, large, non-zero peak in the resulting SACF corresponds to the pitch percept, as shown in **Figure 5** below for an HTC. Related models of pitch include strobed temporal integration (Patterson *et al.*, 1992) and cancellation (de Cheveigné, 1998).

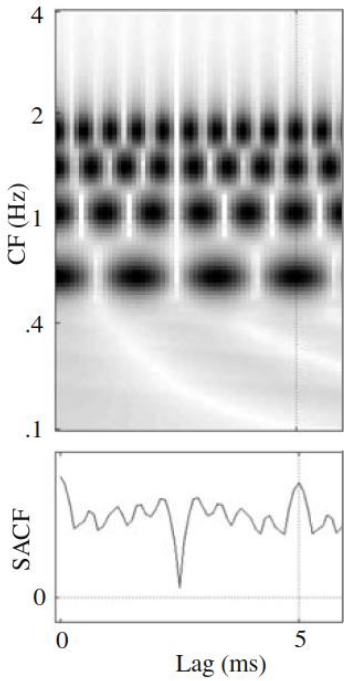

Figure 5 **SACF analysis of an HTC, F0 = 200-Hz, shows a peak at the pitch period, 5-ms**. This peak results from the sum of simultaneous maxima of the autocorrelation function across multiple simulated cochlear filters with varying centre frequencies ('CF' in this figure). Excerpted from de Cheveigné, 2004.

Autocorrelation of interspike intervals (ISIs) at multiple levels of the auditory system (see **Section 1.3.3, page 19**) could be used to determine the pitch of a wide variety of sounds.

However, a key issue for these models is the lack of physiological evidence for such an operation. Some physiological models of pitch are described in **Section 1.3.4 (page 21)**.

1.3.2 Place coding and resolvability

A frequency-to-place code ('tonotopic map') resulting from cochlear filtering is preserved at all stages of the auditory system, from the AN through to the auditory cortex. The presence of this map has long been implicated in the perception of pitch; models of pitch invoking this property are termed 'place models'.

Processing the pitch of wideband harmonic stimuli, those which span a frequency range greater than the bandwidth of a single cochlear filter, requires combination of information from multiple filters (see **Figure 6, page 18**). If the harmonics of an HTC are relatively widely spaced, such that a maximum of 2 harmonics enter each cochlear filter, they are considered 'resolved'; beyond this, harmonics are either considered partially-resolved or fully unresolved (≥3.25 harmonics per filter) (Shackleton & Carlyon, 1994). It has been suggested that the auditory system processes resolved and unresolved harmonics differently (Carlyon & Shackleton, 1994). However, F0-discrimination accuracy for an HTC composed of high, unresolved harmonics is worsened when an interferer of similar F0 but composed of low, resolved harmonics is added (Gockel *et al.*, 2004). This is termed 'pitch discrimination interference' (PDI) and suggests that resolved and unresolved harmonics are not processed separately.

The pitch of an HTC consisting of unresolved harmonics is termed the 'residue' or 'virtual' pitch (Schouten, 1940; de Boer, 1956b). In contrast, the pitch of a resolved complex is termed 'spectral pitch'. Except at very low F0s, the lower harmonics in complex tones that evoke melodic pitch will be fully resolved at the cochlea (see **Figure 6** below). In other words, each

cochlear filter receives a single harmonic, and thus only ANFs with CFs at or close to the frequency of the harmonic will respond with an increased spike rate. Models based on this approach are termed 'rate-place coding of pitch'. These models, however, cannot account for the pitch of unresolved harmonics.

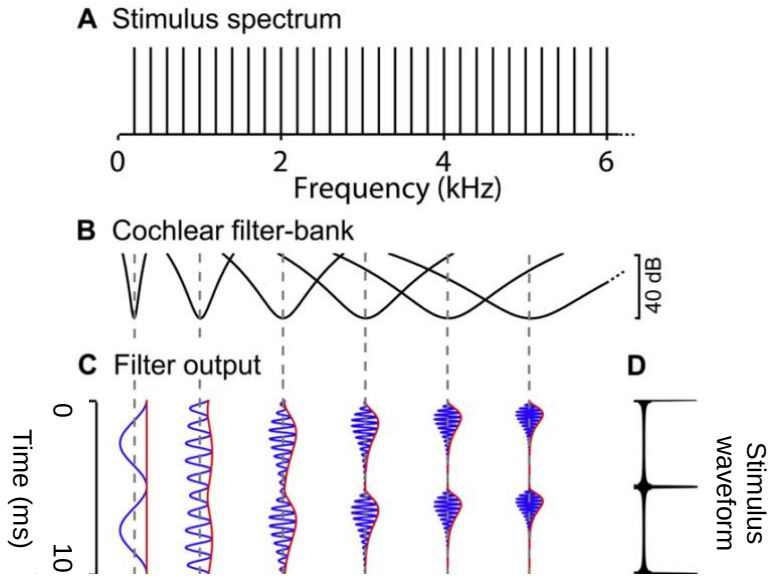

Figure 6 **As more harmonics enter a filter ('unresolved'), filter responses are increasingly modulated to the overall F0 (periodicity) of the stimulus (temporal envelope, ENV), but always respond at their centre frequencies (temporal fine structure, TFS).** Excerpted from Sayles & Winter, 2008b.

A: Schematic spectrum of a 200-Hz F0 HTC.

B: Cochlear filterbank, with filter centre frequencies of 0.2-, 1-, 2-, 3-, 4- and 5-kHz. Absolute filter bandwidth increases with increasing centre frequency.

C: Waveforms at the output of the corresponding filters.

D: Waveform of a 200 Hz F0 harmonic complex.

Each harmonic of an HTC repeats at both its own frequency and also at the F0 of the whole complex. This property makes the F0 a 'subharmonic' of a given component; subharmonic summation has been suggested as one method of obtaining a residue pitch (Hermes, 1988).

Another mechanism of this type is the comparison of filtered, resolved harmonics to some template, for example an HTC of equivalent F0/pitch – 'pattern matching'. Some models of this

type consist of a weighted sum of the individual harmonics (Goldstein, 1973; Terhardt, 1979). Others are based on learned templates that are acquired during development from exposure to harmonic sounds such as speech (Terhardt, 1978, 1979) or to any broadband stimulus (Shamma & Klein, 2000). A candidate for a physiological substrate of these models would be the recent discovery of 'harmonic template neurons' in the auditory cortex (macaque – Feng & Wang, 2017). These cells respond weakly to tones at BF but more robustly when stimulated by multiple tones, close to BF, that are consecutive harmonics of a particular F0.

1.3.3 Temporal coding of pitch

A property of the firing patterns of neurons at many stages of the auditory system is that action potentials occur locked to a narrow phase range in response to pure tones (see **Figure 7** below). This is termed 'phase-locking' and results from the properties of the mechanically-gated ion channels in the stereocilia of IHCs. The probability of channel opening is direction-dependent: displacement towards the largest stereocilium maximally opens the channels. In addition, the ribbon synapses between IHCs and ANFs minimise spike timing jitter. Although maximal spike rates generally do not exceed ~200 spikes/s, precluding an action potential for every cycle of a pure tone of higher frequency, across the population of ANFs there will be at least one spike per cycle: this is termed the 'volley theory'.

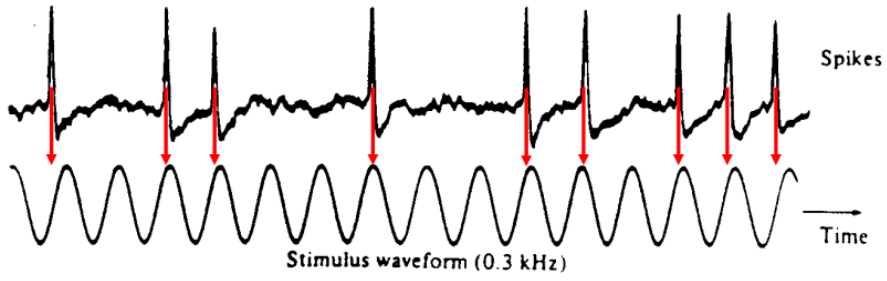

Figure 7 **Phase locking of action potentials in a cat auditory nerve fibre**. Action potentials do not necessarily occur on each period, however, the variation of phases at which they occur is small. From Arthur *et al.*, 1971.

Phase-locking efficacy to pure tones decreases with increasing tone frequency and is no longer detectable in single units for tones above 3.5-kHz in guinea pig ANFs (Palmer & Russell, 1986). At each stage in the ascending auditory pathway, some signal fidelity and temporal precision will be lost due to added jitter. The mammalian auditory system shows synaptic adaptations to minimise this factor, with the most notable examples being the endbulbs of Held (see **Section 5.1.1.1, page 140**) present in the ANF>SBC synapse and the calyx of Held (see **Section 3.1.1, page 48**) in the globular bushy cell (GBC)>MNTB synapse. However, in spite of this, the majority of cells in the guinea pig IC have upper limits of phase-locking below 1-kHz (Liu *et al.*, 2006). At the level of the cortex, the highest frequency that significant phase-locking was found in single units was 200- to 250-Hz (guinea pig – Wallace *et al.*, 2000, 2002). This means it is unlikely for pitch extraction from temporal codes to occur at a late stage of the auditory pathway.

In both resolved and unresolved cases, ANFs sensitive to low frequencies will phase-lock to the filtered harmonic with greatest energy, generally that closest to their CF – a TFS response. In the unresolved case, harmonics will interact within single filters, causing an amplitude modulation at the common periodicity equal to the harmonic spacing (typically F0 in harmonic sounds) – an ENV response (see **Figure 6, page 18**). Phase locking may still occur to the TFS of harmonics near CF, however the strong modulation of spike rates means that spike intervals are likely to occur at the period of the ENV, *i.e.* to F0.

Temporal codes can be assessed by analysing ISIs. The largest non-zero peak in ISI distributions measured at multiple stages of the auditory system occurs at the period of the pitch reported by human subjects and has been termed the ‘neural correlate of pitch’ (cat CN – Rhode, 1995; cat AN – Cariani & Delgutte, 1996a, 1996b); see **Figure 8 (page 21)**. The first

effect of pitch shift has been shown physiologically (cat CN – Rhode, 1995; cat AN – Cariani & Delgutte, 1996b; guinea pig CN – Sayles & Winter, 2008b). Additionally, a neural correlate of the pitch of IRN has been shown in the temporal responses of CN units (chinchilla – Shofner, 1991; guinea pig – Winter *et al.*, 2001; guinea pig – Sayles & Winter, 2008b). These factors together highlight the importance of phase-locking to TFS in pitch perception.

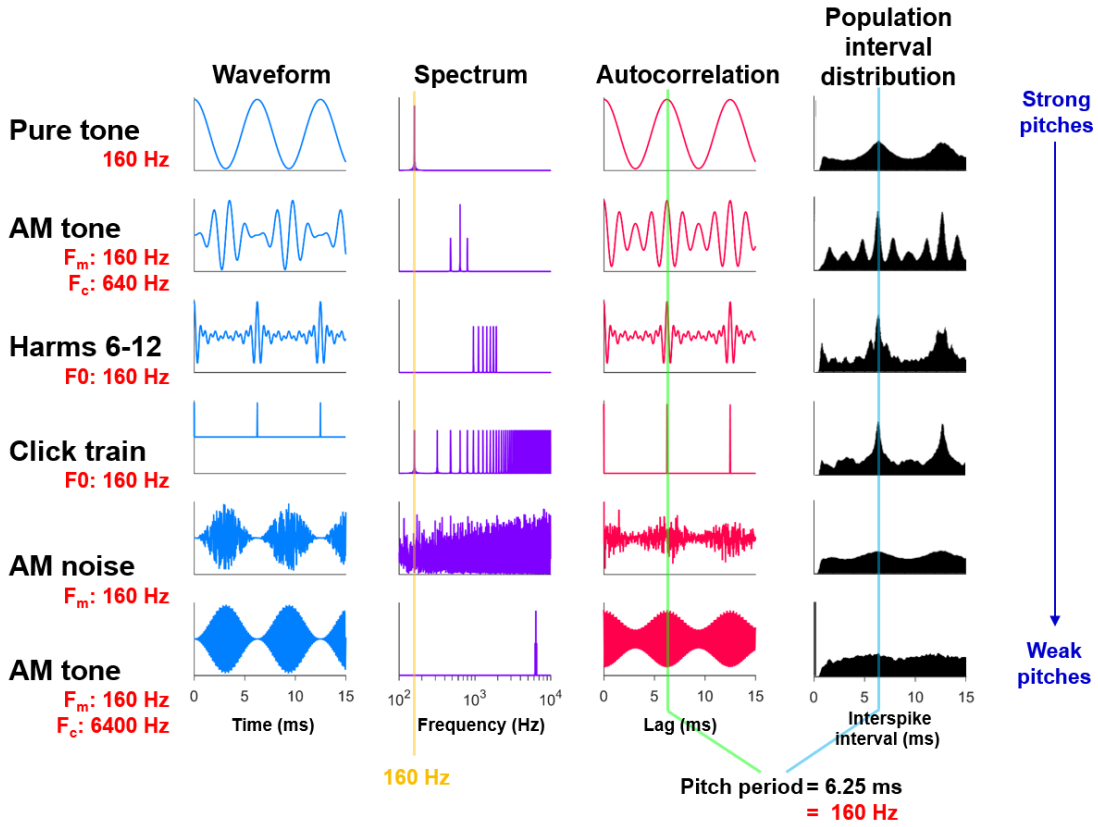

Figure 8 **Stimuli that evoke the same pitch in humans share the same pitch period, even if there is a strong difference in their spectra; a neural correlate of this pitch occurs at the same period.** The peak in the population interval distribution of the cat auditory nerve occurs at the pitch period. Adapted from Cariani & Delgutte, 1996a.

1.3.4 Brainstem and midbrain models of pitch

Chopper units of the CN fire periodically at a 'chopping frequency' independent of their BF. The response of a chopper unit to a stimulus with an F0 equal to the chopping frequency is

strongly locked to the F0: the peaks in the ENV reinforce the peaks in firing rate (gerbil – Frisina *et al.*, 1990; model – Hewitt *et al.*, 1991; guinea pig – Winter *et al.*, 2001). Convergence at cells of the IC of choppers with similar BFs but across all chopping frequencies would lead to a response dominated by F0 across all BFs. One key limitation of this theory is that the range of chopping frequencies found across a population of chopper units is not great enough to account for the range of perceptible F0s, however it may contribute to the perception of low pitches with F0s below around 500-Hz (Wiegrebe & Meddis, 2004).

It has been shown that cells of the CNIC have independent unit BFs and best modulation frequencies (BMFs) (cat – Langner & Schreiner, 1988; cat – Schreiner & Langner, 1988), forming a modulation filterbank (Dau *et al.*, 1996). Cells with BMFs close to F0 respond maximally, irrespective of BF. Convergence of high-BF IC cells with low BMFs onto low-BF IC cells would lead to these cells responding to the F0 of, for instance, a high-pass HTC. However, it has been suggested that low-BF IC cells respond instead to the vibration at the F0 place in the cochlea originating from intermodulation distortion products (guinea pig – McAlpine, 2004).

An alternate model of pitch suggests that IC cells act as coincidence detectors of the 'oscillator' choppers, sensitive to ENV, and 'reducer' principal cells of the DCN, which have longer latency responses and whose response is related to TFS, cross-correlating them (Langner, 1981, 1983, 2015).

1.4 BINAURAL INTEGRATION

Many computations of the auditory system require comparison of signals generated in the two cochleae. Information arising from stimulation of both ears must be combined and integrated in the ascending auditory pathway. Binaural processing occurs as early in the auditory pathway

as the CN (see **Section 5.1.2, page 143**); a major site of binaural processing is the SOC (see **Section 3.1, page 47**).

1.4.1 Binaural advantage

Human thresholds for detecting signals in maskers improve 1-3 dB, depending on masker and signal, when listening to signals diotically (same signal in each ear at the same level) as opposed to monaurally (Keys, 1947; Cox *et al.*, 1981). However, much greater improvements result if the signals are dichotic (different in each ear), as illustrated by binaural masking level differences (BMLDs). Paradigms assessing BMLDs utilise a combination of different configurations of white noise and different configurations of a pure tone, with differing interaural phase relationships. The addition of noise to one ear can actually decrease the signal detection threshold (*i.e.* increase the BMLD) in some cases. The greatest BMLD, 15-dB, results when presenting noise diotically and the signal phase-shifted by π-radians in one ear relative to the other (termed 'N_0S_π'). BMLDs are greatest for low-frequency signals; the N_0S_π BMLD declines to 2-3 dB when the signal is above 1500-Hz (Durlach & Colburn, 1978). Physiological data have shown that interaural phase difference (IPD)-sensitive cells of the CNIC show correlates of BMLDs (guinea pig – Palmer & Shackleton, 2002).

Findings from human binaural psychophysics and knowledge obtained from monaural neural processing of stimuli led to the development of binaural processing theories. The simplest of these is the equalisation-cancellation (EC) model (Durlach, 1963; Culling & Summerfield, 1995). This model accounts for BMLD by 'equalising' the level of the masking signal (the noise) in each ear, then subtracting these from each other, 'cancelling' the masker. Predictions of binaural advantage to speech intelligibility of a model combining EC with a modulation-

frequency estimation block have been found to correlate strongly with human psychophysics (Cosentino *et al.*, 2014).

1.4.2 Azimuthal sound source localisation

The ILD-sensitivity of the LSO and ITD-sensitivity of the MSO are thought to underpin azimuthal sound source localisation. When a sound source is closer to one side of the head than the other, both time and level differences manifest between the ears. In guinea pig, ILDs can reach 40-dB and ITDs can reach 320-µs (Greene *et al.*, 2014). Behavioural evidence shows that the minimum perceptible ITD is 10-µs in humans (Mills, 1958) and 20-µs in cats (Wakeford & Robinson, 1974).

Neither ITDs nor ILDs adequately cover the entire range of audible frequencies as cues for sound localisation. Low-frequency sound waves diffract around the head to a much greater extent than high-frequencies, reducing the head-shadow and thus limiting the efficacy of ILD to higher frequencies. Conversely, ITD-sensitivity requires strongly phase-locked inputs, more effective at lower frequencies. In addition, the maximum possible ethological ITD decreases with decreasing head size. Note, however, that the maximum ITD perceptible by an animal is greater than that implied by the interaural distance, due to the diffraction around the face increasing the effective path length of the sound source to one ear (Kuhn, 1977). It is thus thought that ITDs are the predominant cue for low frequencies and ILDs for high frequencies, with the sense of azimuthal sound source localisation combining the two: the duplex theory (Lord Rayleigh, 1907).

1.4.3 Binaural pitch

Psychophysics has demonstrated various stimuli that are different in the two ears but produce a single pitch percept. These fall into two categories: stimuli comprised of harmonics and stimuli based on white noise. The fact that dichotic stimuli can evoke similar pitch percepts to monaural stimuli has been taken as evidence of a 'central pitch processor' (see **Section 1.4.3.2, page 28**) that receives binaurally-fused inputs.

Two of the simplest binaural-pitch evoking stimuli are dichotic complex tones (DCTs), a single harmonic in each ear, and dichotic HTCs, consisting of odd-only harmonics presented to one ear and even-only to the other. Both evoke a pitch percept equivalent to monaural presentation of their sum, and both were used in this Thesis.

Applying a frequency-shift to DCTs shifts the pitch percept equally to that observed in monaural-only presentation; this is discussed further in **Section 4.1.1 (page 103)**.

Presenting HTCs dichotically increases the spacing between harmonics at each ear, thereby increasing the resolvability. This might be expected to improve some aspects of pitch perception, however no benefit was seen in human psychophysics over monaural presentations; this is discussed further in **Section 4.1.2 (page 104)**.

Stimuli comprising BBN delayed between ears or comprising otherwise-diotic BBN with introduced interaural phase transitions can elicit weak sensations of pitch. Monaurally, these signals are indistinguishable from white noise and evoke no pitch percept. That IPDs in otherwise diotic white noise can give rise to a pitch percept suggests a possible involvement of the IPD-sensitive cells of the MSO and/or LSO underlying the pitch perception of these stimuli.

Huggins pitch (HP) (Cramer & Huggins, 1958) involves otherwise-diotic white noise with interaural decorrelation introduced over a narrow frequency band (see **Figure 9, page 26**).

The pitch percept resembles that evoked by a narrowband noise (NBN) centred at the same frequency (Gockel *et al.*, 2011). It has been suggested that HP arises as a form of unmasking, *i.e.* the phase differences between channels lead to the hearing-out of the phase-transition frequency band as a separate NBN, and this is where the pitch perception originates (Durlach, 1962; Culling, Summerfield, *et al.*, 1998).

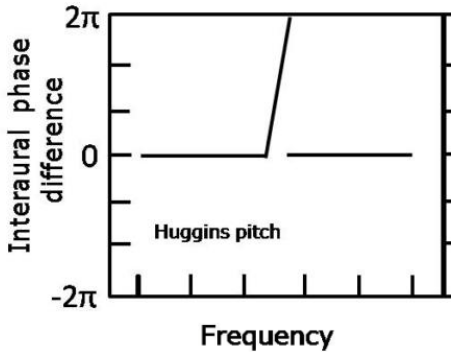

Figure 9 **Schematic of the interaural phase difference spectrum of HP showing the interaural phase transition**. The pitch percept occurs at the centre of the phase transition *i.e.* at the frequency with IPD = π-rads.

HP is, to date, the only dichotic white-noise-based pitch that has been used in neurophysiological studies; this is discussed in **Section 6.3.3 (page 176)**.

Dichotic repetition pitch (DRP) (Bilsen, 1972; Bilsen & Goldstein, 1974), is generated by presenting diotic white noise with a delay τ to one ear. This generates a pitch perception at $^1/_\tau$ Hz. DRP is analogous to IRN. Pitch percepts can results from delays (*e.g.* 31.25-Hz, 32-ms) larger than the bITD range observed in MSO neurophysiological studies and the ±5-ms range suggested in the m-EC model.

1.4.3.1 Binaural pitch models

In the optimum processor theory (Houtsma & Goldstein, 1971, 1972; Goldstein, 1973), spectrum analysis is performed on each ear (cochlear filtering) and the output passed into independent noisy channels (ANFs). The information is combined across ears in an 'optimum central processor', which estimates the F0.

In the central spectrum or central activity pattern model (Bilsen, 1977; Raatgever & Bilsen, 1986), signals from the two ears pass through a gammatone filterbank and responses from filters of the same frequency between ears are added together at varying ITDs. This leads to an activity pattern on a 2D axis of ITD and frequency; a pitch percept would appear as a single peak or a series of peaks on this pattern.

In the modified equalisation-cancellation (m-EC) model (Culling, Marshall, *et al.*, 1998; Culling, Summerfield, *et al.*, 1998; Culling, 2000), based on the original EC model (Durlach, 1963, 1972), signals are bandpass-filtered, then corresponding centre-frequency filters are equalised and cancelled at varying ITDs.

Predictions from mE-C model correlate most strongly with human psychophysical reports of HP (Culling, Marshall, *et al.*, 1998; Culling, Summerfield, *et al.*, 1998; Culling, 2000; Hartmann & Zhang, 2003). The pitch is thought to be detected from an interaural decorrelation: the signals in the majority of filters on each side are identical (zero-IPD) and cancel fully. However, filters with centre frequencies in the IPD transition region do not fully cancel at any given ITD. These therefore show the greatest response after cancellation.

The m-EC model fails to generate an output for DRP (Culling, Marshall, *et al.*, 1998). The central spectrum model, however, predicts that filtered BBN responses reinforce each other at centre frequencies that are integer multiples of $^1/_\tau$ ms, and so the central spectrum has peaks at integer multiples of $^1/_\tau$ ms

1.4.3.2 Central pitch processor

The 'central pitch processor' is a hypothetical structure that receives both monaural and binaural inputs and determines the pitch percept, irrespective of the form of pitch stimulus or whether it was presented monaurally or binaurally (Schouten *et al.*, 1962; Goldstein, 1973).

PDI occurs even if the interferer is presented contralaterally (Gockel, Hafter, *et al.*, 2009). This demonstrates limitations of the central pitch processor: possible differential processing of resolved and unresolved harmonics, and peripheral separation of dichotic signals cannot be used to aid segregation of the interferer from the signal.

Complex HP (Bilsen, 1977; Gockel, Carlyon, *et al.*, 2009), consisting of interaural phase transitions placed at integer multiples of an F0 similar to multiple phase-shift pitch (Bilsen, 1976), but without a transition at the F0 itself, also give rise to PDI (Gockel, Carlyon, *et al.*, 2009). The complex HP signal gave rise to a pitch percept at the F0. Additionally, an F0 percept has been shown to signals comprising 'harmonics' of a narrowband noise (NBN) and HP (Gockel *et al.*, 2011).

These studies combined suggest that interaural phase transitions are 'detected' as interaurally-decorrelated regions prior to pitch processing in the auditory system. The MSO, with its IPD-sensitivity, could be a candidate for this operation. Upstream of this, the central pitch processor would then receive a similar activity pattern for both NBN and HP.

1.5 THESIS OUTLINE

Chapter 2 (page 31) outlines the methods common to all of the projects in this Thesis, detailing the equipment used, experimental preparation, stimuli presented, and analyses performed on

recorded unit spike times. **Chapters 3 (page 45)**, **4 (page 101)**, and **5 (page 137)** are results Chapters. Each contains a separate introduction, methods, results and discussion section. An abstract is provided at the beginning, and a summary at the end, of each Chapter. **Chapter 6 (page 171)** is a general discussion on the results of this Thesis, highlighting implications for the field and potential future experiments.

Chapter 2

Methods

2.1 EXPERIMENTAL PROCEDURE

2.2 SURGERY

Data was obtained from 72 pigmented guinea pigs (*Cavia porcellus*), 39 males and 33 females, of weights 340-725g. In total, 58 experiments were conducted in the superior olivary complex (SOC) and 14 in the ventral cochlear nucleus (VCN). All experimental preparation, surgery and recordings were conducted in a single-walled sound-attenuating chamber (*IAC*).

Animals were anaesthetised with urethane (1.0-g/kg, *intraperitoneal*). For most (49) experiments, Hypnorm (*Vetapharma*, 1.0-ml/kg, *intramuscular*) was administered sufficient to abolish the pedal withdrawal reflex. For the remaining (23) experiments, where Hypnorm was not available, a combination of fentanyl (0.3-mg/kg, *intraperitoneal*) and midazolam (5-mg/kg, *intramuscular*) was administered. Subsequent urethane and Hypnorm, or urethane and fentanyl/midazolam, injections were given upon indication.

The trachea was cannulated and, in the event of respiratory distress, connected to a respiratory pump (*Bioscience UK*). Core body temperature was monitored via a rectal probe and maintained at 38°C using a heating blanket (*Harvard Apparatus*). An incision was made in the tragus on each side. Animals were positioned in stereotactic equipment and held in place using a pair of hollow Perspex speculae that fit into the ear canals. Skin and periosteum overlying the bullae were removed, and the skull levelled at 5-mm and 13-mm anterior of the interaural plane (Luparello *et al.*, 1964).

A small fenestration was made in each bulla, and a silver wire electrode was placed on the ridge of the round window of the cochlea. From this, the compound action potential (CAP) was recorded, reflecting the summed activity of auditory nerve fibres. Tone bursts, duration 10-ms, of frequencies 2-, 3-, 5-, 7-, 10-, 15- and 20-kHz, were presented in alternating phase; the resulting CAP waveform was averaged. An automated 3-down-1-up tracking algorithm was

used to find the threshold at each frequency. Hearing thresholds were not measured below 2-kHz as isolating the CAP from the cochlear microphonic, a summed potential of outer hair cells, becomes particularly difficult. Thresholds were compared with an average of normal-hearing guinea pigs obtained from our lab. Data from animals with raised CAP thresholds, defined as 10-dB above normal at any frequency, were not used. After verifying normal hearing thresholds in both ears, the bullae were then vented with a tube in order to equalise middle ear pressure (cat – Guinan & Peake, 1967); this was glued into place and any remaining fenestration was covered with *Vaseline*.

A craniotomy was made starting to the left of midline just above the lambdoidal suture; the extent was approximately 3-mm lateral and 2-mm anterior of the starting position. Dura overlying the exposed cerebellum was cut and reflected, and the cerebellum was partially aspirated in order to visualise the dorsal cochlear nucleus (DCN) and brainstem midline. Glass-insulated tungsten microelectrodes manufactured according to Merrill & Ainsworth, 1972 (*Ainsworth, Stasiak*) were inserted parasagittally into the left DCN at an angle of 45° to the horizontal under visual guidance. For SOC recordings, electrodes were initially inserted around 2- to 3-mm lateral of midline (see **Figure 10, page 34**). For VCN experiments, electrodes were inserted targeting the low-frequency anteroventral cochlear nucleus.

This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body. All procedures were carried out in accordance with a Home Office Project Licence; in addition, all experimenters were Personal Licence holders.

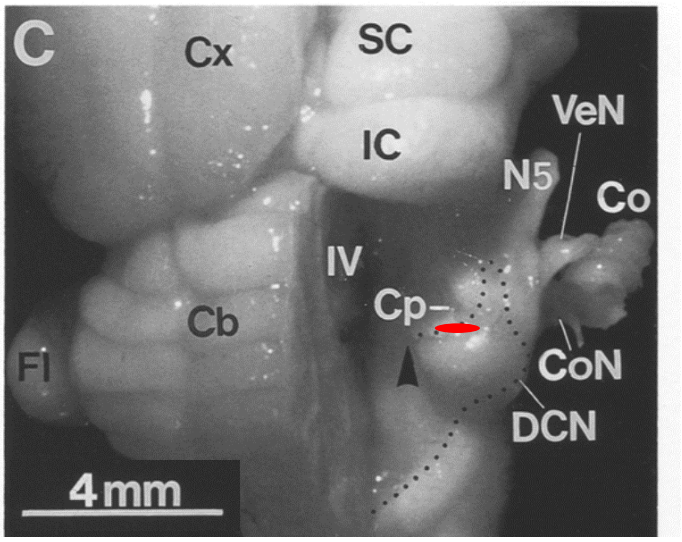

Figure 10 **Approximate electrode insertion sites for SOC experiments, shown in red**. Electrodes were inserted dorsoventrally through the DCN at a 45° angle to the horizontal. Electrodes were placed under visual guidance, with the initial track typically 2- to 3-mm lateral of the midline. Overlaid on figure from Hackney *et al.*, 1990. Note that this shows the right auditory brainstem; experiments were conducted exclusively in the left auditory brainstem. Figure labels: Cx: cerebral cortex; SC: superior colliculus; IC: inferior colliculus; VeN: vestibular nerve; Co: cochlea; CoN: auditory nerve; Fl: flocculus; Cb: cerebellum; IV: fourth ventricle; Cp: cerebellar peduncles; N5: trigeminal nerve.

2.2.1 Signal generation and recording

Digital signals for the stimuli were generated by NeuroSound (custom software, *Lloyd, 2001*) on a Dell workstation. All signals were generated at a sampling rate of 96-kHz at 24-bit. These were converted to analogue signals by an AD/DA converter (ADI-8 DS, *RME*). In the first 39 experiments, signals then passed through an equalizer (EQ 3600, *Apple Sound*); in the later 33 experiments, this was bypassed as it was not required[4]. Signals were amplified (RB971, *Rotel*), then passed through a variable in-line attenuator (custom build), before being

[4] Note that medial superior olive unit ITD-sensitivity, demonstrated by a sharp response to binaural beats at unit BF, and binaural pitch integrative responses, demonstrated by a response to the overall F0 of dichotic harmonic tone complexes, were observed under both configurations.

presented through *Radio Shack* 30-1777 speakers inserted into the speculae. A condenser microphone (4138, *Brüel & Kjær*) attached to a 1-mm probe tube, inserted into the speculae, was used to monitor presented signals within a few millimetres of the eardrum.

Signals from the microelectrode passed through an in-line preamplifier (custom build), before passing to a second amplifier (custom build), which amplified signals 1000-fold and bandpass filtered the signals between 0.3- and 10-kHz. Finally, analogue signals were converted to digital signals in the AD/DA, sampled at 96-kHz with 24-bit depth. Electrode recordings were converted into spike timings via an adjustable spike trigger online in NeuroSound; spike and stimulus waveforms could also be saved and reanalysed offline.

Recorded CAP signals were amplified 10000-fold using a *WPI* DAM 50 amplifier, which bandpass-filtered the signals between 0.3- and 3-kHz. Signals were then passed into the AD/DA and analysed using NeuroSound.

2.2.2 Localisation of unit recording sites

In SOC experiments, electrolytic lesions were created at two sites, the last unit recording site and with the electrode retracted 1-mm from this site, using a current stimulator (custom build) at 10- to 15-µA for 10-seconds. The two lesions were also used to reconstruct electrode tracks. Animals were subsequently perfused using ~200-mls of saline followed by ~200-mls phosphate-buffered 1%-glutaraldehyde/3%-formaldehyde fixative solution. Brains were removed and immersed in the fixative solution until further histological processing.

Prior to embedding, brains were placed in phosphate-buffered sucrose solution for 1-day, followed by gelatine/egg albumen/phosphate-buffered sucrose solution for 2- to 4-days. Brains were then embedded with addition of 1-ml of 25%-glutaraldehyde. Using a rotary freezing

microtome, brains were sectioned parasagittally into sections of 50-µm thickness and mounted onto subbed slides.

After drying, sections were stained using a cresyl fast violet staining protocol. This consisted of defatting with xylene, ethanol baths of decreasing concentrations, cresyl fast violet staining in aqueous solution, ethanol baths of increasing concentrations, and finally defatting once more in xylene. Slides were coverslipped with DePeX. After examination and identification of lesions in sections using a *Zeiss* Axioscop 2 microscope, sections were scanned using a digital slide scanner (*Hamamatsu* NanoZoomer 2.0-RS C10730).

2.3 STIMULI

2.3.1 Search stimulus

To search for neurons sensitive to auditory signals, 50-ms broadband noise (BBN) bursts, bandwidth 0- to 20-kHz, were presented at 60 to 70 dB-sound pressure level (SPL), separated by 200-ms gaps. Evoked potentials at the onset of the BBN indicated the electrode was close to an auditory nucleus: onset activity is the most prevalent across all unit types. The electrode was subsequently advanced or retracted in small steps (0.5- or 1.0-µm) using a hydraulic micropositioner (Model 650 W, *Kopf Instruments*) until a unit was isolated. Isolation was defined as a minimum signal to noise ratio (SNR) of 2:1, but was typically greater than 4:1 (n = 780/826 units).

2.3.2 Unit characterisation

Units were characterised using the auditory threshold and best frequency (BF) – the frequency at which the neuron is most sensitive. This was initially determined audiovisually using 50-ms

pure tones and refined using analysis of the receptive field (see **Section 2.3.3, page 37**). The unit spontaneous discharge rate (SR) was then measured over 10-seconds in the absence of controlled stimulation.

To enable unit type classification, 250 repetitions of a 50-ms BF pure tone with randomised starting phase were presented to each unit. Stimuli were gated with 1-ms $\cos^2$ ramps. Presentation could be varied: ipsilateral-only, contralateral-only, or diotic. Randomising the starting phase allows the differentiation of low-frequency primary-like (PL) units, which would otherwise phase-lock, from low-frequency choppers, which show periodicity in their firing at a chopping frequency independent of their BF (see **Figure 67, page 150**), on the basis of their peri-stimulus time histogram (PSTHs) (see **Section 2.4.1, page 41**).

2.3.3 Receptive fields

Receptive fields were measured using 50-ms pure tones, gated with 5-ms $\cos^2$ ramps. For neurons with BF<5-kHz, the frequency range was 3-octaves below unit BF to 2-octaves above; for units with BFs 5-kHz or above, tone frequencies ranged from 2-octaves below unit BF to 1-octave above. Frequencies were sampled in 0.1-octave steps; levels were sampled in 5-dB steps ranging from ~10-dB below unit threshold up to a maximum of ~94 dB-SPL. Usually, each stimulus combination was presented only once; occasionally, each stimulus was presented 5 or 10 times, and spike rates were averaged across all presentations.

For binaural neurons, receptive fields were measured to ipsilateral-only, contralateral-only and diotic presentations at zero-interaural time difference (ITD). For lateral superior olive units with low-SRs, the contralateral inhibitory receptive field was measured while a BF pure tone was presented to the excitatory ear at 5- to 10-dB above threshold; this paradigm is termed a 'tickle tone receptive field'.

Receptive fields were then fitted with a 10th-order polynomial to responses 2-standard deviations above the unit SR (see **Figure 11** below). Fitted BF and threshold were taken from the minimum of this curve. If the algorithm was not able to adequately fit the data, or no receptive field was measured, online estimates of BF and threshold were used instead.

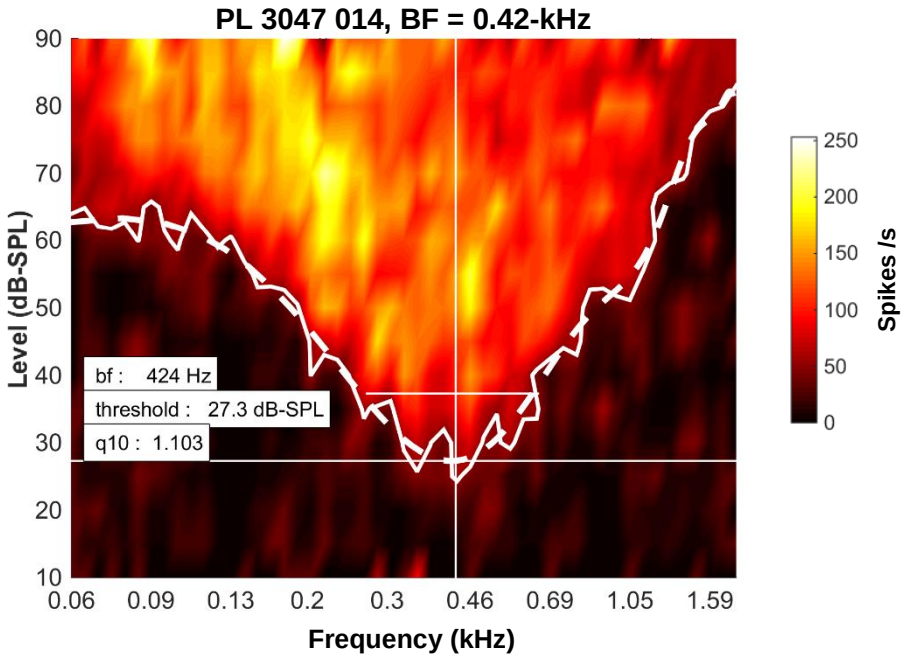

Figure 11 **Example of a fitted receptive field**. An estimate of threshold at each frequency was found using a criterion level, typically two standard deviations above unit spontaneous rate. A 10th-order polynomial was then fitted to this, shown as a dotted white line. Fitted unit BF is the frequency that the unit is most sensitive to, in this case 424-Hz; fitted threshold corresponds to threshold at BF. $Q_{10\text{-dB}}$ is then measured from this fit, 10-dB above threshold, as shown above. The online BF estimate was 457-Hz, a 7.8% difference; the online threshold estimate was 26.8 dB-SPL, a 1.9% difference.

To compare the responses evoked by monaural and diotic stimulation, a contiguous area was defined in the receptive field with strongest overall response. Spike counts were summed across this area, and compared with spike counts in corresponding areas in receptive fields obtained from other presentations. The methodology is shown in **Figure 12 (page 40)** using an example medial superior olive (MSO) unit. These were calculated instead of merely

reporting whether MSO units showed responses to monaural tonal stimulation ('[E·E]'), as has been done in other studies, as this does not assess binaural facilitation. At the worst ITD, for instance, MSO unit responses are weaker than in monaural-only stimulation, even below spontaneous rate (cat – Yin & Chan, 1990). It also does not account for the relative contributions of each side's input *i.e.* ear dominance. However, the MSO population of **Section 3.3.5 (page 87)** contains units that show facilitation even when stimulated diotically, *i.e.* away from their best ITD.

Four metrics were computed using the resulting spike counts:

- Monaural to diotic ratio, used in monaural CN and SOC units. A value of greater than 1 indicates binaural suppression.
- Facilitation index, $\boldsymbol{FI} = \frac{\boldsymbol{diotic}}{\boldsymbol{max(ipsi,contra)}}$. An FI of greater than 1 indicates that the diotic response exceeded the response of the best driven monaural filter.
- Relative binaural asymmetry = $\frac{\boldsymbol{ipsi-contra}}{\boldsymbol{ipsi+contra}}$. This metric ranges from ±1, with values closer to 1 indicating that the ipsilateral filter had greater responses than the contralateral, and *vice versa*. If the filters were of equal strength, this metric equalled zero.
- Absolute binaural asymmetry, the modulus of the absolute binaural asymmetry metric. This tested for particular ear dominance across the MSO population.

Statistical significance of the differences in distribution of these metrics between nuclei was assessed using ANOVA analysis in MATLAB.

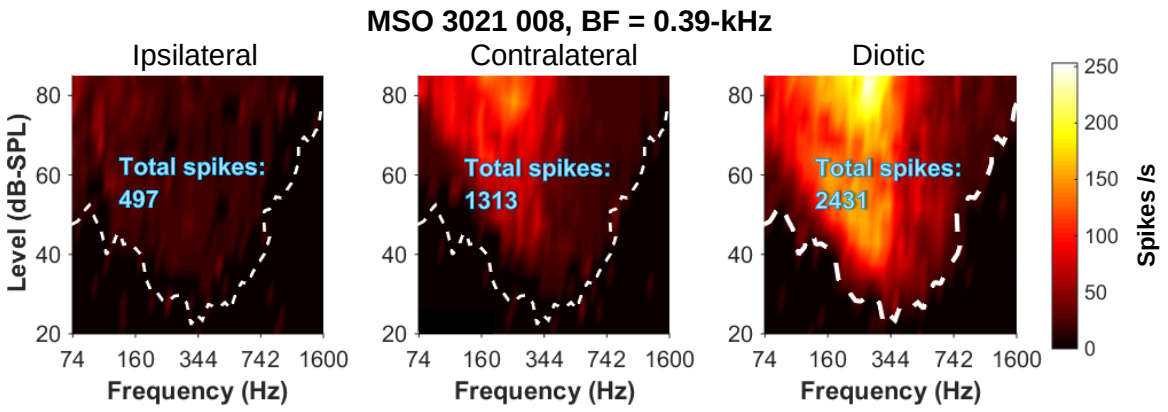

Figure 12 **Receptive field analysis for an MSO unit**. A contiguous region in the receptive field with greatest response (diotic for MSO units) was defined, shown by the dotted line, where spike rates exceeded two standard deviations above SR. Spike counts in this region were summed and compared to the sum from the same region in monaural presentations. Receptive field metrics are then calculated as follows:

$\boldsymbol{FI} = \frac{diotic}{max(ipsi,contra)} = \frac{4862}{1313} \approx 1.85$, *i.e.* the diotic response is 1.85 times the maximum of the monaural filters, in this case the contralateral filter.

Relative binaural asymmetry: $\frac{ipsi-contra}{ipsi+contra} = \frac{497-1313}{497+1313} \approx -0.45$, *i.e.* the contralateral response is greater than the ipsilateral response.

Absolute binaural asymmetry: $\left|\frac{ipsi-contra}{ipsi+contra}\right| = \left|\frac{497-1313}{497+1313}\right| \approx 0.45$, *i.e.* this unit has an asymmetry in the spike rates elicited by monaural stimulation.

2.3.4 Stimulus calibration

Speakers were calibrated *in situ* using white noise and a *Brüel & Kjær* 4138 microphone. **Figure 13** below shows the magnitude spectrum of each speaker.

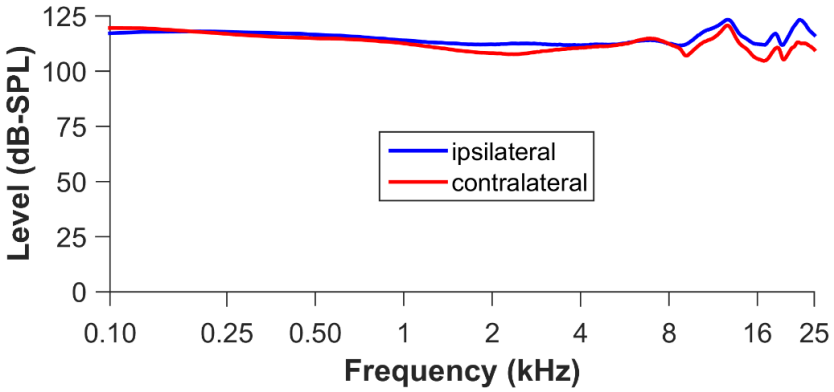

Figure 13 **System average magnitude spectra measured *in situ*.** These have been scaled to system maximum output for pure tones; the maximum level used in experiments was ~20-dB below this. Between 0.1- and 2-kHz, interaural level differed by less than 5-dB.

Complex stimuli were calibrated offline using a *Brüel & Kjær* 4134 microphone in an artificial coupler. Interaural phase was calibrated offline (*i.e.* not *in situ*) using 100-ms pure-tone bursts. Between 100-Hz and 2000-Hz, there was generally a small ipsilateral phase lead of mean±standard deviation: 0.010 ± 0.035 cycles.

2.4 ANALYSES

2.4.1 Spike histograms

PSTHs were constructed using a bin-width of 200-µs.

Period histograms were computed by normalising recorded spike times by the stimulus period. Synchronicity to the period of a binaural beat, or to individual carrier frequencies, could be assessed using vector strength (VS) analysis of period histograms (Goldberg & Brown, 1969). Fixed-phase pure tones were presented either monaurally or diotically. If the spike-timing jitter was close to zero, VS approached 1. Else, if the firing pattern was completely random across stimulus presentations, VS equalled zero. Statistical significance was determined using the Rayleigh test of uniformity with a $p<0.001$ criterion (Buunen & Rhode, 1978).

Entrainment indices were defined as the fraction of periods that evoked at least one spike, using the period of either the pure tone frequency or the modulation frequency.

2.4.2 Spike waveform analysis

During unit recordings, if the signal-to-noise ratio decreased to a point where units were not adequately isolated, the unit was counted as lost and contaminated data discounted. However, a spike sorting algorithm (Quiroga *et al.*, 2004) was used to analyse spike failure units in the

VCN, to confirm that no second action potential waveform shape or violations of the refractory period occurred, and to generate average spike waveforms. The original software was implemented into a MATLAB GUI by postdoc colleague Arkadiusz Stasiak; all analyses using the GUI were carried out by myself.

Saved electrode recordings were low-pass filtered at 5-kHz. Spikes were triggered from vertical crossings of an adjustable threshold. Spike waveforms were centred according to the action potential peak height, then fed into the spike sorting algorithm.

A wavelet transform was performed on the spike waveforms, using 64 Haar wavelets, to extract features of the spike waveforms. The top ten wavelet coefficients that show the greatest evidence of a multimodal distribution are selected, using a Kolmogorov–Smirnov test for normality. A superparamagnetic clustering algorithm then separates the ten-dimensional wavelet coefficient space into separate clusters – these comprise the different spike waveform shapes.

Figure 14 (page 43) shows spike sorting analysis of a single MSO unit's responses to binaural beats at unit BF. In **Figure 14A**, an example electrode recording is shown. Spikes were triggered when crossing the red line. The presence of the stimulus is represented by a black bar. Wavelet coefficient distributions for triggered spikes are all unimodal, reflecting a single cluster and single spike waveform shape. **Figure 14B** depicts the coefficient with largest deviation from normality according to a Kolmogorov-Smirnov test. This particular Haar wavelet did not contribute greatly to spike waveform shapes as the wavelet coefficients are small. In **Figure 14C** an average spike waveform of the cluster is shown. Spike amplitude did vary, however all spikes belong to a single analysed cluster. As all triggered spikes belong to the same cluster, and there are no violations of the refractory period (not shown), all spikes originated from a single, well-isolated MSO unit.

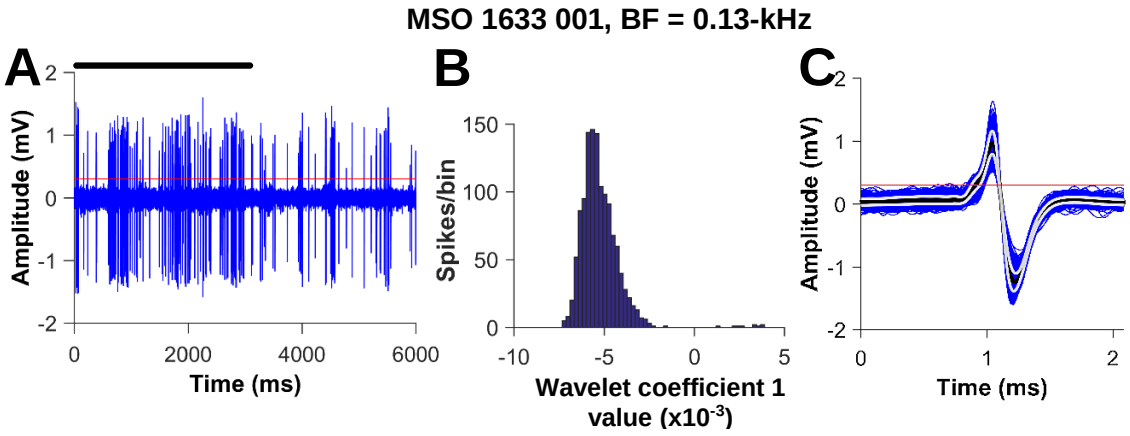

Figure 14 **Spike sorting the recordings of an isolated single unit produces a single cluster**.

A: Electrode recording of a single MSO unit's response to binaural beats at unit BF. Stimulus is represented by a black bar.

B: All wavelet coefficient distributions are unimodal; plotted is the coefficient with largest deviation from normality according to a Kolmogorov-Smirnov test.

C: Average spike waveform of the cluster, in black, plotted on top of all triggered spikes.

Conversely, if two different spike waveform shapes, are present, the software separates the data into two clusters. The only unit type for which spike sorting of electrode recordings showed multiple clusters was the PL with prepotentials [PL(PP)] unit type in the VCN. This is due to spike failures in these units (see **Section 5.1.1.1, page 140**). **Figure 15** (**page 44**) shows spike sorting of responses of a PL(PP) unit to pure tones at BF. In **Figure 15A**, an electrode recording of a single stimulus presentation is shown. There are two discrete sets of amplitude distributions. In **Figure 15B**, waveforms triggered at the red line in **A** are sorted into two clusters, shown in red and blue. The waveforms for each cluster of spike waveform shapes are shown in **Figure 15C** and **D**. Note the similarity in the initial segments in these two waveforms: this, along with the absence of violations of the refractory period (not shown), support the interpretation that this activity reflects responses of a single unit. The larger cluster of **Figure 15C** contains prepotential (PP)-excitatory postsynaptic potential (EPSP) events that do not result in action potentials; this comprises the majority of triggered events. The cluster

shown in **Figure 15D** consists of PP-EPSP events that elicited action potentials. Distributions of wavelet coefficients show bimodal distributions, reflecting two distinct clusters. In **Figure 15E**, the coefficient with greatest deviation from normality shows a clear bimodal distribution.

All units presented in this Thesis, in particular those from the MSO, are well-isolated single unit recordings that do not require spike sorting.

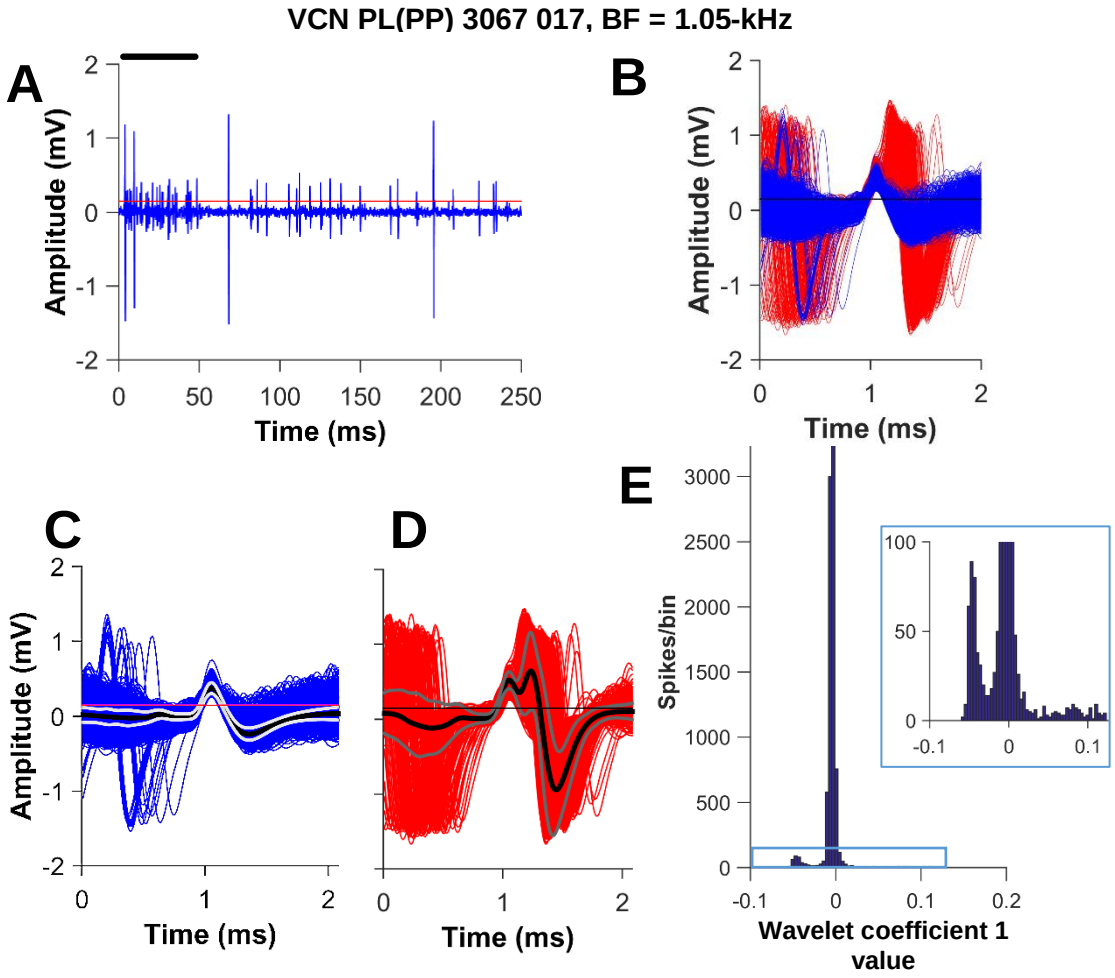

Figure 15 **Exemplar VCN PL(PP) unit showing multiple clusters due to spike failure**.

A: An electrode recording of unit responses to BF short tone bursts, represented by the black bar.

B: Waveforms triggered at the red line in **A** are sorted into two clusters, shown in red and blue.

C: PP- EPSP events that did not result in action potentials.

D: PP-EPSP events that elicited action potentials.

E: An example wavelet coefficients showing a bimodal distribution, reflecting two distinct clusters. The scale has been magnified in the inset plot to better see the two peaks.

Chapter 3

Basic response properties of single units in the superior olivary complex

Abstract

Recordings were made from presumed cells of the SPN, MNTB, LSO and MSO. Responses of cells of these nuclei have been extensively characterised in other species, however, to my knowledge, the responses of the guinea pig SOC have yet to be characterised. This Chapter compared the basic response properties of guinea pig SOC units with those reported in the SOC of other mammals.

Responses to pure tones at unit BF, plotted as PSTHs, were used to form an initial classification, in-keeping with studies of the CN and IC. SPN units show offset-chopper responses to stimulation with BF pure tones, and showed sensitivity to gaps within tones. MNTB units showed PL or PN responses to BF pure tones and the same three-component spike waveform shape reported in other species. LSO units were sensitive to interaural level differences of BF tones. MSO units responded weakly to monaural stimulation but more strongly to diotic stimulation, with many units showing binaural facilitation. IPD-sensitivity of low-BF binaurally-responsive cells of the LSO and MSO was confirmed using binaural beat stimuli.

These responses show close similarities to those reported of other mammals and supports the use of the guinea pig SOC as a model for studying binaural interactions.

3.1 INTRODUCTION

This Chapter focusses on the responses of cells from the four main nuclei of the superior olivary complex (SOC) with ascending projections (guinea pig – Saint Marie & Baker, 1990; guinea pig – Schofield & Cant, 1992): two monaural periolivary nuclei, the medial nucleus of the trapezoid body (MNTB) and superior paraolivary nucleus (SPN), and the two binaural principal nuclei, the medial superior olive (MSO) and lateral superior olive (LSO). These nuclei are highlighted in **Figure 16** below.

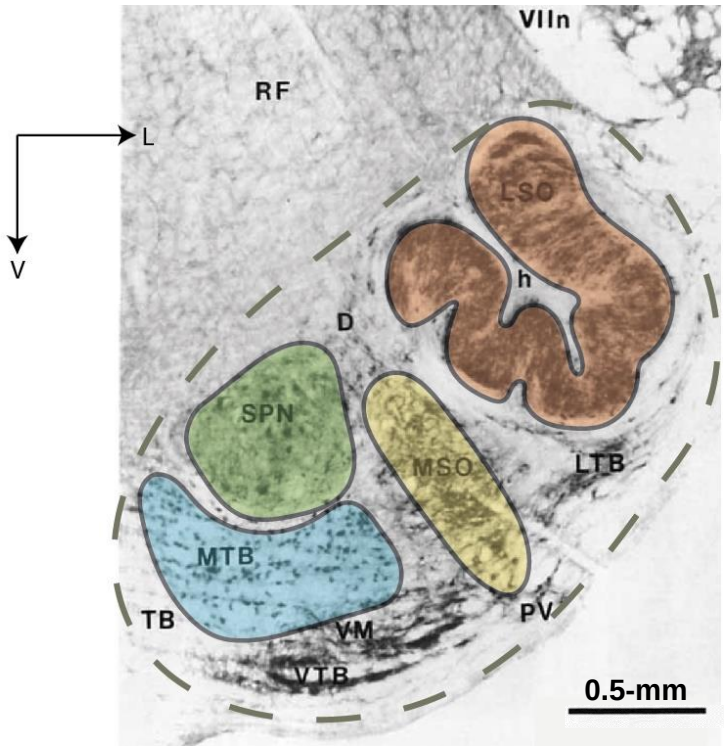

Figure 16 **The guinea pig superior olivary complex in transverse section**. SOC is demarcated by the dotted grey line. Recordings were made from the four shaded SOC nuclei; note that not all SOC nuclei are present in this section. The LSO is the U-shaped nucleus highlighted in orange. The MSO, highlighted in yellow, is a sheet-like nucleus, with cell bodies in the central plane and dendrites extending medially and laterally. The largest of the periolivary nuclei are the SPN (highlighted in green) and MNTB (highlighted in blue). Section adapted from Schofield & Cant, 1991.

Labels: h = hilus of LSO; MTB = medial nucleus of the trapezoid body; RF = reticular formation; VIIn = facial nucleus; D = dorsomedial periolivary nucleus; LTB = lateral nucleus of the trapezoid body; PV = posteroventral periolivary nucleus; VM = ventromedial periolivary nucleus; VTB = ventral nucleus of the trapezoid body; TB = trapezoid body. L = lateral, V = ventral.

To start, I review the properties of these four nuclei using data from other species; no data recorded *in vivo* from the guinea pig SOC have been published. In the Discussion of this Chapter, I compare data recorded in the guinea pig to the data from other species.

3.1.1 Medial nucleus of the trapezoid body

Cells of the MNTB receive glutamatergic input (mouse – Wu & Kelly, 1992) from globular bushy cells (GBCs) in the contralateral ventral cochlear nucleus (VCN) (cat – Smith *et al.*, 1991). The synapse between GBCs and MNTB cells is highly specialised and among the largest in the brain, known as the calyx of Held (cat – Held, 1893; guinea pig – Robertson & Winter, 1988; guinea pig – Thompson & Thompson, 1991; rat – Forsythe & Barnes-Davies, 1993; gerbil – Kil *et al.*, 1995). In mice, there is evidence of multiple calyceal inputs to each MNTB cell (anatomy – Bergsman *et al.*, 2004; physiology – Matho, 2013).

The MNTB also receives inhibitory GABAergic (cat – Adams & Mugnaini, 1990) and glycinergic (gerbil and mouse – Kuwabara & Zook, 1991) inputs. GABAergic inputs may arise from the ventral nucleus of the trapezoid body (VNTB) (guinea pig – Helfert *et al.*, 1989) and SPN (rat – Kulesza & Berrebi, 2000). Glycinergic input to the MNTB arises, at least in part, from collaterals of the MNTB itself (cat – Smith *et al.*, 1998). However, evidence of resulting inhibitory post synaptic potentials (IPSPs) or inhibition-induced spike failure has not been found (cat – Guinan & Li, 1990). Inhibition does not underpin the adaptation to sustained or repeated tonal stimulation seen in MNTB (mouse – Lorteije & Borst, 2011). In addition, it has been shown that glycinergic input, rather than depressing MNTB output activity, enhances MNTB cell neurotransmitter release (rat – Turecek & Trussell, 2001).

As with the endbulb of Held synapse in the VCN (see **Section 5.1.1.1, page 140**), MNTB extracellular spike waveform recordings can show three components (cat – Guinan & Li, 1990; cat – Smith *et al.*, 1998; gerbil – Kopp-Scheinpflug *et al.*, 2003, rat and mouse – Kopp-Scheinpflug *et al.*, 2008): prepotentials comprising presynaptic activity ('PP' component), excitatory postsynaptic potentials (EPSPs, 'A' component) and action potentials ('B' component). Depending on the species, EPSPs either sometimes fail to evoke spikes (gerbil – Kopp-Scheinpflug *et al.*, 2003), fail in up to half of all units (mouse – Lorteije *et al.*, 2009), or never fail to evoke spikes (cat – Mc Laughlin *et al.*, 2008; guinea pig – Stasiak *et al.*, unpublished). Spike failure has been shown to increase temporal enhancement of output over input (gerbil – Kopp-Scheinpflug *et al.*, 2003), as in the spherical bushy cells (SBCs) of the VCN (see **Section 5.1.1.1, page 140**).

Output fibres of MNTB cells are glycinergic (guinea pig, present in cell bodies – Wenthold *et al.*, 1987; guinea pig, descending projection to CN – Benson & Potashner, 1990; guinea pig, projection to LSO – Bledsoe *et al.*, 1990). They provide strong inhibitory inputs to ipsilateral SOC nuclei (see **Figure 17, page 50**; see **Sections 3.1.2, page 51**; **3.1.3.1, page 53**; and **3.1.4.1, page 61**), and also to the ipsilateral ventral nucleus of the lateral lemniscus (cat – Spangler *et al.*, 1985).

It was previously thought that humans do not have an MNTB, implying a fundamentally different type of SOC architecture compared to other mammals. This has been rebuked upon careful reconsideration of the previous anatomical evidence (Kulesza & Grothe, 2015). However, bilateral genetic ablation of the MNTB via a Cre-loxP system does not eliminate the ability to localise sounds; inhibitory inputs to other SOC nuclei were replaced by other unknown sources (mouse – Jalabi *et al.*, 2013). Additionally, interaural level difference (ILD)-sensitivity is maintained in the inferior colliculus (IC) after bilateral kainic acid lesions of the SOC (rat – Li

& Kelly, 1992); this suggests that the LSO and, indeed, inhibitory output of the MNTB that gives rise to its ILD-sensitivity, are not required for ILD-sensitivity in the IC. Additional evidence of this is discussed in **Section 1.2.6 (page 11)**.

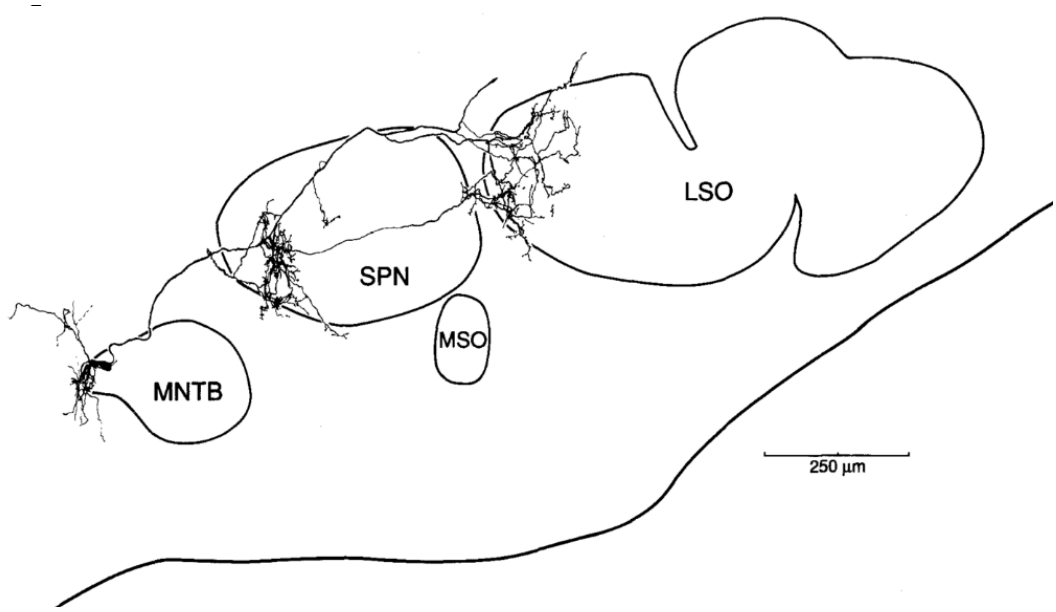

Figure 17 **A labelled rat MNTB cell showing extensive innervation of the ipsilateral SPN and LSO**. Cells were found that innervated the MSO as well. From Banks & Smith, 1992.

Resulting from their near one-to-one GBC input, MNTB cells show a primary-like (PL) or primary-like with notch (PN) response to short tone bursts at unit best frequency (BF) (cat – Smith *et al.*, 1998; gerbil, rat and mouse – Kopp-Scheinpflug *et al.*, 2008). Cells show [0·E][5] responses, originating from contralateral input from [E·0] GBC(s).

Reflecting their high-synchrony input from GBCs (see **Section 5.1.1.1, page 140**), MNTB cells show enhanced temporal encoding compared to auditory nerve fibres (ANFs) (cat – Smith *et*

[5] The convention used in this Thesis is '[ipsilateral response]·[contralateral response]', where 'E' is excitatory, 'I' is inhibitory, and '0' is no effect.

al., 1998; gerbil – Kopp-Scheinpflug *et al.*, 2003; gerbil – Dehmel *et al.*, 2005; cat – Tollin & Yin, 2005).

The MNTB comprises an important input to the other SOC nuclei, giving rise to ILD-sensitivity in the LSO and thought to refine the interaural time difference (ITD)-sensitivity in MSO. The MNTB also sends projections to the SPN (see **Figure 17, page 50**), the subject of the next section.

3.1.2 Superior paraolivary nucleus

SPN units are thought to underpin gap detection (rat – Kadner *et al.*, 2006; rat – Kadner & Berrebi, 2008). SPN cells do not respond to gaps in BF pure tones until the gap exceeds a gap detection threshold (GDT). Human GDTs are around 5-ms for gaps in pure tones (Shailer & Moore, 1987). SPN units with GDTs close to this value have been shown (rat – Kadner *et al.*, 2006; rat – Kadner & Berrebi, 2008).

In response to short tone bursts at BF, SPN cells fire a burst of spikes after tone offset (gerbil – Behrend *et al.*, 2002; gerbil – Dehmel *et al.*, 2002; rat – Kulesza *et al.*, 2003; mouse – Felix *et al.*, 2011). This is caused by a rebound excitation from the profound glycinergic inhibition (rat – Kadner & Berrebi, 2008) received from the ipsilateral MNTB (cat DMPO – Morest, 1968; rat – Banks & Smith, 1992)[6]. When tones are interrupted by very short gaps, SPN units do not respond as the inhibition does not subside enough for a rebound excitation to occur. The rebound excitation effect is increased with increasing build-up of inhibition, as would occur with

[6] The dorsomedial periolivary nucleus (DMPO) is thought to be a homologue of the rodent SPN present in other mammals (Grothe & Park, 2000).

longer duration tones. It has been proposed that human sensitivity to sound duration (Creelman, 1962; Abel, 1972) might be underpinned by the increasing number of offset spikes SPN units produce in response to increasing sound durations (rat – Kadner *et al.*, 2006).

The main current responsible for rebound excitation is hyperpolarisation-activated current (I_h) arising from hyperpolarisation-activated cyclic nucleotide-gated (HCN) cation channels, with hyperpolarisation-activated calcium ion channel-mediated current (I_{TCa}) increasing the number of spikes per rebound excitation (mouse – Kopp-Scheinpflug *et al.*, 2011). Other ion channels also play a role: gap detection thresholds were decreased when a drug that positively modulates Kv3 was administered (rat – Rybalko *et al.*, 2014).

Extensive GABAergic collateral inhibition from other SPN cells and autoinhibition (rat – Kulesza & Berrebi, 2000) also contributes to the abolition of peri-stimulus responses (rat – Kulesza *et al.*, 2007). The slower time course of GABAergic inhibition is thought to enhance firing to low modulation rates but suppress the response to higher modulation rates, setting a lower limit on modulation rates that SPN cells can follow (rat – Kadner & Berrebi, 2008). In the guinea pig, SPN cells also receive bilateral inputs from octopus and multipolar cells in the VCN, predominantly from the contralateral VCN (Schofield, 1995); however, in the mouse, a recent paper using anatomical tracers found evidence for only a single excitatory input, arising from octopus cells in the contralateral VCN (Felix II *et al.*, 2017).

Separate classes of SPN cells have been described (guinea pig – Schofield, 1991). Round, multipolar cells project to both IC (rat – Saldaña *et al.*, 2009; rat – Felix *et al.*, 2014), with ipsilaterally-projecting cells using GABA as a neurotransmitter; the identity of the neurotransmitter of contralaterally-projecting cells is not known (guinea pig – Schofield, 1991). Cells that project to the ipsilateral IC are selectively contacted by collaterals from MSO cells (gerbil – Kuwabara & Zook, 1999). Small, elongated cells project to the cochlear nuclei, with

ipsilaterally-projecting cells using glycine and/or GABA as a neurotransmitter; contralaterally-projecting cells are thought to use other neurotransmitter(s) (guinea pig – Schofield, 1991). In addition, it has been demonstrated that there is a substantial projection from the ipsilateral SPN and, to a lesser extent, from the contralateral SPN, directly to the medial geniculate body (MGB), either bypassing the IC or projecting to both the IC and MGB (guinea pig – Schofield *et al.*, 2014).

Next, I discuss the two principal ascending nuclei of the SOC, the MSO and LSO. The responses of these nuclei are thought to underpin azimuthal sound source localisation. The vast majority of physiological studies on MSO and LSO have focussed on their ITD- and ILD-sensitivity and mechanisms affecting these properties; these studies are reviewed in the following sections. However, as the MSO and LSO receive input from both sides of the auditory system, they are good candidates for other binaural integrative computations such as binaural pitch. This is investigated in **Chapter 4 (page 101)**.

3.1.3 Medial superior olive

MSO cells have a fusiform geometry with dendrites extending medially and laterally; cell bodies are located in the mediolateral and dorsoventral centre of the MSO, along the full rostrocaudal extent (guinea pig – Schofield & Cant, 1991; guinea pig – Smith, 1995).

3.1.3.1 Inputs and response properties

In low-frequency-hearing mammals, a majority of low-BF MSO units are binaural[7] and sensitive to ITDs (cat – Hall, 1965; dog – Goldberg & Brown, 1969; chinchilla – Langford, 1984; cat – Yin & Chan, 1990; gerbil – Spitzer & Semple, 1995; rabbit – Batra *et al.*, 1997a, 1997b; Mexican free-tailed bat – Grothe & Park, 1998; guinea pig IC – McAlpine *et al.*, 2001; rabbit – Fitzpatrick & Kuwada, 2001; gerbil – Seidl & Grothe, 2005; gerbil – Pecka *et al.*, 2008; gerbil – Pecka & Siveke, 2010; gerbil – Day & Semple, 2011; gerbil – Stange *et al.*, 2013; gerbil – van der Heijden *et al.*, 2013; chinchilla LL – Bremen & Joris, 2013; gerbil – Franken *et al.*, 2015; gerbil – Plauška *et al.*, 2016)[8]. They act as coincidence detectors of the excitatory, glutamatergic, phase-locked inputs from 4-8 (gerbil – van der Heijden *et al.*, 2013) SBCs in both anteroventral cochlear nuclei (AVCN) (cat – Smith *et al.*, 1993).

MSO cell EPSP duration varies depending on input strength, varying between 400-µs and 2-ms (guinea pig – Smith, 1995). Cells show a voltage-dependent conductance (guinea pig – Smith, 1995), resulting from a low-voltage-activated potassium current (I_K) (gerbil – Franken *et al.*, 2015) (see **Section 1.2.5.1, page 9**). It has been suggested that I_K electrically insulates the soma and dendrites, where binaural integration take place, from distortions from currents caused by action potential initiation (gerbil – Scott *et al.*, 2005). Cells also have little to no spontaneous rate (SR) (guinea pig – Smith, 1995).

Due to a difference in relative arrive time of the inputs to a given MSO cell from each side of the auditory system, coincidence requires stimuli to be presented with a compensatory ITD, the best ITD (bITD). The ITD-sensitivity of the MSO at low frequencies is thought to underpin

[7] It is likely that these units are binaurally-facilitatory but this is generally not assessed.

[8] MSO data is not available for all species; when proxy data from other nuclei have been recorded, the structure and method (if appropriate) will be delineated. LL = lateral lemniscus.

the sense of sound-localisation (Lord Rayleigh, 1907; Jeffress, 1948; Goldberg & Brown, 1969).

The proportion of monaural units is quite high in some studies (dog – Goldberg & Brown, 1969; cat – Yin & Chan, 1990; gerbil – Brand *et al.*, 2002), but zero in others (chinchilla LL – Bremen & Joris, 2013). Sensitivity to envelope ITDs in high-BF MSO cells has been shown (cat – Yin & Chan, 1990), however, due to the use of binaural beat search stimulation, high-BF MSO units are likely to be undersampled.

In some species, the MSO is entirely monaural and thus ITD-insensitive. The MSOs of these species act as a low-pass filters for envelope amplitude modulation (AM) (mustached bat – Grothe & Neuweiler, 2000). MSO cells are thought to detect self-generated AM in frequency sweeps interacting with wings of the prey (Grothe, 2000).

High-frequency-hearing mammals also can show ITD-sensitivity to AM imposed on high-frequency carriers (free-tailed bat – Grothe & Park, 1998; free-tailed bat – Grothe & Neuweiler, 2000), similar to high-frequency MSO units in low-frequency hearing mammals (cat – Yin & Chan, 1990).

bITD can be assessed by presenting pairs of pure tones with varying ITD (dog – Goldberg & Brown, 1969) or by binaural beat stimulation, where the interaural phase difference (IPD) changes continuously (cat – Yin & Chan, 1990). bITD varies across the population of MSO cells, resulting from differences of relative signal arrival time from each side of the auditory system. The mechanism(s) underpinning this input lag have been vigorously debated (see **Section 1.2.5.1, page 9**).

MSO cells receive glycinergic inputs from cells of the ipsilateral MNTB (guinea pig – Smith, 1995; cat – Smith *et al.*, 1998) (see **Section 3.1.1, page 48**) and ipsilateral lateral nucleus of

the trapezoid body (LNTB) (gerbil – Cant & Hyson, 1992; guinea pig – Smith, 1995; cat – Spirou & Berrebi, 1997). The time courses of these inputs are similar in spite of the contralateral origin of MNTB input (gerbil – Roberts *et al.*, 2014). The role of these inputs in ITD-sensitivity has been debated (see **Section 1.2.5.1, page 9**). Ultimately, the responses of the MSO can be simulated well by a single excitatory input from each side (Colburn *et al.*, 1990).

MSO nonprincipal cells have been described in a single study (guinea pig – Smith, 1995). These do not show a fusiform geometry, with dendrites instead branching randomly. They do not show the voltage-dependent I_K seen in MSO principal cells show have larger SRs.

Little-to-no data exist about the basic response properties of MSO cells, such as frequency tuning curves and responses to BF tones. The few tuning curves that have been published show monaural V-shaped filters (dog – Goldberg & Brown, 1969; cat – Yin & Chan, 1990).

3.1.3.2 Binaural properties

MSO cells act as coincidence detectors, responding maximally when EPSPs arrive simultaneously at the cell (model – Jeffress, 1948; cat – Yin & Chan, 1990; gerbil – van der Heijden *et al.*, 2013; gerbil – Plauška *et al.*, 2016). Coincidence is highly tuned to a particular interaural phase, the best phase (BP) for a given frequency of pure-tone, and a bITD for wideband stimuli.

Binaural responses can be predicted by applying a coincidence window to recordings of MSO unit responses to monaural stimuli (cat – Yin & Chan, 1990; gerbil – van der Heijden *et al.*, 2013; gerbil – Plauška *et al.*, 2016). SBC inputs elicit mostly subthreshold EPSPs that sum linearly within MSO cells. Weak, if any, responses result from monaural stimulation of either

ear, however, substantial responses result when stimulated binaurally; this has been described as a nonlinear spiking probability (gerbil – van der Heijden *et al.*, 2013). The fusiform geometry of MSO cells is thought to enhance the binaural coincidence detection effect by each side acting as a current sink when unstimulated by the other, as would occur in monaural stimulation (Agmon-Snir *et al.*, 1998).

MSO cells were initially thought to have constant bITDs across frequency, *i.e.* frequency-independent bITDs to tones. This response was termed a 'peak-type' response (cat IC – Yin & Kuwada, 1983; cat – Yin & Chan, 1990), meaning that the maxima of ITD curves measured to different pure tone frequencies all occur at the same ITD. A plot of BP or mean interaural phase of binaural beat response against carrier frequency, using binaural beats at different carrier frequencies, gives the 'characteristic delay' (CD) as the gradient and the 'characteristic phase' (CP) at which the coincidence occurs as its y-intercept. A peak-type cell gives rise to a linear distribution, with the slope being equal to the CD and the CP being near-zero. Contrastingly, cells of the LSO generally give rise to a CP of close to 0.5 cycles, termed a 'trough-type' response – coincidence occurs in the minima of the ITD curves (see **Section 3.1.4.2, page 62**). A schematic of an ideal peak-type and trough-type unit can be seen in **Figure 18 (page 58)**.

However, MSO cells with non-zero CP, termed 'non-peak-type' or 'intermediate' units, have been described (gerbil – Day & Semple, 2011). In these cells, the CD occurs at a specific frequency-invariant phase delay; inhibitory synaptic currents may act to shift the time-course of summation in these units (see **Section 1.2.5.1, page 9**).

The peak-type classification requires linearity of the phase-frequency relation; frequency-dependent bITDs are predicted by the cochlear delay hypothesis (see **Section 1.2.5.1, page 9**) – mismatches in the BF of inputs to each side of the MSO cells.

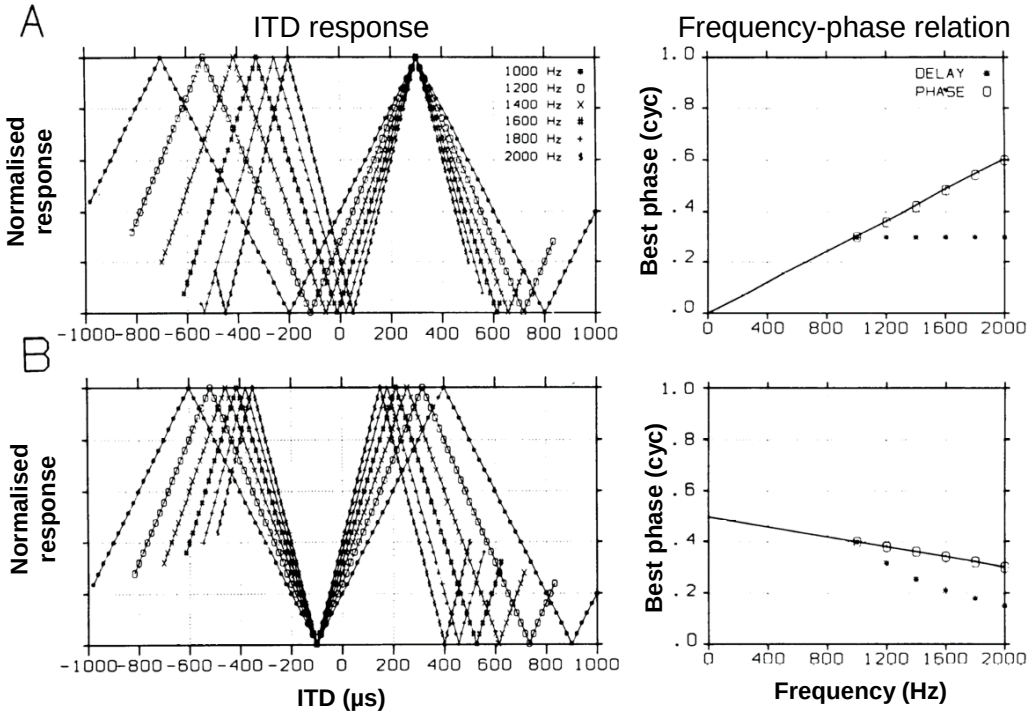

Figure 18 **Simplified peak- (**A**) and trough-type (**B**) units, showing how simulated ITD responses, shown left, give rise to frequency-phase relations, shown right**. Across all carrier frequencies, a peak-type unit maximally responds at the same ITD; this corresponds to a different interaural phase for a given frequency. All curves peak at the same ITD, and the slope of the frequency-phase relation gives the CD, with a y-intercept (CP) of 0 cycles. Conversely, a trough-type unit has coincidence in the minima across frequency; this gives rise to a CP of 0.5-cycles. Excerpted from Yin & Kuwada, 1983.

3.1.3.3 Experimental challenges

The MSO is considered a difficult area of the auditory system from which to make recordings. Previous studies of MSO have reported problems with cell isolation and low yields (dog – Goldberg & Brown, 1969; cat – Yin & Chan, 1990; gerbil – Spitzer & Semple, 1995). This Thesis reports the first *in vivo* responses recorded from the guinea pig MSO.

The close-packing of inputs to the MSO and its sheet-like geometry give rise to a large local field potential (cat – Lorente de Nó, 1947), termed the 'neurophonic' as it resembles the stimulus (cat – Mc Laughlin *et al.*, 2010). Its presence decreases the signal-to-noise ratio

(SNR) of unit recordings. Although neurophonics can be seen in other auditory nuclei, the neurophonic in the MSO is thought to be particularly strong due to MSO cells and their inputs strongly phase-locking to tones, especially at low frequencies (chinchilla – Langford, 1984; cat – Yin & Chan, 1990). Evidence of the neurophonic can be seen in the electrode recordings shown in **Figure 50 (page 117)**.

Action potential amplitude size and variability further complicate recordings. Average *in vitro* action potential amplitude in MSO cells is one seventh that of MNTB cells (gerbil – Scott *et al.*, 2005). Spike amplitude has been found to be level-dependent (gerbil – Day & Semple, 2011), possibly due to EPSPs with longer rise times giving rise to lower-amplitude spikes (gerbil – Scott *et al.*, 2007). Action potentials have been shown to have graded amplitudes at the soma (gerbil – Scott *et al.*, 2007), heavily attenuating as they backpropagate from the site of initiation, the first 25-µm of the output axon, to the soma. This is thought to be due to I_K mediated by Kv1-channels (gerbil – Scott *et al.*, 2007).

MSO units were only found in 26 of the 58 SOC experiments conducted in this Thesis; in those experiments, it took mean±standard deviation (sd): 6.7 ± 4.4 hours to find the first MSO unit. Holding times ranged from 9 minutes to 6 hours, mean±sd: 1.2 ± 1.0 hours.

3.1.3.4 Factors affecting interpretation of published results

The variation across studies in the methodology of finding MSO units lead to different biases that are not immediately apparent. Some studies did not anatomically-verify MSO unit recording location and so may inadvertently include non-MSO units or exclude atypical MSO units. Many studies search for units using only binaural beat stimuli, which leads to rejection of high-frequency, monaural or ITD-insensitive MSO units. Some studies have assumed units

demonstrating ‘trough-type’ properties to be of low-frequency LSO origin, leading to possible exaggeration of the proportion of reported ‘peak-type’ units in MSO. In contrast, in this study, the search stimulus consisted of monaural or diotic white noise and/or tones, and anatomical verification was undertaken for all low-BF ITD-sensitive units to place unit location in MSO or low-BF LSO.

Due to the difficulty of recording from the MSO, cells from other nuclei receiving MSO input, such as the dorsal nucleus of the lateral lemniscus (DNLL) and central nucleus of the IC (CNIC), have been treated as reliable surrogates of the MSO in some studies. This ignores any transformations or computations that may take place above the MSO.

One approach that has been used to circumvent the recording difficulties described in **Section 3.1.3.3 (page 58)** is to, rather than record from the cell bodies of MSO neurons, instead record from their output axons in the lateral lemniscus (LL) (chinchilla – Bremen & Joris, 2013). This approach increases the rate of unit sampling, however, a key weakness is that first spike latency (FSL) is used to determine unit origin: low-BF MSO and LSO unit populations have indistinguishable FSL distributions and can have similar properties of ITD-sensitivity. Additionally, axonal recordings are generally thought to be less stable over time than cellular recordings, meaning that less data can be gathered from a given MSO unit. Typical holding times using this technique were around 20 minutes (Sayles, personal communication), compared to the mean±sd: 1.2 ± 1.0 hours holding time in this Thesis.

3.1.3.5 Outputs

MSO cells send their main output to the ipsilateral IC, using glutamate as the neurotransmitter (cat – Oliver *et al.*, 1995). Low-, medium- and high-BF regions of MSO also send collaterals to

the lowest BF region of the MSO (guinea pig – Smith, 1995). SPN cells that project to the ipsilateral IC are also contacted by collaterals from MSO cells (gerbil – Kuwabara & Zook, 1999).

MSO-to-IC connectivity varies across species, leading to differences in the way ITD information is reprocessed, with either an increase in SNR by convergence of MSO units of similar BFs, or convergence across unit-BF of MSO units with similar bITDs (Vonderschen & Wagner, 2014).

Finally, I discuss the LSO, which is sensitive to level differences between the ears and is also thought to underpin sound source localisation.

3.1.4 Lateral superior olive

3.1.4.1 Inputs and properties

Units of the LSO are sensitive to ILDs (see **Section 3.1.4.2, page 62**). This arises as a result of ipsilateral glutamatergic input from SBCs (cat – Glendenning *et al.*, 1985; mouse – Wu & Kelly, 1992; cat – Smith *et al.*, 1993) and contralaterally-derived glycinergic input from the ipsilateral MNTB (cat – Glendenning *et al.*, 1985; cat – Bledsoe *et al.*, 1990; mouse – Wu & Kelly, 1992; cat – Tsuchitani, 1997; cat – Smith *et al.*, 1998) converging at LSO cells. Strychnine blockade of the glycinergic inhibition abolishes ILD-sensitivity in LSO (chinchilla – Moore & Caspary, 1983). The ILD-sensitivity of the LSO is thought to contribute to azimuthal sound source localisation (reviewed in Tollin, 2003).

The LSO is an S-shaped nucleus in cat, and has a tonotopic gradient, with low-BF cells in the lateral limb and increasing BFs towards the medial limb (cat – Guinan *et al.*, 1972a). In guinea

pig, the nucleus is U-shaped (Schofield & Cant, 1991). LSO cells are fusiform (cat – Cant, 1984; guinea pig – Schofield & Cant, 1991), oriented perpendicular to the axis of the nucleus. Multipolar LSO cells have also been reported (cat – Helfert & Schwartz, 1986, gerbil – 1987; guinea pig – Schofield & Cant, 1991).

The 'subtraction model' of LSO responses is complicated by evidence of both ipsilateral inhibitory projections, thought to originate from multipolar cells (rat – Doucet & Ryugo, 2003) or VCN/LNTB (mouse – Wu & Kelly, 1994), and contralateral excitatory projections, originating from SBCs (cat – Glendenning *et al.*, 1985; gerbil – Kil *et al.*, 1995).

LSO tuning curves to ipsilateral stimulation are V-shaped, resembling those of SBCs in the VCN; to contralateral stimulation, the inhibitory tuning curve arising from the MNTB inputs resembles the inverse of the ipsilateral, and the two are closely matched in shape (cat – Tsuchitani & Boudreau, 1969; cat – Tsuchitani, 1977, 1997; cat – Tollin & Yin, 2005).

3.1.4.2 Binaural properties

LSO cells show a sigmoidal response to BF tones presented at varying ILDs, with minimal response where the contralateral level is greater than the ipsilateral[9], maximal response where ipsilateral level is greater, and the slope having maximal gradient around zero-ILD (cat – Tsuchitani & Boudreau, 1969; cat – Joris & Yin, 1995; cat – Smith *et al.*, 1998; bat – Park *et al.*, 2004, gerbil – 2008; cat – Tollin & Yin, 2005; gerbil – Magnusson *et al.*, 2008; cat – Tollin *et al.*, 2008; cat – Tsai *et al.*, 2010). Pedestal level affects ILD response, *i.e.* a greater spike

[9] By convention, this is positive ILD.

rate results at the same ILD for a sound with greater overall stimulus level (cat – Tsai *et al.*, 2010).

In individual LSO cells, ILD increments of as low as 0.5-dB can be discriminated, with a median of 4.35-dB (cat – Tollin *et al.*, 2008). It has been suggested that these thresholds set the lower bound psychophysical ILD thresholds: in both cats and humans, ILD difference limens are 1-dB, with behavioural thresholds worsening as the pedestal level is increased or as the IPD is increased (cat – Wakeford & Robinson, 1974; human – Yost & Dye, 1988). Additionally, LSO cell responses to the same ILD can vary depending on the ILD of the previous stimulus (gerbil – Park *et al.*, 2008), thought to be mediated by retrograde GABAergic mechanisms (gerbil – Magnusson *et al.*, 2008).

High-BF LSO cells are sensitive to ITDs and IPDs in the temporal envelope (ENV) of high-frequency signals (cat – Joris, 1996). Low-BF LSO cells show IPD- and ITD-sensitivity to temporal fine structure (TFS), arising from phase-locked inputs (cat – Tollin & Yin, 2005). Maximal response occurs when the ipsilateral excitation is greatest and the contralaterally-derived inhibition is minimal, independent of frequency. In this way, the LSO mechanism is also mediated by coincidence detection of their inputs, responding maximally when there are no coincidences. Coincidence occurs in the minima of ITD curves, irrespective of ILD (cat – Joris & Yin, 1995), leading to a 'trough-type' (see **Figure 18, page 58**) response (cat, high-frequency LSO – Joris, 1996; cat, low-frequency LSO – Tollin & Yin, 2005).

LSO cells generally show binaural suppression when stimulated diotically, compared to ipsilateral-only. However, binaural facilitation has been shown in labelled LSO cells (gerbil – Kil *et al.*, 1995). It was suggested that this might arise from the non-uniform distribution of glycine receptors throughout the LSO, with a lesser density at the low-frequency limb of LSO

(gerbil – Sanes & Wooten, 1987). In response to diotic tones, LSO cells show either transient chopper or sustained responses (cat – Tsuchitani, 1988a, 1988b).

Phase-locking of LSO cells with BFs<1.2-kHz to ipsilateral pure tones, binaural pure tones at best ITD and to contralateral AM is enhanced compared to ANFs (cat – Joris & Yin, 1998; cat – Tollin & Yin, 2005).

3.1.4.3 Outputs

The principal output of the LSO is to the IC: around 40% of cat principal LSO cells are glycinergic and project to the ipsilateral IC; around 55% are glutamatergic and project mostly to contralateral IC and/or DNLL (Glendenning & Masterton, 1983; Glendenning *et al.*, 1992). Cells from the (low-frequency) lateral limb of the LSO project to ipsilateral IC; cells from the medial limb project to the contralateral IC (cat – Glendenning *et al.*, 1985).

3.2 METHODS

3.2.1 Online unit classification

The dorsoventral electrode approach used in this Thesis meant that the precise location of the electrode could not be known until histological processing of lesioned unit recording sites. This necessitated the use of an online classification scheme, which also gave an estimate of where the electrode was within the SOC. This information was used to decide which direction to move the electrode in subsequent tracks. The principal aim was to maximise the chance of isolating MSO units.

SOC units were defined as those found at depths >3.300-mm from surface of DCN using the electrode approach described in **Section 2.1 (page 32)**. The location of a lesion created at the final electrode position, usually at a putative MSO or LSO unit recording site, was used for confirmation or refutation of this hypothesis. A second lesion was created with the electrode retracted 1-mm to enable reconstruction of the track trajectory.

The presence of a neurophonic in response to monaural and/or binaural stimulation with low-frequency pure tones gave an initial indication that the electrode was close to MSO or a low-frequency area of LSO. The neurophonic was generally of equal amplitude for monaural stimulation when the electrode was near MSO.

The online criteria used to identify presumed SOC unit types were as follows:

a) MSO units

- low-BF (<2-kHz);
- ITD-sensitive, assessed with binaural beats;

- [E·E][10] and showing stronger responses to diotic[11] tonal stimulation than to monaural stimulation;
- at depths ~4.500- to 6.000-mm from the surface of the DCN.

In no cases did the lesion evidence contradict the online classification. Note that this approach rejects possible monaural and/or high-frequency MSO units: consequently, the highest BF of the MSO unit population was ~2-kHz, and no monaural units were found.

b) LSO units

- showed contralateral inhibition and ipsilateral excitation, [E·I];
- showed clear binaural suppression in response to diotic tonal stimulation;
- at depths ~3.300- to 6.000-mm from the surface of the DCN.

Disambiguating low-BF LSO cells, which are ITD-sensitive, from MSO cells was of particular concern. Typical responses to binaural beats and unit receptive fields allowed the formation of an online hypothesis (see **Figure 39, page 90**). In no case was the online hypothesis refuted by the lesion evidence, however lesions were not generally created at LSO recording sites as the experimental priority was to accrue data from MSO units.

c) MNTB units

- showed clear evidence of a three-component spike waveform shape indicating presynaptic activity (‘PP’ component), the resulting EPSP (‘A’ component) and the evoked spike (‘B’ component);
- showed a PL or PN response, as would be expected from their (likely) single input from GBCs;

[10] In this Thesis, binaural responses are abbreviated as ‘[ipsilateral response]·[contralateral response]’. ‘E’ is excitatory, ‘I’ is inhibitory, and ‘0’ is no effect, based on peri-stimulus responses to pure tones.

[11] Diotic throughout this Thesis means that tones were presented at zero system ITD.

- showed contralateral-only excitatory responses, [0·E].

In practice, the second and third criteria did not rule out any units.

d) SPN units

- showed offset-only responses;
- responded only to contralateral pure tones [0·I][12].

Note that these criteria bias rejection of possible binaural units or non-offset units of the SPN, however there were no cases where the second criterion ruled out units.

3.2.2 Confirmation of recording sites

The methodologies for creation of lesions at unit recording sites, histological processing and scanning of brainstem sections are outlined in **Section 2.2.2 (page 35)**.

A parasagittal section from Schofield & Cant, 1991 was used as a model section of the centre of MSO. Scanned sections showing evidence of lesions were imported into *Adobe Illustrator* and scaled or rotated as appropriate to fit the model section. Camera lucida drawings of lesions were made using a *Microsoft Surface Pro 2* tablet, then lesion centroids were fitted using *Illustrator's* 'Ellipse' transform (see **Figure 19, page 68**). Symbols corresponding to units were then overlaid, centred on the centroids (see **Figure 20, page 69**).

[12] This means contralateral stimulation causes peristimulus inhibition. However, there was always a rebound excitation following this.

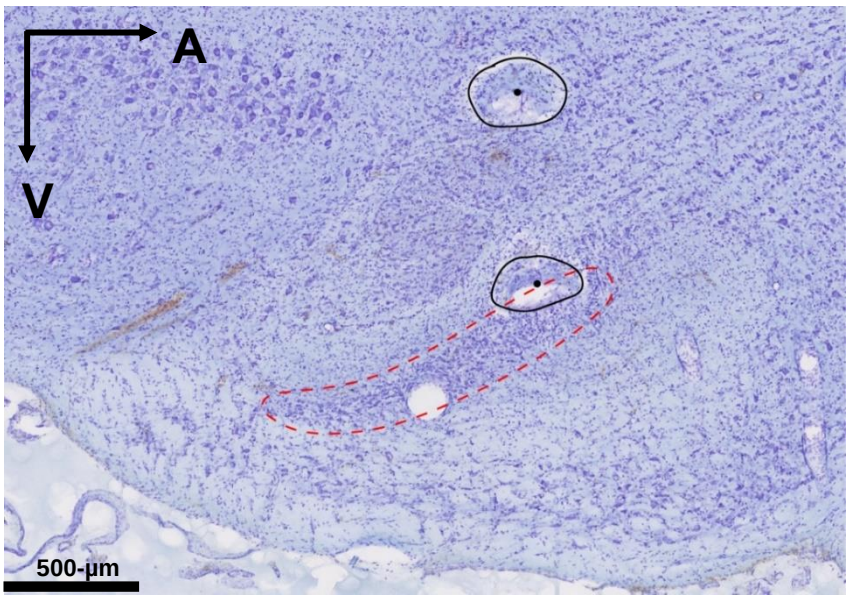

Figure 19 **Stained parasagittal section of guinea pig brainstem confirming unit location in MSO (outlined in red)**. Lesions (outlined in black) were made at the recording site and 1-mm retracted from **MSO 1633 001** (see **Figure 14, page 43** for spike waveform analysis and **Figure 37, page 88** for the unit binaural beat response and receptive field). Lesion centroids are shown as dots. The distance between the lesions in the section is shorter than 1-mm due to tissue shrinkage from histological processing. A = anterior; V = ventral.

Lesions were created at MSO and LSO unit recording sites in order to confirm or reject the online experimental hypothesis of unit type and electrode location. In earlier experiments, recordings were made from a single MSO or LSO unit and then the animal was perfuse-fixed; lesion administration not only damaged a sizeable portion of the nucleus in question (see **Figure 19** above) but may have also interfered with the responses of neurons in areas that appear unharmed. In later experiments, recordings were made from multiple MSO or LSO units and a lesion was made at the last unit recording site.

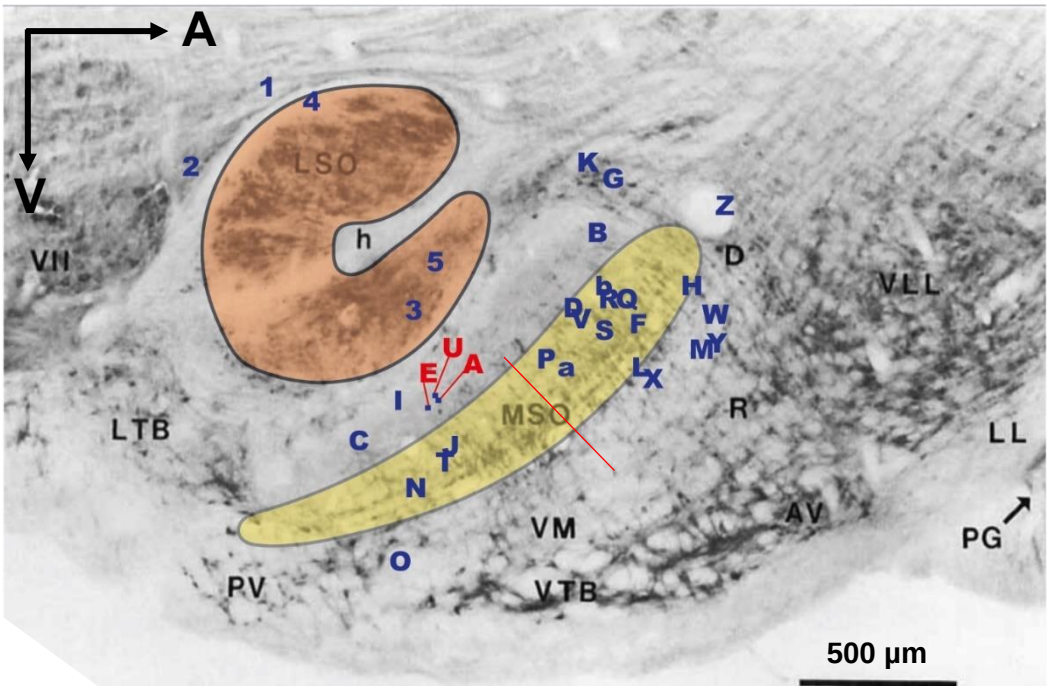

***Figure 20* Composite figure of lesioned recording sites across all SOC experiments**.

Lesions are demarcated by letters for MSO recording sites or numbers for LSO recording sites. Labels are positioned at the centroids of the lesions in all but three cases; for clarity, a dot is placed at these sites and the corresponding labels are in red. The red line corresponds to the location of a blood track through MSO seen in the sections from one experiment. A full list of the corresponding unit/experiment numbers can be found in **Appendix 7.1 (page 180)**. The parasagittal section of the guinea pig SOC is from Schofield & Cant, 1991.

Figure labels: VII = facial nucleus; LTB = lateral nucleus of the trapezoid body; PV = posteroventral periolivary nucleus; LSO = lateral superior olive; h = hilus of the LSO; MSO = medial superior olive; VM = ventromedial periolivary nucleus; VTB = ventral nucleus of the trapezoid body; D = dorsal periolivary nucleus; R = rostral periolivary nucleus; AV = anteroventral periolivary nucleus; VLL = ventral nucleus of the lateral lemniscus; LL = lateral lemniscus; PG = pontine grey. A = anterior, V = ventral.

3.2.3 Stimuli

See **Section 2.3 (page 36)** for a description of search and unit characterisation stimuli.

Binaural beats were generated using two pure tones of the same sound level, one presented to each ear. The frequency presented to the contralateral ear was always [ipsilateral tone frequency+1]-Hz. This generates a continuously-varying IPD over the 1-second period of the binaural beat. Initially, 10 repetitions of a 3-second binaural beat were presented at BF; in

some units, this was followed by 5 repetitions at a range of carrier frequencies, 100- to 1000-Hz in 50-Hz steps. All stimuli were presented at ~20-dB sensation level (SL, level above threshold).

3.3 RESULTS

3.3.1 Population data

A total of 726 units were isolated in the left SOC of 58 guinea pigs. Of these, 325 units were discarded after determining unit BF and threshold. The remaining 401 units entered the classification scheme delineated in **Section 3.2.1 (page 65)**. The classified SOC population totalled 144 units: MSO: *n* = 60, LSO: *n* = 28; MNTB: *n* = 24; SPN: *n* = 32. To maximise search time for MSO units, anatomical verification was reserved to the MSO and, in a few cases, the low-frequency LSO. This led to the use of highly stringent classification requirements (see **Section 3.2.1, page 65**), which biased rejection of 'atypical' SPN, MNTB, LSO or MSO units and those of other SOC nuclei.

First, I consider the distributions of MNTB, SPN, LSO and MSO cell populations of unit BFs, thresholds, spontaneous rates (SRs), FSLs, and receptive field widths assessed using filter 10-dB quality factor ($Q_{10\text{-}dB}$).

BFs in the classified SOC population varied from 0.13- to 23.93-kHz, mean±sd: 3.11 ± 4.69-kHz; the ranges can be seen for each nucleus in **Figure 22A (page 73)**. Unit BF was determined in monaural presentation for the ear evoking greater responses except for MSO units, where diotic BFs were determined.

Unit thresholds varied from 12- to 71-dB SPL, mean±sd 36 ± 13 dB-SPL. Unit SR varied from 0.0 to 92.9 spikes/s, mean±sd: 13.3 ± 20.8 spikes/s. **Figure 21A** below shows unit threshold plotted against unit BF.

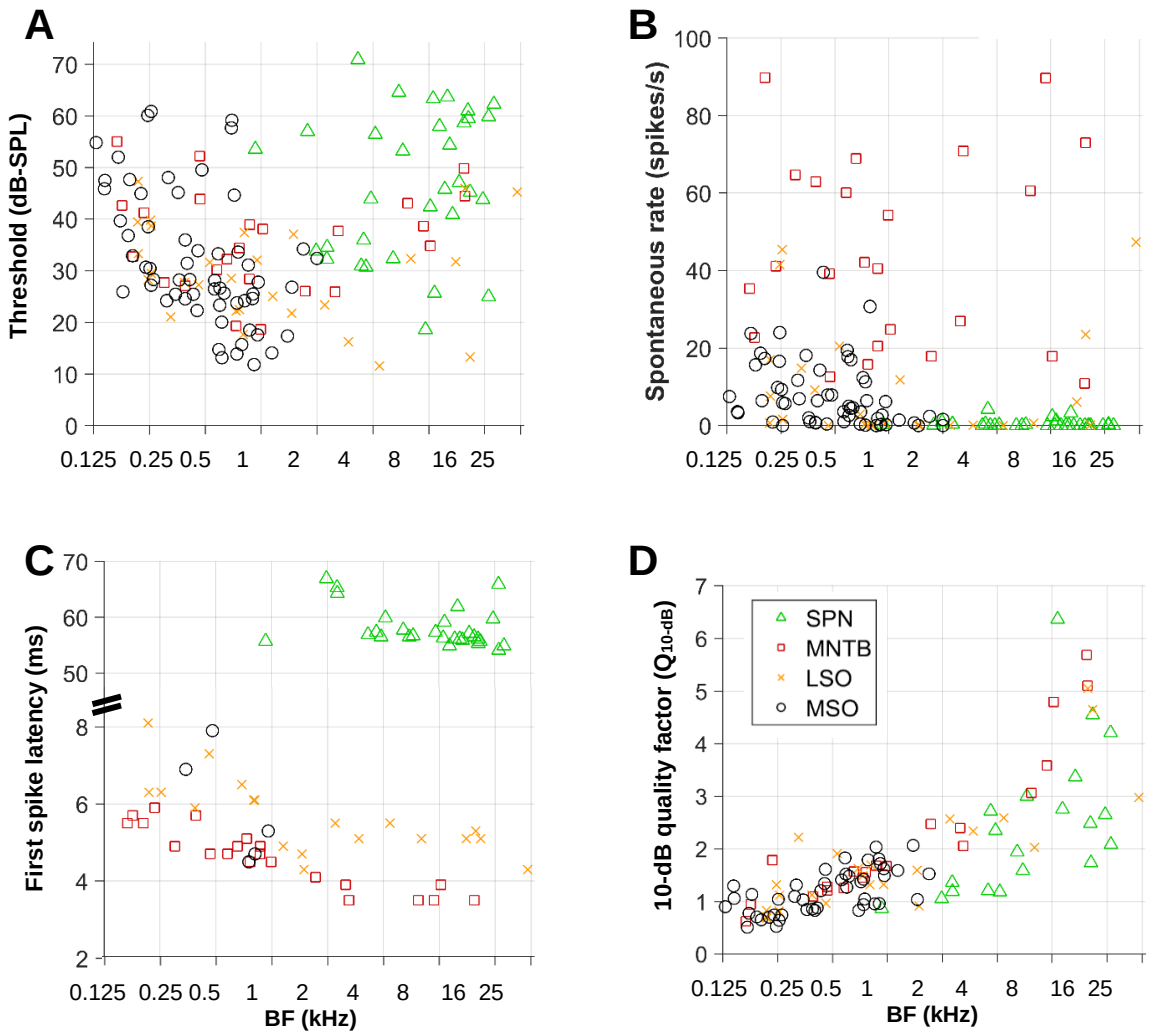

Figure 21 **Distributions of basic response properties against BF for the SOC population**. *n* = 144 (by nucleus: MSO: *n* = 60, LSO: *n* = 28; MNTB: *n* = 24; SPN: *n* = 32).

A: Unit threshold against unit BF. Note the low-BF bias in all nuclei except SPN.

B: Unit SR against unit BF. MNTB units showed the highest SRs; SPN the lowest.

C: Unit FSL decreases with BF, however SPN FSLs are greater than those in the other SOC nuclei as they do not respond during stimulus presentation (the stimulus used was a 50-ms pure tone at BF). Note the broken y-axis.

D: Filter sharpness, assessed via $Q_{10\text{-}dB}$, increases with increasing BF.

MNTB units were characterised with high SRs (mean±sd: 44.3 ± 24.5 spikes/s) and low thresholds (mean±sd: 36 ± 10 dB-SPL). SPN units were generally characterised as low-SR (mean±sd: 0.4 ± 1.0 spikes/s) and had higher thresholds (mean±sd: 47 ± 14 dB-SPL). LSO units had mostly low SRs, mean±sd: 9.0 ± 14.3 spikes/s. MSO units were low-BF (mean±sd: 0.60 ± 0.43-kHz) with mostly low SRs (mean±sd: 9.8 ± 14.8 spikes/s). **Figure 21B (page 71)** shows unit SR plotted against unit BF.

FSLs of SPN units were large, reflective of the absence of response during stimulus presentation (mean±sd: 58.1 ± 3.5-ms); the stimulus was a 50-ms pure tone at unit BF. By comparison, the FSLs of MNTB units, which supply the inhibitory input to SPN, were mean±sd: 4.6 ± 0.8-ms across the population. **Figure 21C (page 71)** shows unit FSL plotted as a function of unit BF.

FSL, $Q_{10\text{-}dB}$ and unit threshold are strongly BF-dependent. FSL decreases with increasing BF (see **Figure 21C, page 71**), and $Q_{10\text{-}dB}$ increases with increasing BF (see **Figure 21D, page 71**). Minimum unit thresholds are ultimately determined by the audiogram *i.e.* the lowest threshold ANFs at each frequency. Due to the differences in BF distributions between the unit populations of each nucleus, only low-BF (≤2-kHz) subpopulations were compared (SPN: n = 3; MNTB: n = 17; LSO: n = 20; MSO: n = 60). **Figure 22 (page 73)** compares the distributions of these properties between nuclei subpopulations and, using ANOVA analysis, assess the differences of these distributions for significance. Neither the threshold distributions nor the $Q_{10\text{-}dB}$ distributions of these subpopulations differed significantly at the $p<0.05$ level. FSLs differed significantly ($p<0.001$) between LSO and SPN, MSO and SPN and SPN and MNTB; all other comparisons were nonsignificant at the $p<0.05$ level.

The responses to diotic and monaural-only stimulation were compared in the binaural MSO and LSO nuclei. Facilitation or suppression was assessed using a facilitation index (FI) (see

Section 2.3.3, page 37). An FI of greater than one indicates that the diotic response exceeds the maximum of the monaural responses.

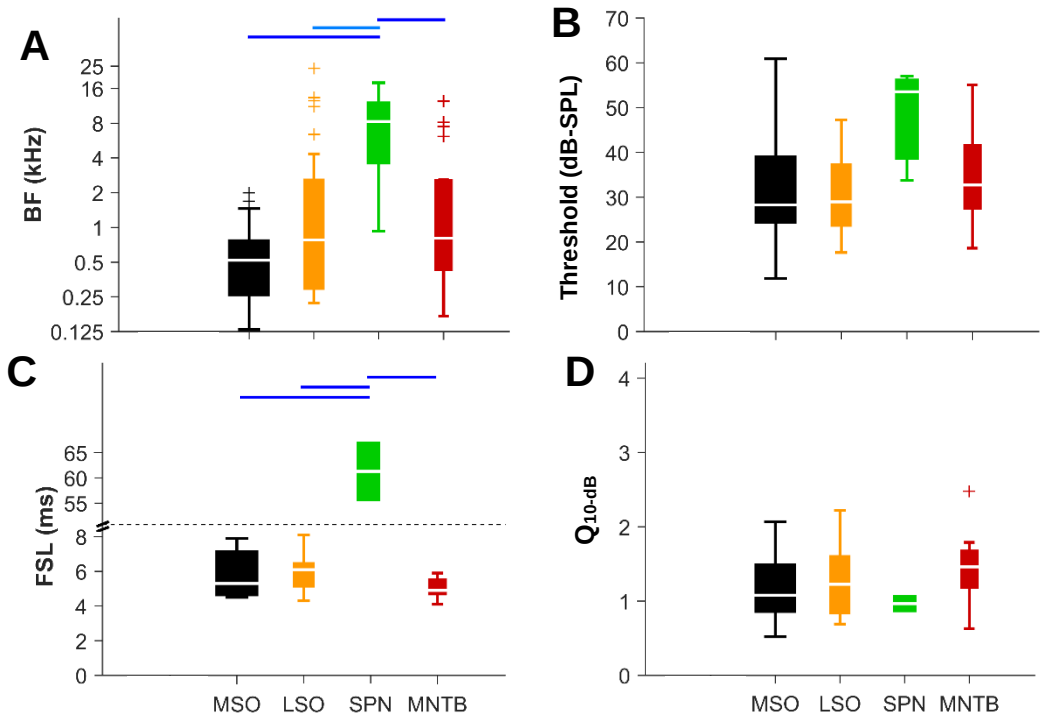

Figure 22 **SOC nuclei BF, threshold, FSL and $Q_{10\text{-}dB}$ distributions**.

A: BF distributions for the nuclei highly differed; **B**, **C** and **D** show data from units with BF≤2-kHz.

B: Threshold distributions for units with BF≤2-kHz are not significantly different at the p<0.05 criterion.

C: The SPN FSL distribution is significantly different than the FSL distribution of other SOC nuclei (p<0.001).

D: $Q_{10\text{-}dB}$ distributions do not differ significantly across the low-BF subpopulations of nuclei of the SOC.

Significance legend: **p<0.01**; **p<0.001**.

MSO units showed [E·E] responses, with stronger responses elicited in diotic stimulation, and weak responses to monaural-only tone stimulation. In all units, FI significantly exceeded 1 (p<0.001). On average, diotic stimulation of MSO units produces spike rates a factor of mean±sd: 2.6 ± 1.6 times the spike rate of the monaural-only stimulation. In contrast, LSO units showed strong ipsilateral receptive fields and weak responses to diotic tones, due to the

contralaterally-derived inhibition from the MNTB; diotic stimulation resulted in significantly lower spike rates than ipsilateral-only ($p<0.01$). Statistically, the FI distributions of MSO and LSO cells were significantly different ($p<0.001$) (see **Figure 23B** below).

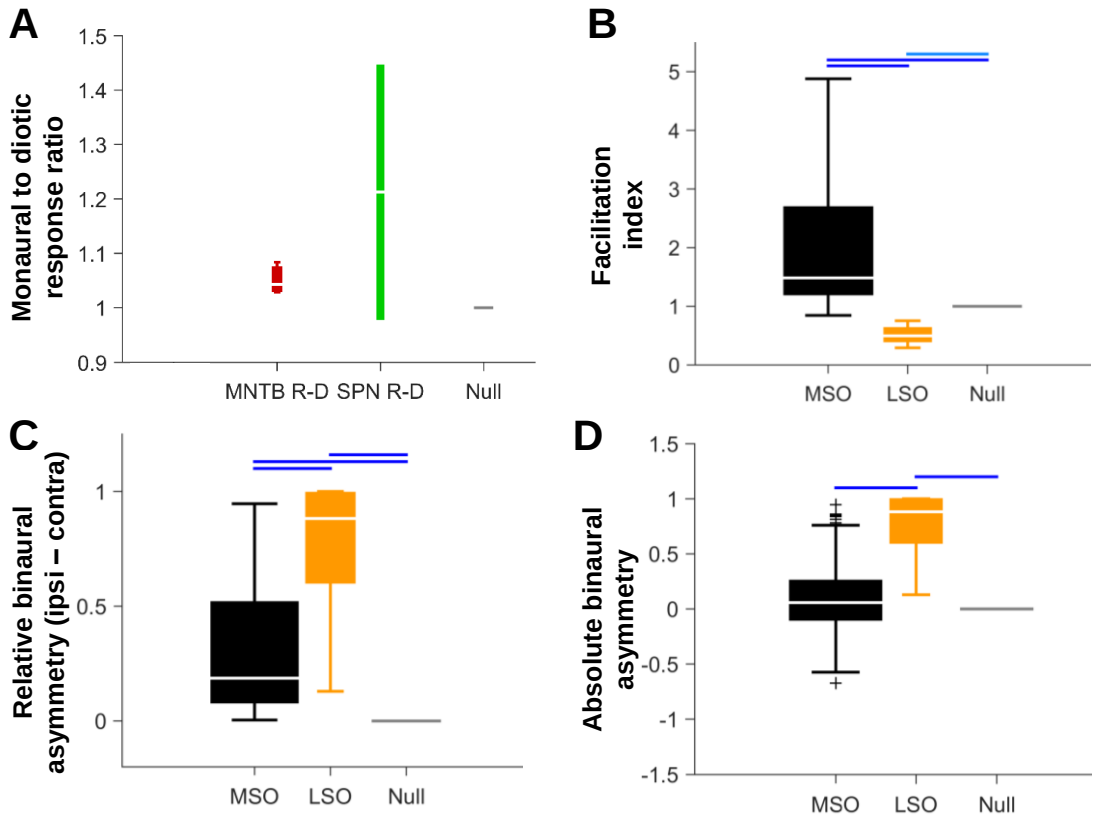

Figure 23 **MSO and LSO units show differing monaural and binaural responses**.

A: Monaural SOC units do not show significant contralateral inhibition at the $p<0.05$ level (y-axis values of >1) (SPN: $n = 3$; MNTB: $n = 3$).

B: MSO cells ($n = 44$) show significantly different ($p<0.001$) FIs compared to the LSO and null hypothesis of no facilitation.

C: MSO units exhibit significant differences between the strengths of the monaural receptive fields ($p<0.001$).

D: The MSO population as a whole does not show a significantly greater response to a particular ear.

Significance legend: **$p<0.01$**; **$p<0.001$**.

To assess ear dominance in MSO units, the response magnitudes of ipsilateral-only and contralateral-only receptive fields were compared (see **Section 2.3.3, page 37**). In MSO, there was a significant asymmetry between the strengths of responses to monaural stimulation of either ear ($p<0.001$) (see **Figure 23C, page 74**). However, there was no dominance of a particular ear across the entire population ($p\geq0.05$) (see **Figure 23D, page 74**). In contrast, the LSO showed significant dominance ($p<0.001$) of the ipsilateral ear over the contralateral, expected due to its inhibitory input from MNTB.

Next, I discuss the basic response properties of the individual nuclei.

3.3.2 Medial nucleus of the trapezoid body

MNTB units show the same three-component spike waveform shape seen in other species and in SBCs of the VCN (see **Sections 3.1.1, page 48** and **5.1.1.1, page 140**). They show PL or PN responses to tones presented to the contralateral ear ([0·E]). **Figure 24 (page 76)** shows example PSTHs from MNTB units.

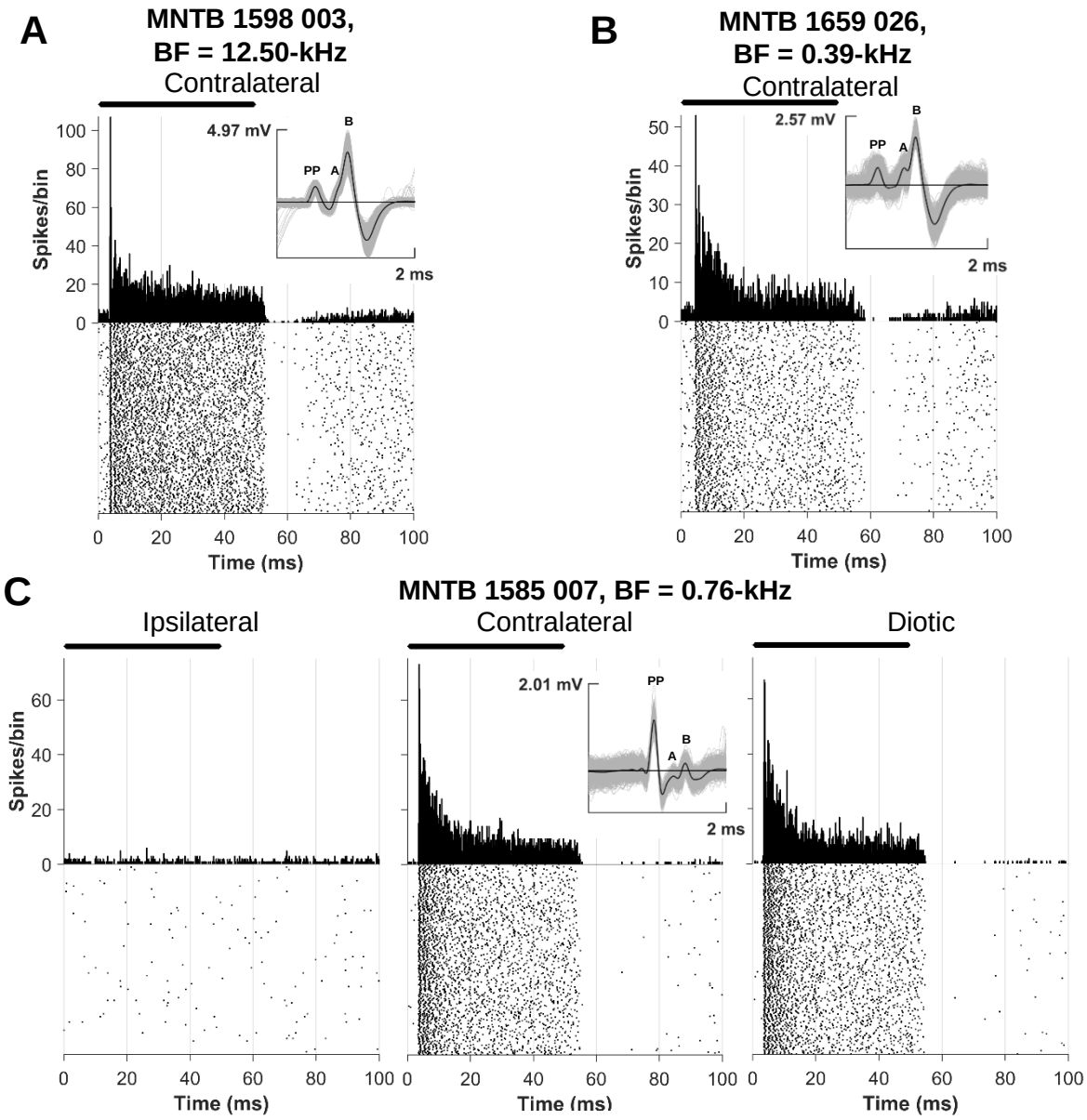

Figure 24 **MNTB units show a PL or PN response to short BF tone bursts**. Average spike waveforms are inset. MNTB units are characterised by prepotentials ('PP') in their spike waveform shapes. Stimulus is represented by a black bar.

A: PSTH of a high-BF MNTB unit. There is a precisely-timed first spike shortly after stimulus onset, followed by a depression in firing rate, followed by the resumption of a PL response: a clear PN response.

B: PSTH of a low-BF MNTB unit. There is a post-stimulus depression in firing rate below SR, reflective of stimulus adaptation.

C: Diotic stimulation of MNTB units does not result in a greater response magnitude than contralateral stimulation; ipsilateral-only stimulation does not cause unit activity above SR. This particular unit shows a larger PP component than the evoked spike ('B'); 'A' is the EPSP evoked by the PP.

MNTB unit receptive fields show only responses to the contralateral ear. Diotic receptive fields resemble those of contralateral-only; ipsilateral stimulation does not elicit a response above spontaneous rate. See **Figure 25** below and **Figure 40 (page 92)** for some example MNTB unit receptive fields.

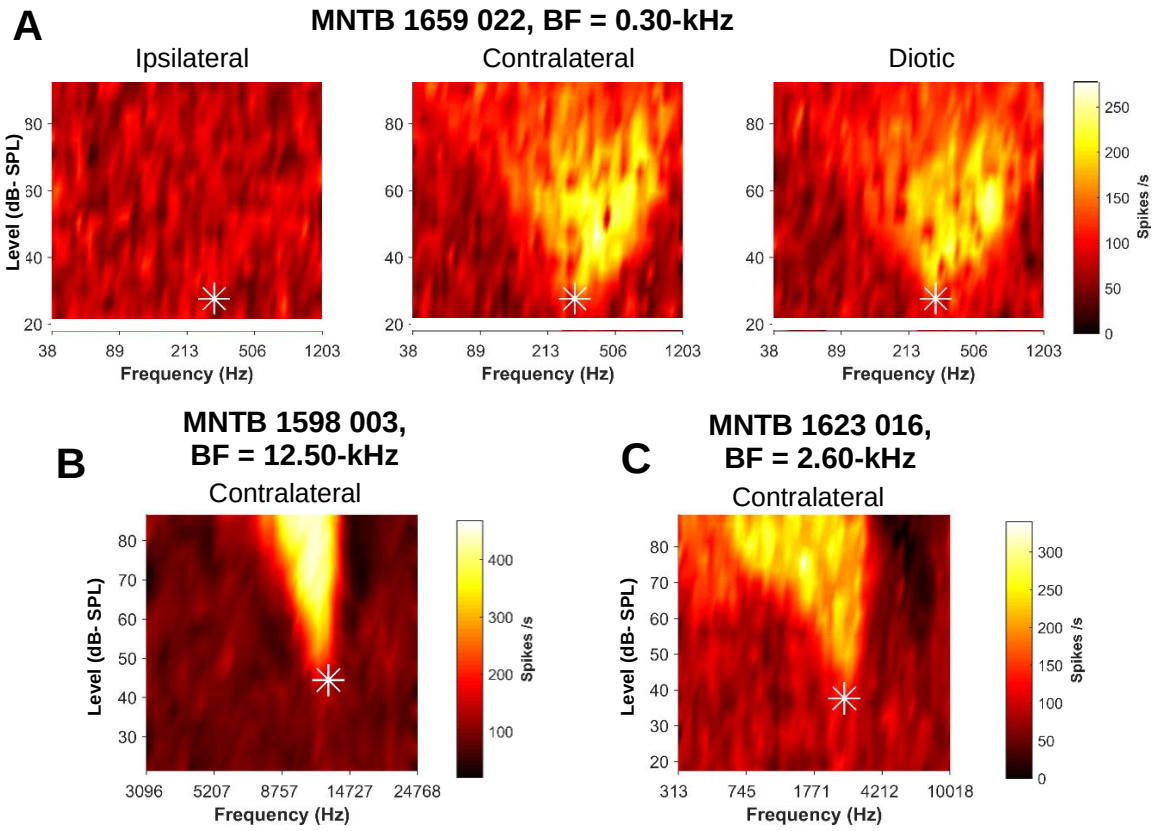

Figure 25 **MNTB unit receptive fields show [0·E] responses**. Fitted BF and threshold are marked by the white asterisks.

A: Diotic receptive fields do not show differences from contralateral-only receptive fields; additionally, ipsilateral stimulation does not evoke a response above spontaneous rate.

B and **C**: Receptive fields show greater relative filter bandwidths at lower BFs.

3.3.3 Superior paraolivary nucleus

SPN units show [0·I] offset-only responses to pure tones, as described in other species. During stimulus presentation, cells do not show responses (see **Figure 26** below). Following a short period of time after stimulus presentation ('continued inhibition'), a rebound excitation occurs as a burst of spikes.

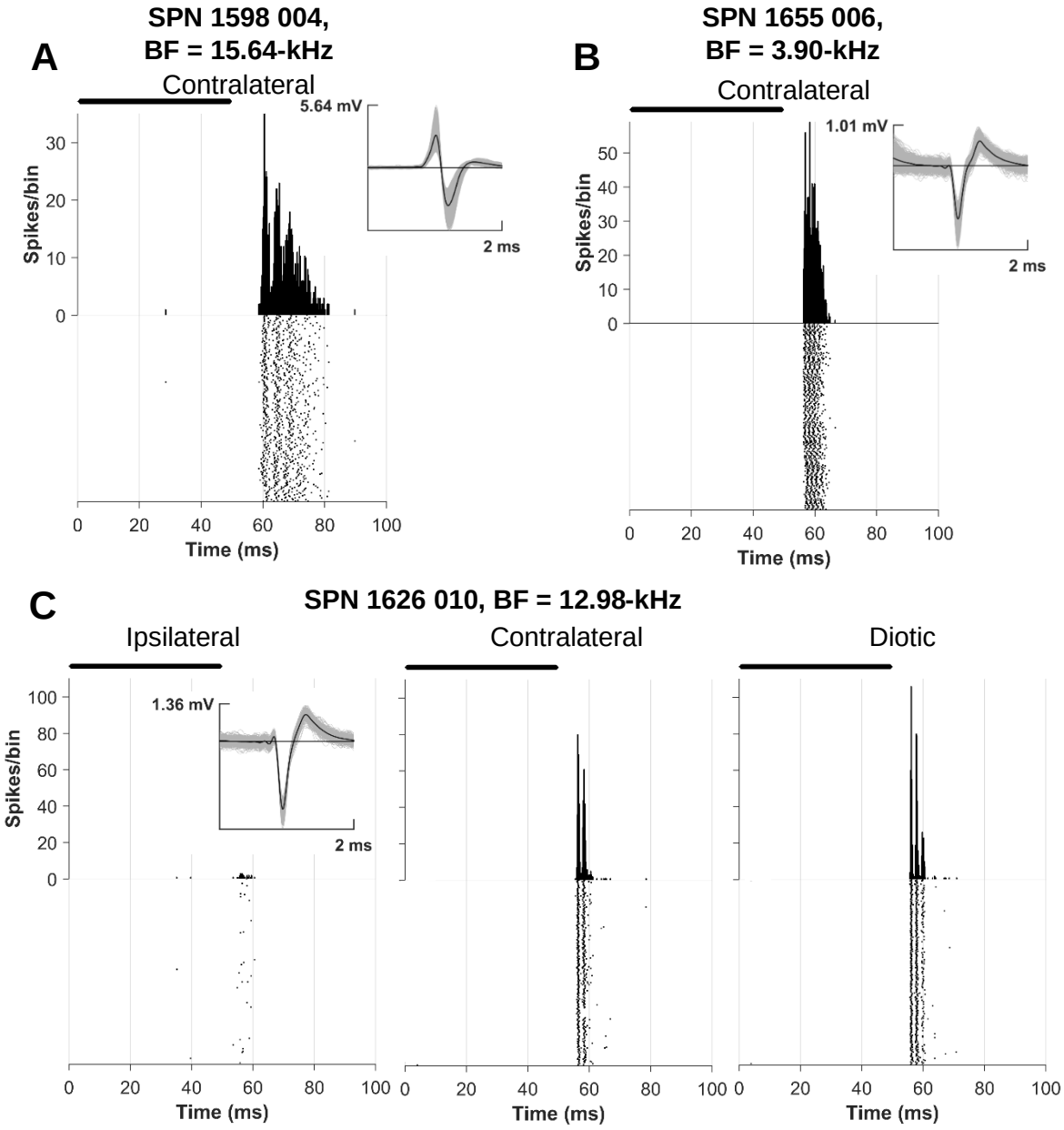

Figure 26 **SPN units respond at the offset of stimulation of the contralateral ear with a multipeaked 'chopping' response**. Stimulus time course is represented with a black bar, spike waveforms are inset.

A and **B**: SPN units show little-to-no SR, but a clear offset response to contralateral BF tones.

C: Unit shows weak binaural facilitation, with a greater response to diotic stimulation than contralateral-only.

SPN units show an inhibitory receptive field when analysed during stimulus presentation. Following stimulus presentation, the inhibited region shows an increased response over spontaneous, reflecting the rebound excitation. See **Figure 27** below for some example SPN unit receptive fields.

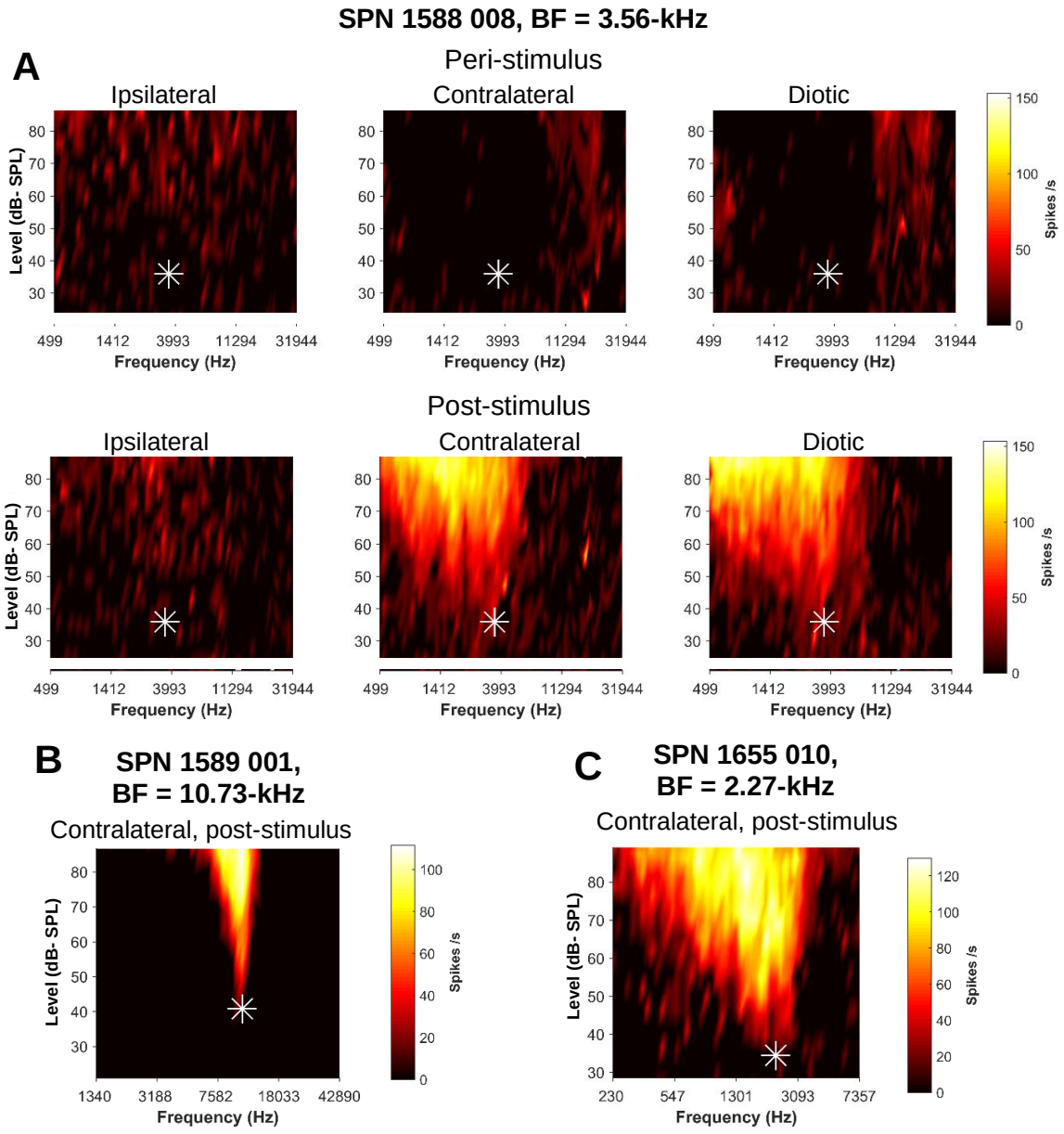

Figure 27 **Receptive fields measured from SPN**. Fitted BF and threshold are marked by the white asterisks.

A: Inhibition occurs during contralateral stimulus presentation; following stimulus offset, the inhibited region shows a rebound excitation.

B and **C**: SPN receptive fields are monotonic.

Units of the guinea pig SPN show sensitivity to gaps in tones, as described in other species (see **Section 3.1.2, page 51**). **Figure 28** below shows responses of two SPN units to varying gap durations in BF tones. For small gaps, the offset response that would ordinarily be evoked by the first tone is prevented by inhibition arising from the second tone. For larger gaps, however, SPN units fire at the offset of both stimuli. GDTs, assessed as the duration evoking at least one spike in at least 76% of presentations, varied from 1- to 15-ms across the SPN population, *n* = 7 (see **Figure 29, page 81**).

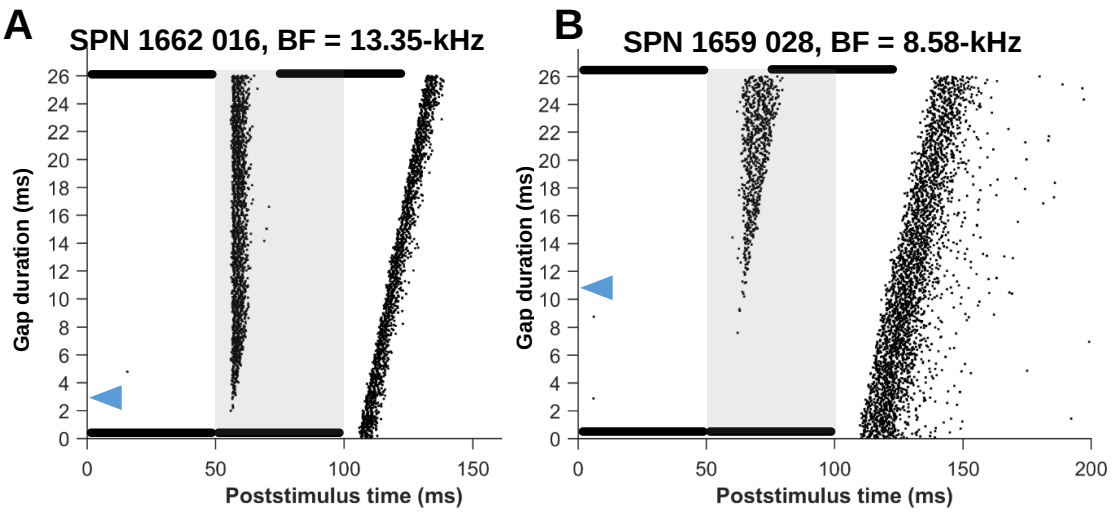

Figure 28 **SPN units do not respond to very short gaps; instead, they show varying gap detection thresholds**. Rasters are shown for two SPN units. GDTs are shown as blue triangles, stimuli are depicted as black bars. 50-ms pure tones at unit BF were presented with a varying gap, from 0-ms (the stimuli were still gated on and off) to 25-ms. Spikes were analysed in the grey-shaded area. For short gaps, neither unit responds at the offset of the first tone, essentially treating the two stimuli as continuous.

A: This unit shows weak firing starting at around gaps of 2-ms, with the offset response increasing up to a saturation for gaps 8-ms and higher.

B: In contrast, this unit does not start firing at the offset of the first tone, and thus detect the gap, until around 11-ms.

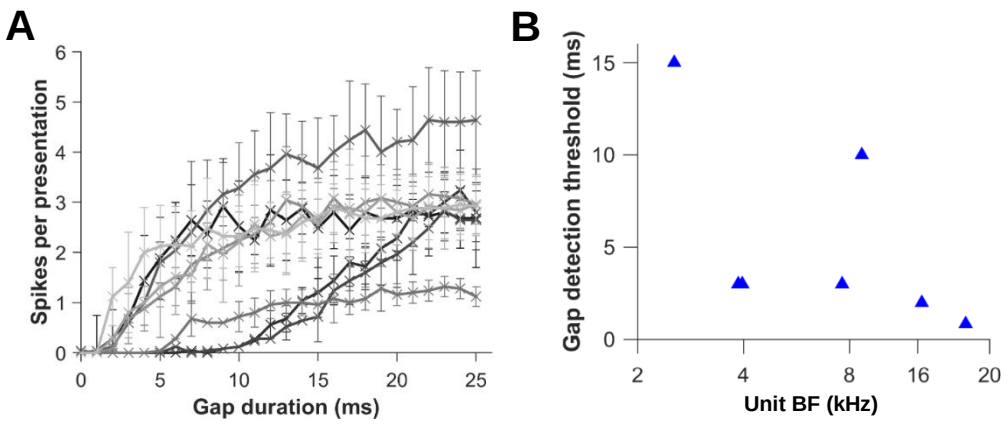

Figure 29 **SPN gap detection threshold varies across the unit population**.

A: SPN units show a sigmoidal response to increasing gap durations, with the threshold varying between units.

B: Across the SPN population (*n* = 7), GDT generally decrease with increasing BF.

SPN units (*n* = 6) respond with a lower FSL and a greater number of spikes per presentation as the duration of a BF pure tone increases, up to a saturation (see **Figure 30** below). Similar properties have been reported in other species.

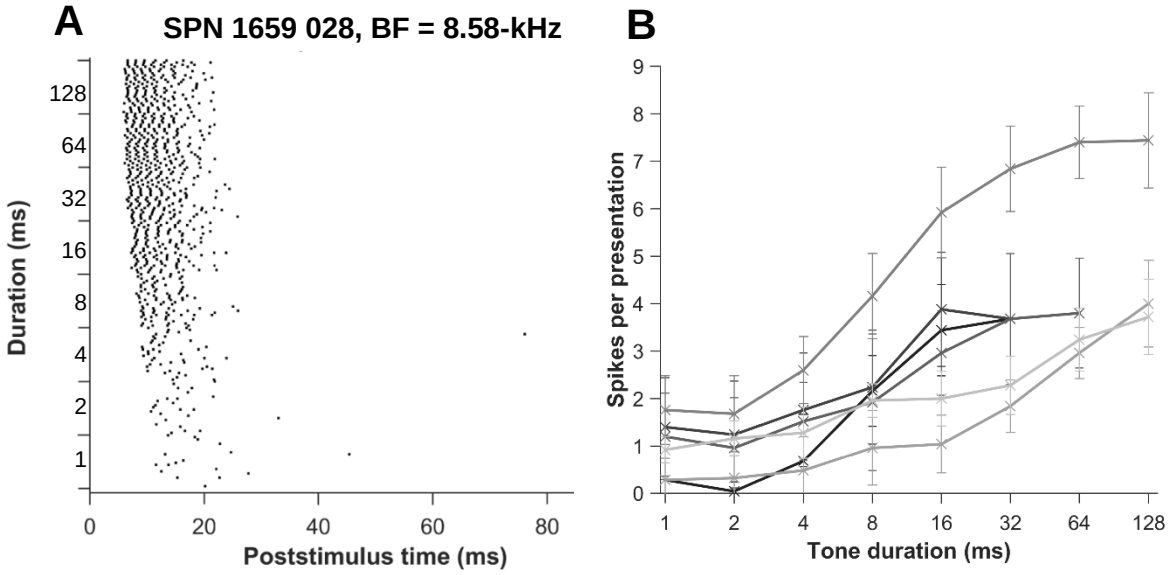

Figure 30 **SPN unit first spike latency decreases and spikes per presentation increase with increasing tone duration**.

A: Raster plot for an SPN unit. X-axis starts from stimulus offset.

B: Average spikes per presentation follow a sigmoidal function when tone duration is plotted on a logarithmic axis. *n* = 6.

SPN units were assessed for entrainment, the proportion of periods of modulation that evoked at least one spike, to sinusoidally amplitude-modulated (SAM) tones at unit BF. There was variability in the upper frequency limit of modulation across units: this is likely due to the variability of the strength and number of GABAergic inputs from SPN collaterals and autapses (see **Section 3.1.2, page 51**), which have a slower time course than glycinergic inhibition. Units showed a bandpass tuning: at low modulation frequencies, the inhibition builds up to a great-enough extent in periods of high stimulus energy that it is not fully cleared to allow a rebound excitation in periods of low stimulus energy. At high modulation frequencies, the level of inhibition does not subside enough to allow a rebound excitation. For middle-frequencies, SPN units show high entrainment, responding after each ENV period of the SAM. **Figure 31** below shows raw responses for SAM-BF tones for one SPN unit and the entrainment indices across the population, *n* = 9.

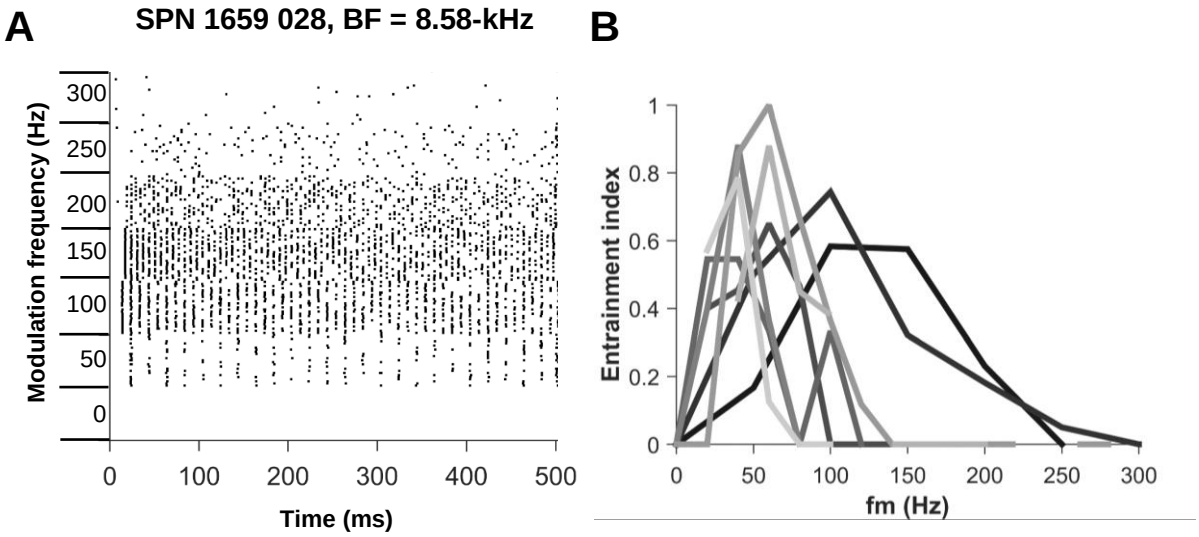

Figure 31 **SPN units are bandpass tuned to the modulation of SAM-BF tones**.

A: Spike rasters from a single SPN unit. Entrainment – responding with at least one spike to each period of the SAM tone – increases as modulation frequency increases, until it reaches a maximum. At high modulation frequencies, entrainment decreases: units respond only to the offset of the whole stimulus.

B: SPN entrainment to SAM-BF tone modulation rate is bandpass (*n* = 9). However, the upper frequency limit of entrainment is highly variable due to varying levels of GABAergic input to each cell.

3.3.4 Lateral superior olive

Units of the guinea pig LSO show sensitivity to ILDs for tones at BF (see **Figure 32** below). This originates from the inhibition resulting from contralateral stimulation. Spike rate decreases as the contralateral level increases.

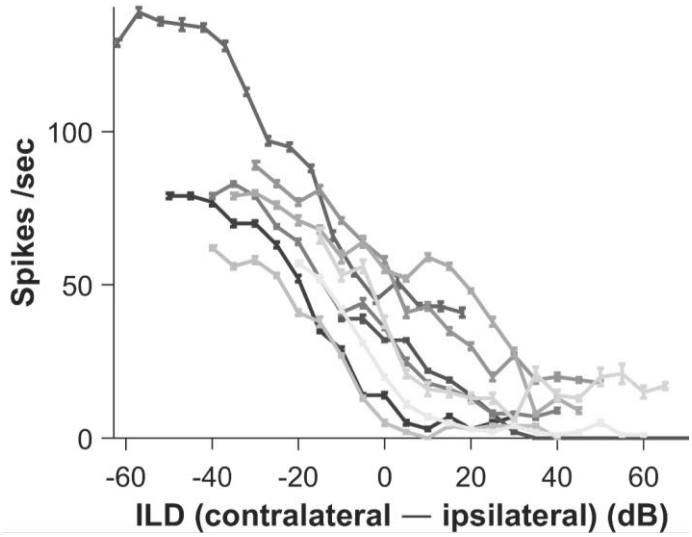

Figure 32 **Guinea pig LSO units show ILD-sensitivity similar to that described in other species**. *n* = 9. Stimuli were BF pure tones with varying ILD; the level of the ipsilateral tone was held constant and the contralateral level roved. As the contralateral level increases, the spike rate decreases sigmoidally.

LSO units showed chopper, PL or PN responses to ipsilateral-only tones, and an almost entirely cancelled out response to diotic stimulation. This is reflective of the contralaterally-derived MNTB inhibition. See **Figure 33 (page 84)** for example PSTHs.

Receptive fields to ipsilateral-only stimulation show monotonically-increasing spike rates with increasing sound level. In contralateral-only stimulation, either zero response was elicited, or if the unit had non-zero SR, a suppressive field is seen, resulting from the inhibition from the MNTB. In cases where units had zero or low SRs, an 'artificial' SR could be created by

stimulating the unit ipsilaterally with a BF-tone at low sound level, usually 5- or 10-dB SL, a 'tickle-tone' receptive field. See **Figure 34 (page 85)**, **Figure 35 (page 86)** and **Figure 39 (page 90)** for receptive fields of some example LSO units.

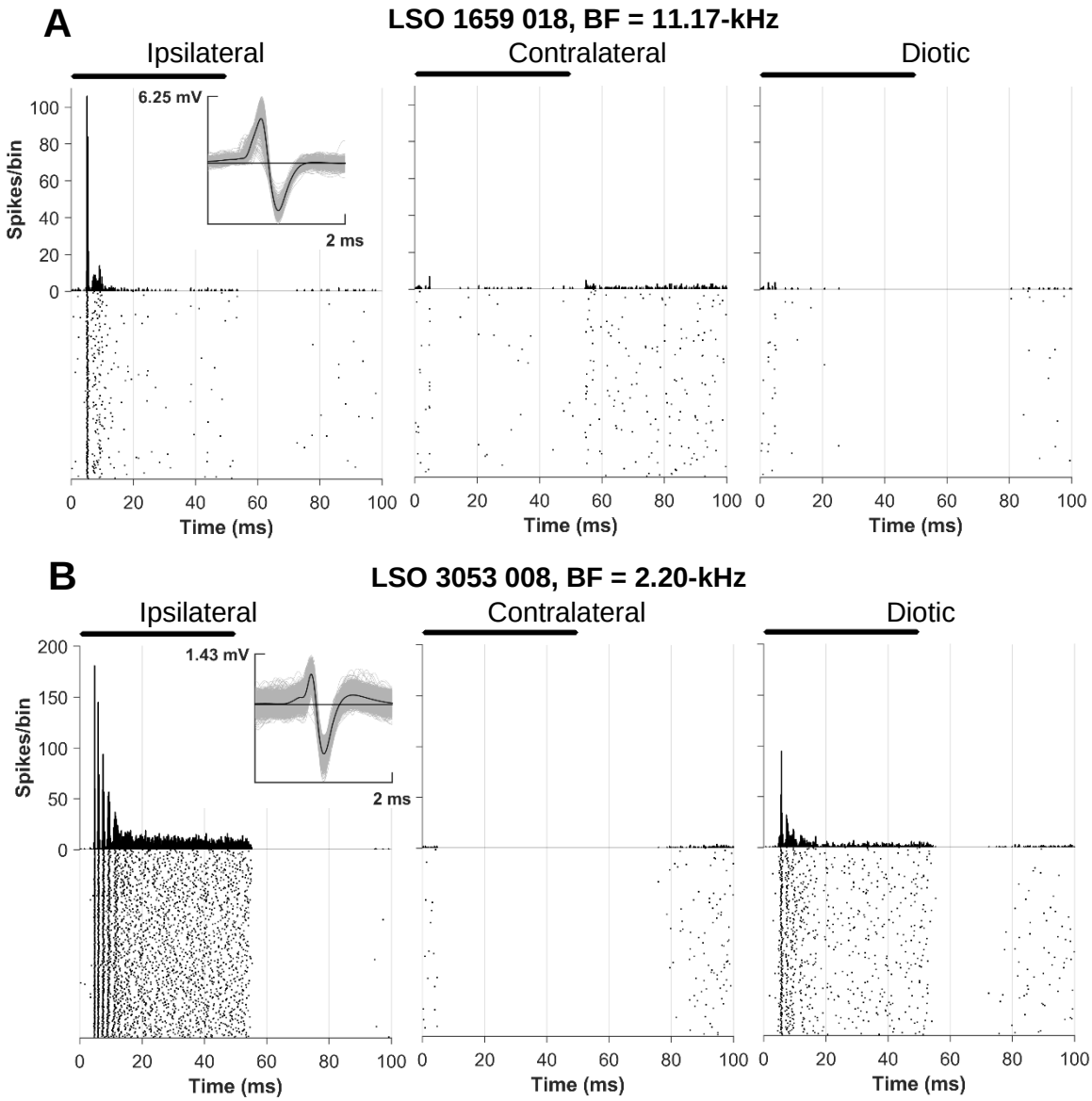

Figure 33 **LSO units show PN (unit** A, **termed 'bimodal' in Tsuchitani, 1988a) or onset chopper (unit** B**, termed 'slow chop' in Tsuchitani, 1988a) responses to ipsilateral-only BF tones**. Black bar represents stimulus duration. In response to contralateral-only stimulation, inhibition from MNTB depresses firing rates below SR. In diotic presentation, LSO units either do not respond at all or respond similarly to ipsilateral-only stimulation but with reduced firing rates.

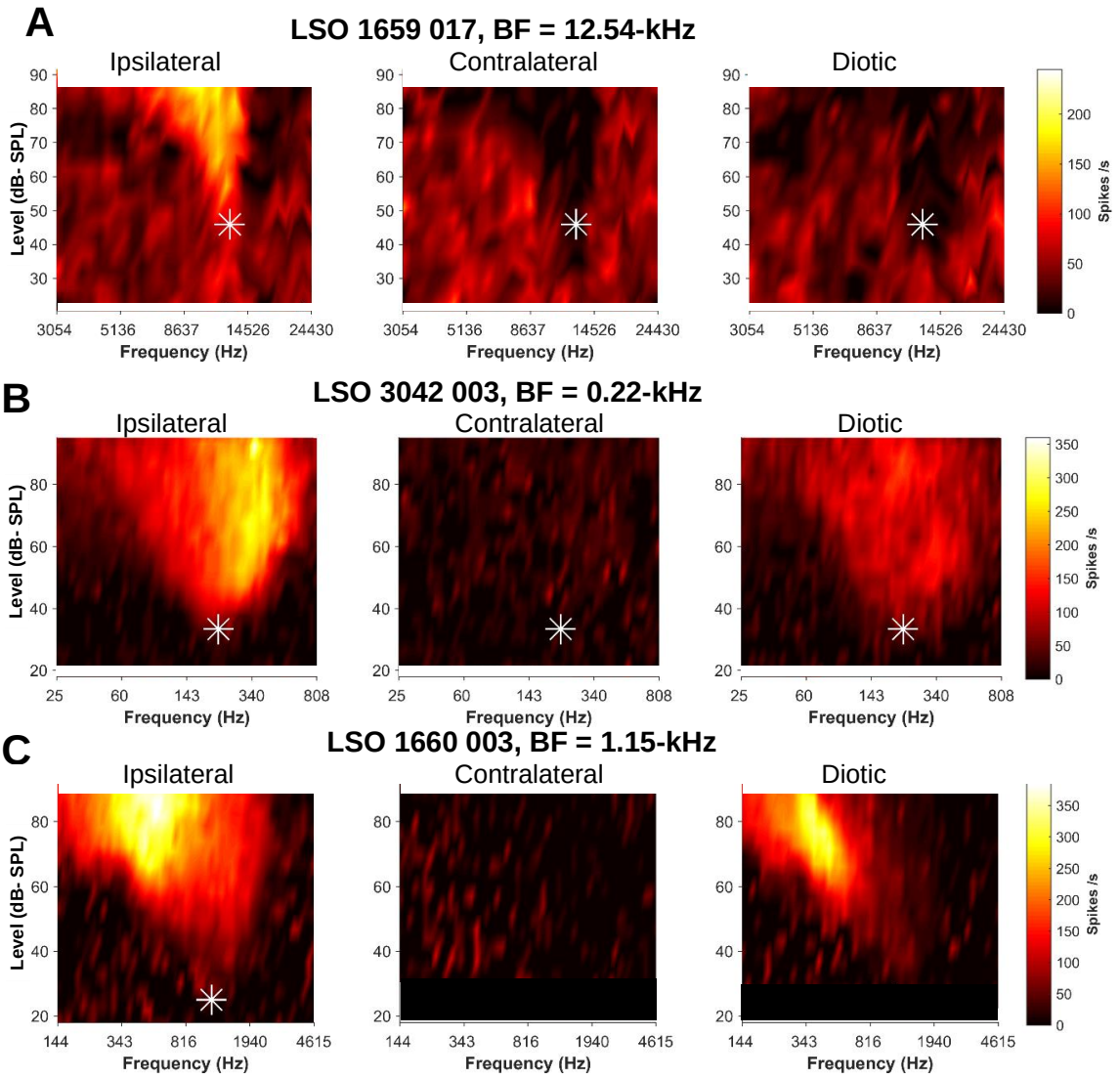

Figure 34 **Binaural suppression is seen in diotic stimulation of LSO units**. Fitted BF and threshold are marked by the white asterisks.

A: High-BF LSO unit. In response to diotic stimulation, the inhibitory inputs are strong enough to completely cancel the excitatory field seen in ipsilateral stimulation. This unit had a large enough SR to allow the inhibitory filter to be seen in contralateral stimulation.

B and **C**: Low-BF LSO units. In **B**, the diotic receptive field shows the same shape as the ipsilateral receptive field but is considerably weaker. In **C**, the diotic receptive field is also weaker than the ipsilateral field, however there is a residual response at low frequencies, presumably because the inhibitory filter does not entirely overlap the ipsilateral.

As reported in other species, low-BF LSO units of the guinea pig showed ITD- and IPD-sensitivity. Binaural beat responses for LSO units are shown in **Figure 35** below and **Figure 39 (page 90)**.

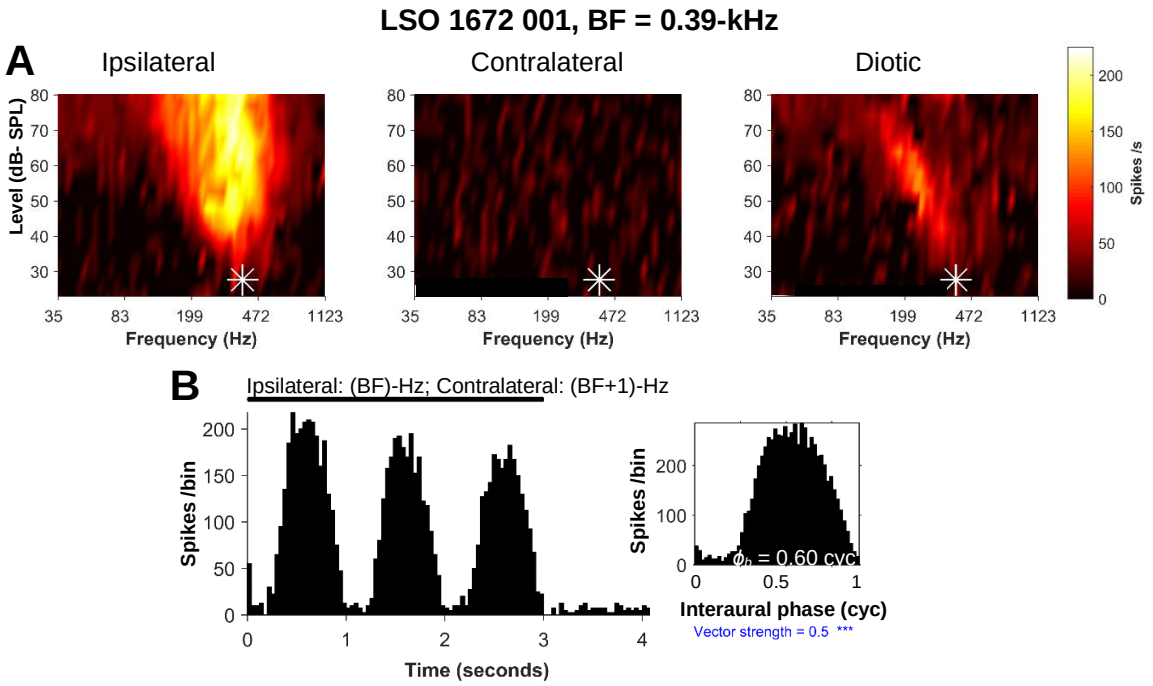

Figure 35 **Low-BF LSO units show 'trough type' binaural beat responses**.

A: Receptive fields of a low-BF LSO unit. As with other low-BF LSO units (see **Figure 34, page 85**), the response to diotic tones is weaker than ipsilateral-only. Fitted BF and threshold are marked by the white asterisks.

B: Binaural beat response at unit BF. The unit is IPD-sensitive, showing a broad response to binaural beats (vector strength [VS] = 0.5). Inset is the period histogram, showing a mean interaural phase (ϕ_b) of close to 0.5 cycles (0.60), indicative of a trough-type response (see **Section 3.1.4.2, page 62**).

3.3.5 Medial superior olive

Guinea pig MSO units showed weak responses to monaural stimulation, but stronger responses to diotic tones. For all units, diotic responses were greater than monaural responses of either ear (see **Section 3.3.1, page 70**). Receptive field examples are shown in **Figure 12 (page 40)**, **Figure 36** below, **Figure 37 (page 88)**, **Figure 39 (page 90)** and **Figure 40 (page 92)**.

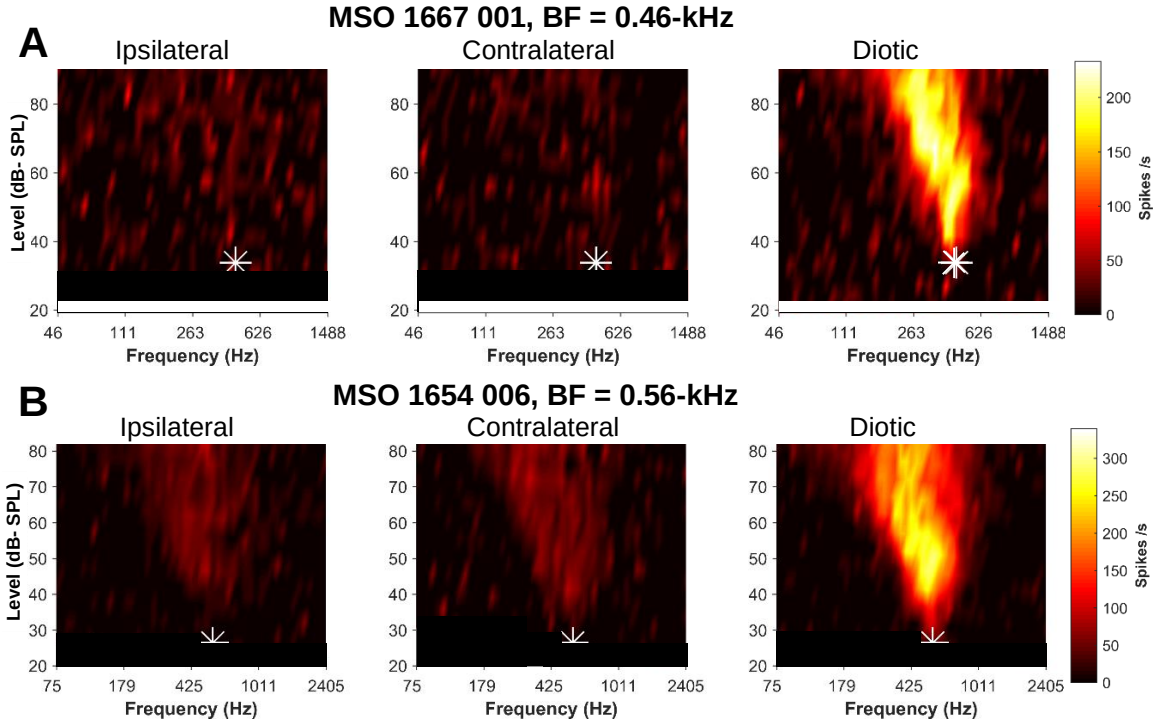

Figure 36 **MSO units showing weak monaural receptive fields but clear binaural facilitation**. Fitted BF and threshold are marked by the white asterisks.

A: In this unit, no activity is evoked by presentation of monaural tones, however, diotic stimulation elicits a strong response.

B: In this unit, weak V-shaped receptive fields are seen monaurally, but diotic stimulation results in a strong, binaurally-facilitatory response.

Guinea pig MSO units showed similar ITD- and IPD-sensitivity to that described in other species. Binaural beat responses are shown in **Figure 37 (page 88)** and **Figure 39 (page 90)**.

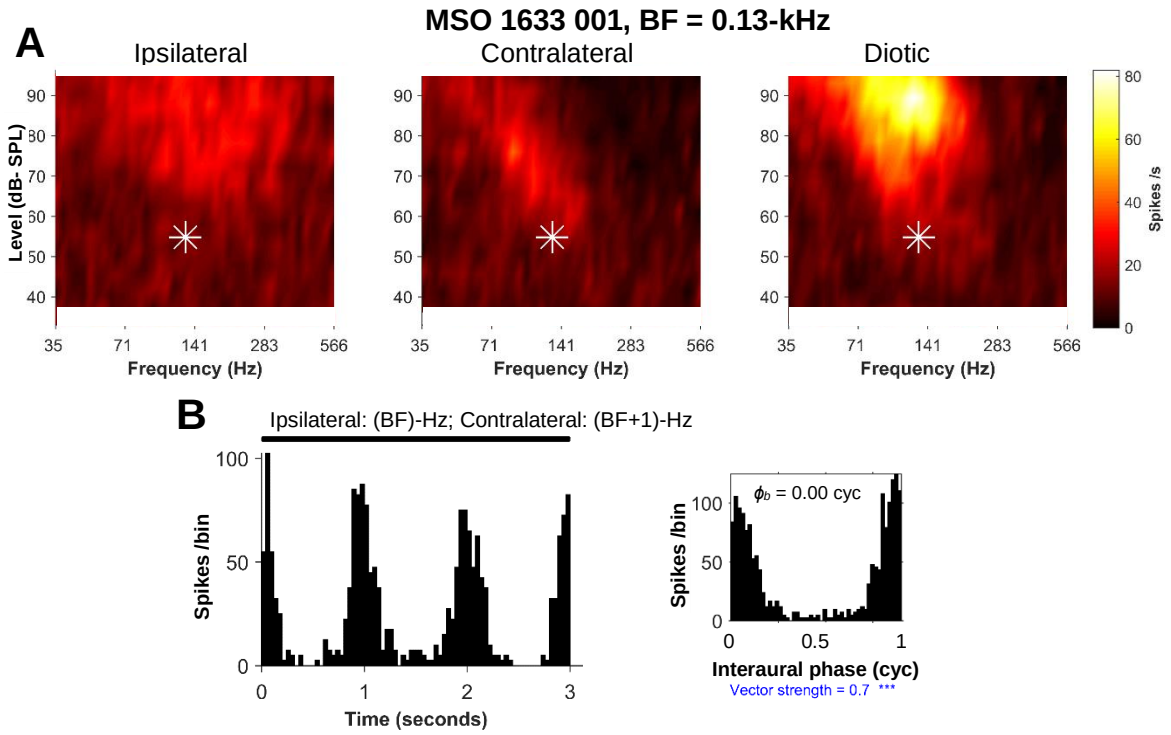

Figure 37 **MSO units are ITD-sensitive**.

A: Receptive fields measured over 10-repetitions. As with other MSO units (see **Figure 36, page 87)**, this unit showed binaural facilitation. Differences in the ipsilateral and contralateral filter shapes show that this only occurs in the overlap of the two monaural filters. Fitted BF and threshold are marked by the white asterisks.

B: Binaural beat response at BF. Contrary to the low-BF LSO unit shown earlier (see **Figure 35, page 86**), this unit show a sharper response to binaural beats (VS = 0.70) and a mean interaural phase (ϕ_b) of 0.00 cycles. This is indicative of a peak-type response (see **Section 3.1.3.2, page 56**). Average spike waveform and waveform analysis of the binaural beat response is shown in **Figure 14 (page 43)**. Lesion evidence for the location of this unit is shown in **Figure 19 (page 68)**. Average spike waveforms for the responses in **B** are shown in **Figure 38A (page 89)**.

MSO units showed either negative-going (n = 40) or positive-going (n = 20) spike waveform shapes. Typical examples of these types be found in **Figure 38 (page 89)**. Units showing positive-going spikes had greater average peak-to-peak amplitudes, range: 0.34 to 2.27-mV; mean±sd: 1.12 ± 0.60 mV, than units showing negative-going spikes, range: 0.20 to 0.68-mV; mean±sd: 0.38 ± 0.38 mV. A two-sample t-test showed that the means were significantly different ($p<0.001$).

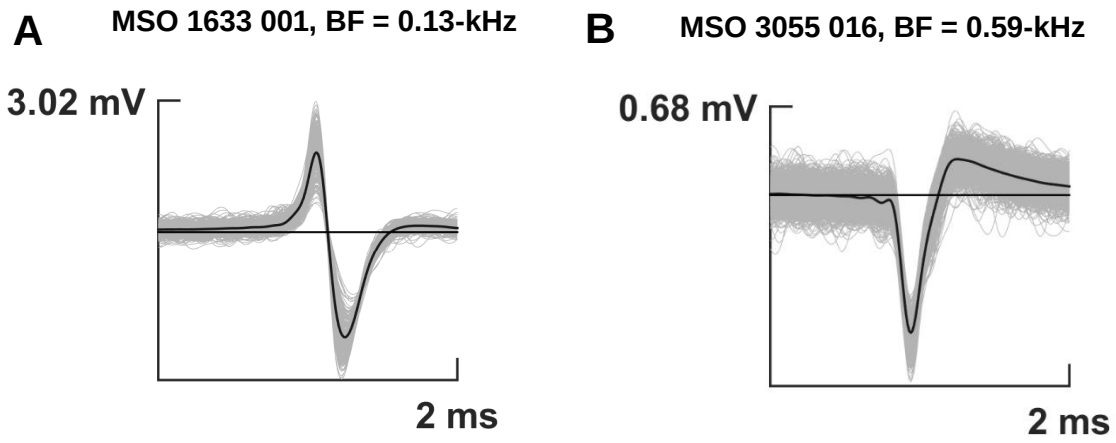

Figure 38 **Typical examples of MSO unit average spike waveforms.**

A: MSO unit showing a positive-going, biphasic spike waveform shape.

B: MSO unit showing a negative-going spike waveform shape.

Disambiguating MSO and low-frequency LSO units online was possible due to the differences in their responses to diotic tones and binaural beats (see **Figure 39, page 90**). LSO units show a broad response to binaural beats, with mean interaural phase (ϕ_b) close to 0.5 cycles (see **Figure 35, page 86**); MSO units show a sharper response with ϕ_b near 0 cycles (see **Figure 37, page 88**). LSO unit diotic receptive fields are weaker than their ipsilateral fields (examples in **Figure 34, page 85** and **Figure 35, page 86**); MSO units show stronger diotic receptive fields than ipsilateral (see **Figure 36, page 87**). This was also true across the unit population (see **Figure 23, page 74**). Although not a substitute for anatomical proof, as the MSO and LSO distributions of these properties overlap, these differences were useful in forming an online hypothesis of electrode location.

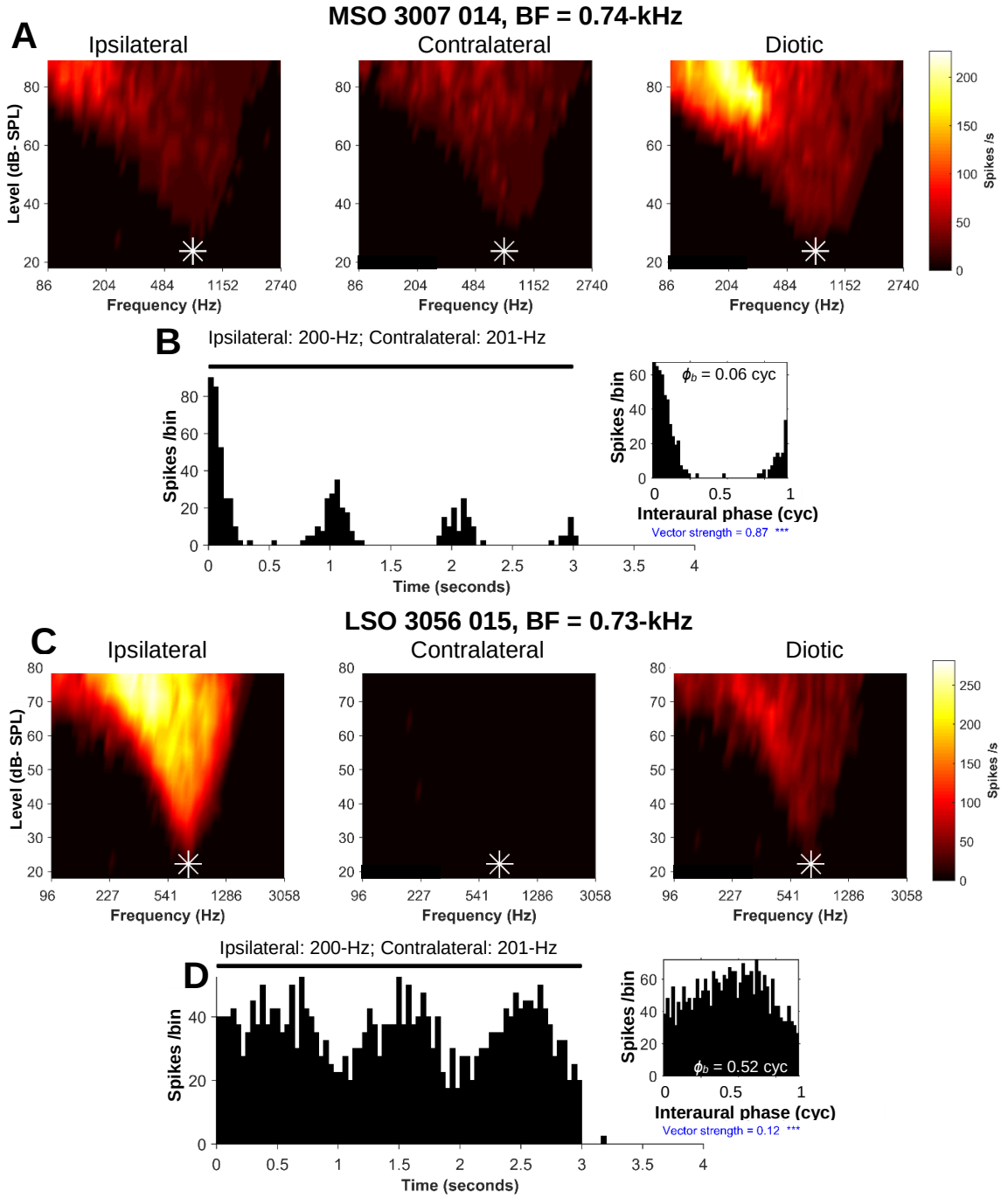

Figure 39 **Comparison of the binaural beat and receptive field responses of similar-BF MSO and LSO units**. Fitted BF and threshold are marked by the white asterisks.

A: MSO units show weak monaural receptive fields but a strong response to diotic stimulation.

B: MSO units show frequency-dependent mean interaural phases (ϕ_b), generally close to zero cycles (in this case, 0.08 cycles). The range of interaural phase that elicits responses is also narrow: VS is 0.87 in this case.

C: In contrast, LSO units show a stronger response to ipsilateral-only stimulation than diotic stimulation, and zero response to contralateral-only stimulation.

D: LSO units show ϕ_b of close to 0.5 cycles across frequency, in this case 0.52 cycles. Additionally, the response is broader: VS is 0.12 in this case.

3.3.6 Temporal synchrony

Units of the guinea pig SOC show strong phase-locking in response to pure tone stimulation, as has been reported in other species. MSO units showed the highest maximum vector strengths of all units sampled in the SOC. This likely reflects their coincidence-detection operation of several inputs that themselves likely show temporal enhancement (see **Sections 3.1.3.1, page 53** and **5.3.8, page 163**).

Figure 40 (page 92) shows an exemplar MSO unit's synchrony of responses to pure tone stimuli. Monaural receptive fields, shown in **Figure 40A**, are weak, but binaural facilitation occurs in diotic stimulation. In **Figure 40B**, VS is plotted as a function of pure tone frequency. Very high VSs are seen at all frequencies, exceeding the range found in guinea pig ANFs, plotted as grey dotted lines (Palmer & Russell, 1986), however, fewer conditions in the monaural presentations reached a high enough spike rate to reach statistical significance (see **Section 2.4.1, page 45**). **Figure 40C** shows PSTHs and spike rasters of the responses to a 325-Hz tone. Monaural responses are weak, but firing is extremely temporally-precise when it does occur, likely reflecting of monaural coincidence detection. A much greater response occurs to diotic stimulation, maintaining high temporal precision (VS = 0.97).

In **Figure 41 (page 93)**, I have shown responses of an exemplar MNTB unit to pure tones. **Figure 41A** shows the PSTH and raster of unit responses to a 200-Hz tone. This elicits highly precise temporal firing (VS = 0.9), which has been described as a 'picket-fence'. **Figure 41B** shows the unit's receptive field; in **Figure 41C**, VSs for pure tone stimulation are plotted using the same frequency scale. At all frequencies sampled, VS is high, with most points exceeding the range of VS found in guinea pig ANFs, plotted as grey dotted lines (Palmer & Russell, 1986).

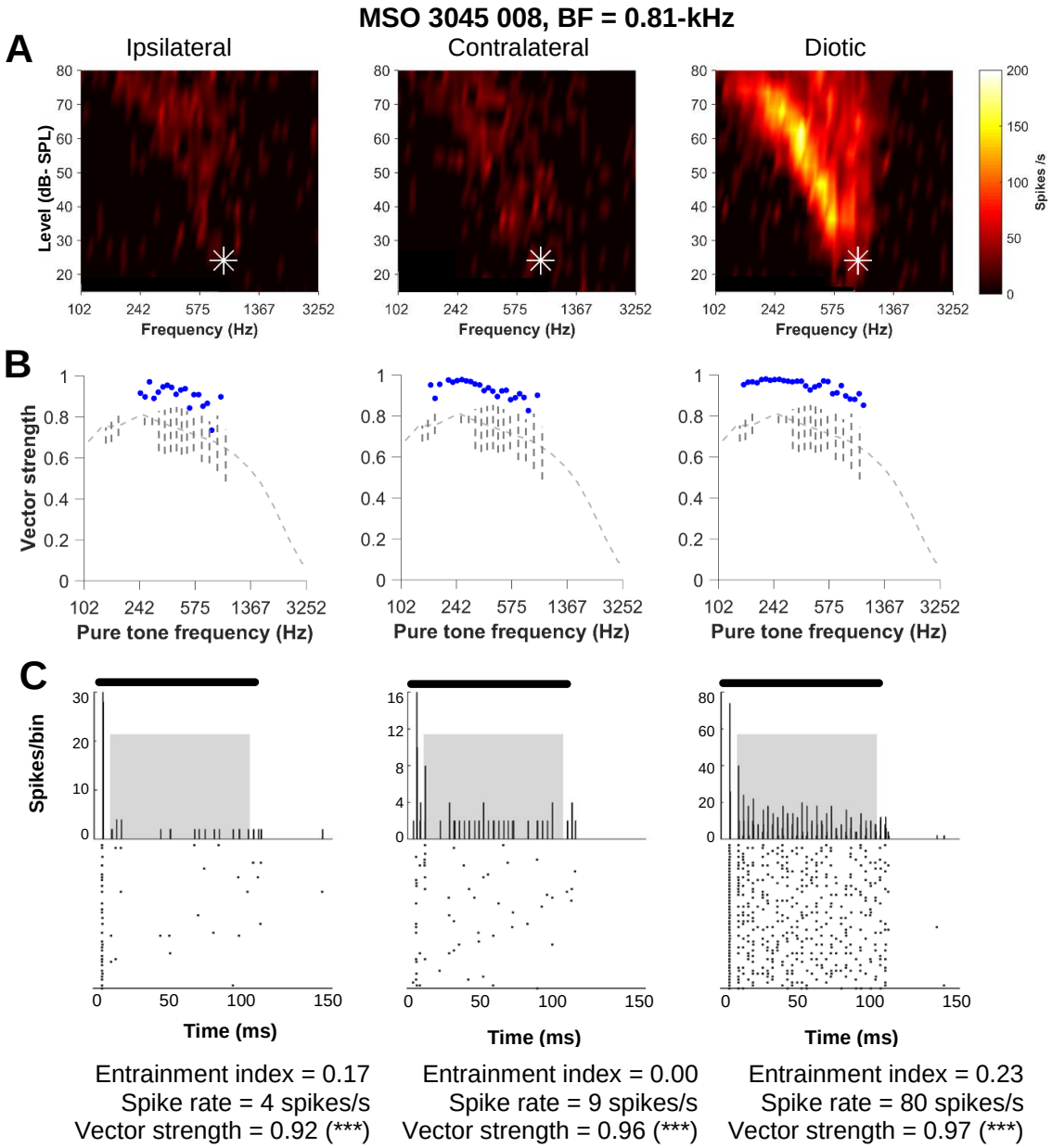

Figure 40 **Precise temporal coding in an MSO unit**.

A: Unit receptive fields. Fitted BF and threshold are marked by the white asterisks.

B: Vector strength as a function of pure tone frequency, shown on the same frequency scale as **A**.

C: PSTHs and spike rasters of the responses to a 325-Hz tone at 61-dB SPL. In order to avoid onset and offset effects, only spikes occurring in the grey shaded area were analysed in **B**.

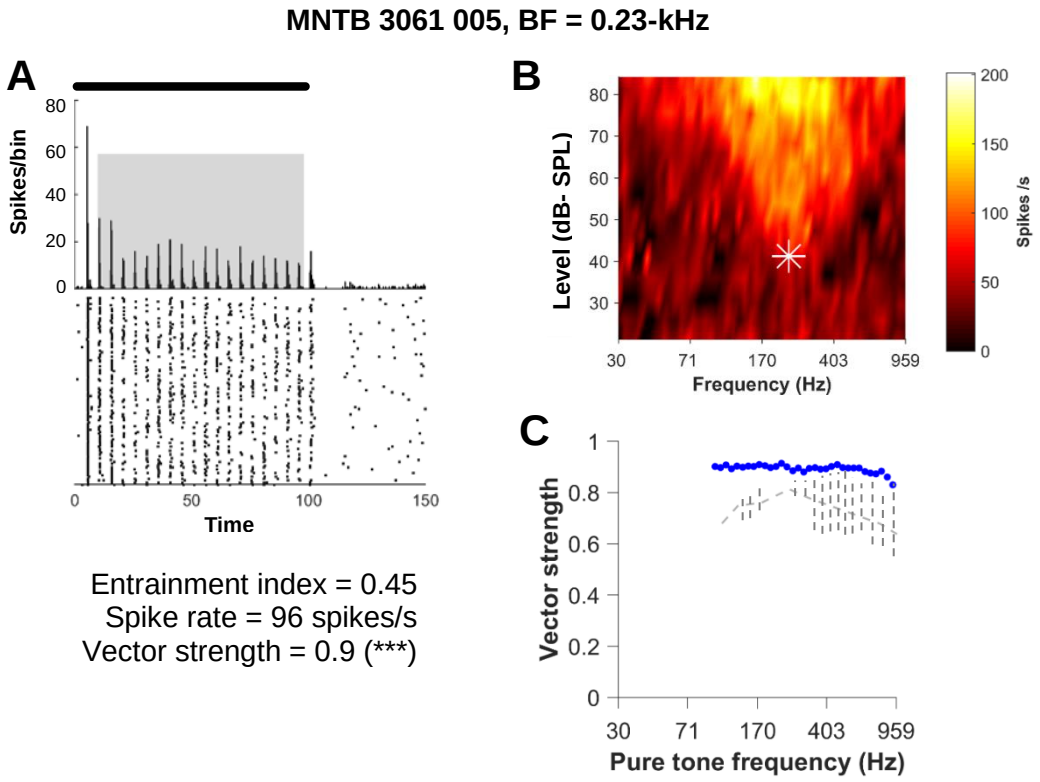

Figure 41 **Low-frequency MNTB neurons show high synchrony**.

A: PSTH of unit responses to a fixed-phase 200-Hz tone at 84-dB SPL to the contralateral ear (represented by black bar).

B: Receptive field for this unit. Fitted BF and threshold are marked by the white asterisk.

C: VS against pure tone frequency. To avoid analysing onset and offset responses, spikes were only analysed if they occurred in the shaded region in **A**. All points were sampled using tones of at least 30-dB SL.

The population data from SOC (MNTB: *n* = 3; MSO: *n* = 13; LSO: *n* = 8) has been plotted in **Figure 42 (page 94)**; note that the majority of points show higher synchrony than that found in ANFs, and that VS are greatest in the ipsilateral stimulation for the LSO and in diotic stimulation in the MSO.

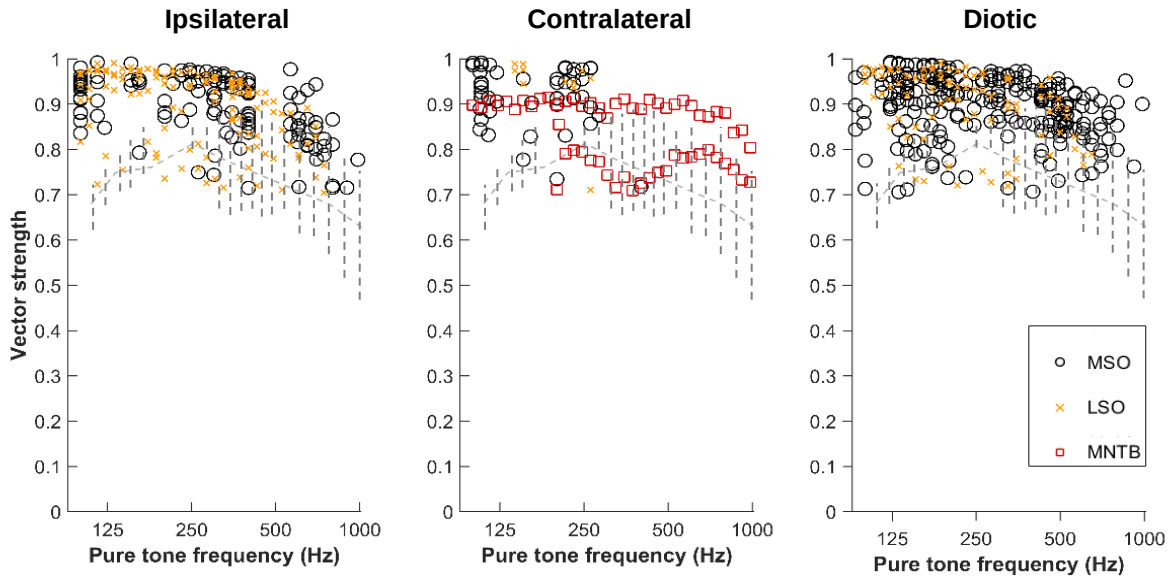

Figure 42 **High synchrony to pure tones of varying frequency in units of the SOC.** The majority of points in all SOC nuclei are above the range found in guinea pig ANFs, shown as grey dotted lines (Palmer & Russell, 1986). Each point represents a significant VS from pure tone stimulation at one frequency from one unit. MNTB: *n* = 3; MSO: *n* = 13; LSO: *n* = 8.

3.4 DISCUSSION

In this section, I will discuss the basic response properties of the guinea pig SPN, MNTB, MSO and LSO in comparison with other species. Note, however, that the highly-stringent classification scheme outlined in **Section 3.2.1 (page 65)** likely led to the exclusion of MNTB and SPN units with atypical properties. In addition, the low-BF bias of this population qualifies the comparisons of the BF-dependent properties, unit FSL, $Q_{10\text{-}dB}$ and threshold, across species.

3.4.1 Medial nucleus of the trapezoid body

MNTB thresholds are higher in guinea pig than reported in other species (see **Appendix 7.2: Table 1a, page 183**), however this is likely due to the low-frequency bias of the SOC population

in this Thesis. Probably for the same reason, mean $Q_{10\text{-}dB}$ reported in other species is also higher than the sampled population (see **Appendix 7.2: Table 1b, page 183**). However, guinea pig MNTB SRs were similar to those reported in other species (see **Appendix 7.2: Table 1c, page 183**).

MNTB cells in this study were found to have entirely [0·E] responses, with nonsignificant inhibition occurring in response to diotic tonal stimulation. This is consistent with other studies, however the only studies that report a percentage state that around 80% of MNTB cells were [0·E] (cat – Guinan *et al.*, 1972b; rat – Sommer *et al.*, 1993).

The presence of prepotentials was used as a criterion for MNTB in this study; prepotentials have been reported in MNTB cells of other species (cat – Guinan & Li, 1990; cat – Smith *et al.*, 1998; gerbil – Kopp-Scheinpflug *et al.*, 2003, rat and mouse – Kopp-Scheinpflug *et al.*, 2008). However, in some species, MNTB units do not show always show prepotentials (gerbil – Kopp-Scheinpflug *et al.*, 2003).

Finally, evidence of high VSs to low-frequency pure tones, as seen in MNTB units in this study, was found in cat (Joris & Yin, 1998; Smith *et al.*, 1998; Tollin & Yin, 2005) and in gerbil (Kopp-Scheinpflug *et al.*, 2003; Dehmel *et al.*, 2005).

3.4.2 Superior paraolivary nucleus

The SPN population BF range of this study, mean±sd: 8.16 ± 4.98-kHz, differs from that found in the SPN of the gerbil, where the majority of neurons have BFs below 5-kHz (Behrend *et al.*, 2002).

The prevalence of offset-only responses varies across the SPN of different mammals. In the gerbil, one study found that only 5% of SPN units showed offset-only responses (Behrend *et al.*, 2002); a different study found that 34% of SPN units were offset-only (Dehmel *et al.*, 2002). Both the SPN of mouse (64%, Kopp-Scheinpflug *et al.*, 2011) and rat (95%, Kulesza *et al.*, 2003) have a majority of units with offset-only responses. In this study, SPN was identified using offset-only responses as a criterion; this does not preclude the existence of other unit types in the guinea pig SPN.

For the remaining response properties, where possible, only the gerbil offset-only subpopulations will be compared with the data of guinea pig and other species.

The range of thresholds of guinea pig SPN units is in line with that reported in other species, however the mean threshold is higher (see **Appendix 7.2: Table 2a, page 183**). $Q_{10\text{-}dB}$s of the guinea pig SPN are lower than those reported in other species (see **Appendix 7.2: Table 2b, page 183**). SRs of the guinea pig SPN were low, in line with reports from studies in other species (see **Appendix 7.2: Table 2c, page 183**).

Gap detection sensitivity of guinea pig SPN resembles that of rat SPN (Kadner & Berrebi, 2008). Additionally, the response to increasing tone duration in guinea pig SPN has been reported in rat (Kadner *et al.*, 2006). Finally, entrainment to SAM-BF tones of guinea pig SPN resembles that of rat SPN (Kulesza *et al.*, 2003).

3.4.3 Lateral superior olive

LSO units showed spontaneous rates of mean±sd: 9.0 ± 14.3 spikes/s. To my knowledge, there are no other published data of other species with which to compare.

LSO units of the guinea pig showed sensitivity to ILDs in BF pure tones similar to the LSO of other species (cat – Tsuchitani & Boudreau, 1969; cat – Joris & Yin, 1995; cat – Smith *et al.*, 1998; bat – Park *et al.*, 2004, gerbil – 2008; cat – Tollin & Yin, 2005; gerbil – Magnusson *et al.*, 2008; cat – Tollin *et al.*, 2008; cat – Tsai *et al.*, 2010).

Guinea pig LSO unit thresholds were mean±sd: 30 ± 10 dB-SPL. This is higher than in cat, mean±sd: –4 ± 8 dB-SPL (Tsuchitani, 1997), however it is possible that the BF distributions of the two samples are not comparable due to the low-BF oversampling in this Thesis.

LSO $Q_{10\text{-}dB}$s in guinea pig were mean±sd: 1.79 ± 1.12; this is broader than in cat, mean±sd: 6.96 ± 2.58 (Tsuchitani, 1997), however, this difference could be due to the BF skew in the population of this Thesis.

FSLs in guinea pig LSO units were mean±sd: 5.7 ± 1.0-ms, similar to that reported in cat, mean±sd: 6.53 ± 0.71-ms (Tsuchitani, 1997).

Low-frequency units of the LSO of guinea pig showed sensitivity to TFS IPD and ITD, as reported in the cat (Tollin & Yin, 2005).

All LSO units of the guinea pig showed [E·I] responses, similar to what is seen in cat LSO (92%, Tsuchitani, 1977).

Finally, vector strengths in the guinea pig LSO responses to pure tones in the ipsilateral ear are greater than VSs of ANFs; this is consistent with studies in the cat (Joris & Yin, 1998; Tollin & Yin, 2005).

3.4.4 Medial superior olive

All guinea pig MSO units showed ITD-sensitivity to low-frequency binaural beats. This is comparable to the chinchilla (100%, Langford, 1984), but higher than the cat (79%, Yin & Chan, 1990) and gerbil (56%, Brand *et al.*, 2002).

All units of the guinea pig MSO showed stronger diotic responses than monaural responses, with 41% of units also showing [E·E] responses. This is in line with what has been shown in dog (65%, Goldberg & Brown, 1968), cat (58%, Yin & Chan, 1990) and gerbil (67%, Brand *et al.*, 2002).

In the guinea pig MSO, 16% of units showed [E·0] responses and 14% showed [0·E] responses. This is similar to that reported in cat: the combined [E·0] and [0·E] percentage reported in cat is 36% (Yin & Chan, 1990).

In guinea pig MSO, 29% of units did not show monaural responses ([0·0]); this is lower than in chinchilla MSO, where 100% of units were [0·0] (Langford, 1984).

Guinea pig MSO $Q_{10\text{-}dB}$s and FSLs are similar to those reported in other species (see **Appendix 7.2 Table 3a** and **3b, page 184**).

MSO units in guinea pig showed SRs mean±sd: 9.8 ± 14.8 spikes/s, with 23% of units having SRs less than 1 spike/s and 90% with SRs less than 20 spikes/s. This is higher than the proportion found in cat, where 74% of units had SRs less than 1 spike/second, and 92% had SRs less than 20 spikes/s (Yin & Chan, 1990).

Finally, guinea pig MSO showed enhanced phase-locking to low-frequency pure tones compared to ANFs; this has also been found in other species (chinchilla – Langford, 1984; cat – Yin & Chan, 1990).

3.4.5 Summary

The findings of this chapter can be summarised as follows:

- A class of units in the guinea pig SOC show a characteristic three-component waveform, primary-like responses and [0·E] response resembling MNTB units reported in other species. These are presumably MNTB units.
- Offset-only units exist in the guinea pig SOC, responding to stimulation in the contralateral ear with a rebound excitatory response, consistent with those reported in the SPN of other rodents; these are presumed to be SPN units.
- Anatomically-localised LSO units of the guinea pig show the characteristic sensitivity to ILDs and a binaural suppression to diotic tones observed in other species; additionally, low-BF LSO units also show sensitivity to IPDs and ITDs as reported in other species.
- Anatomically-localised MSO cells of the guinea pig are ITD- and IPD-sensitive, as reported in other species. Additionally, they show pronounced responses to diotic tonal stimulation, with many units showing binaural facilitation.
- The basic response properties of the nuclei of the guinea pig SOC show many similarities to those of other species, particularly the MSO and LSO, validating it as a suitable model to study binaural interaction underlying the temporal coding of binaural pitch.

Chapter 4

Binaural pitch

Abstract

Psychophysical experiments have found pitch-evoking stimuli that require combination of information originating from both sides of the auditory system; these are termed ‘binaural pitch’ stimuli. However, previous neurophysiological studies using binaural complex tones have not observed correlates of this effect.

Recordings were made of single units of the MSO and LSO, known to perform binaural integrative computations such as ITD-sensitivity and ILD-sensitivity, to stimuli that elicit a binaural pitch percept. Dichotic complex tones consisted of consecutive harmonics of a particular F0, usually 100-Hz, presented one to each ear. Dichotic harmonic tone complexes consisted of odd-only harmonics presented to one ear, and even-only harmonics to the other. Both cosine- and random-phase HTCs were presented.

Analysis of the timing patterns of MSO and LSO cell responses to these stimuli reveals a response to the binaural pitch. In addition, a clear neural correlate of pitch to the overall F0 across ears was evident irrespective of presentation, *i.e.* there was no dominance of one ear over the other in MSO or LSO unit responses to binaural pitch.

These findings suggest a role of the MSO and LSO in the neural machinery underlying the perception of the pitch of binaural complex tones.

4.1 INTRODUCTION

4.1.1 Dichotic complex tones

In response to a dichotic complex tone (DCT), consisting of two consecutive harmonics of a fundamental frequency (F0) presented one to each ear, humans perceive a weak pitch at the F0. When a frequency shift is applied to the DCT, the pitch percept shifts proportionally. This is termed the first effect of pitch shift (see **Section 1.3.1, page 13**) and is equal to that observed in monaural presentation (Houtsma & Goldstein, 1971). Human psychophysical data showing this equivalence are plotted in **Figure 43** below.

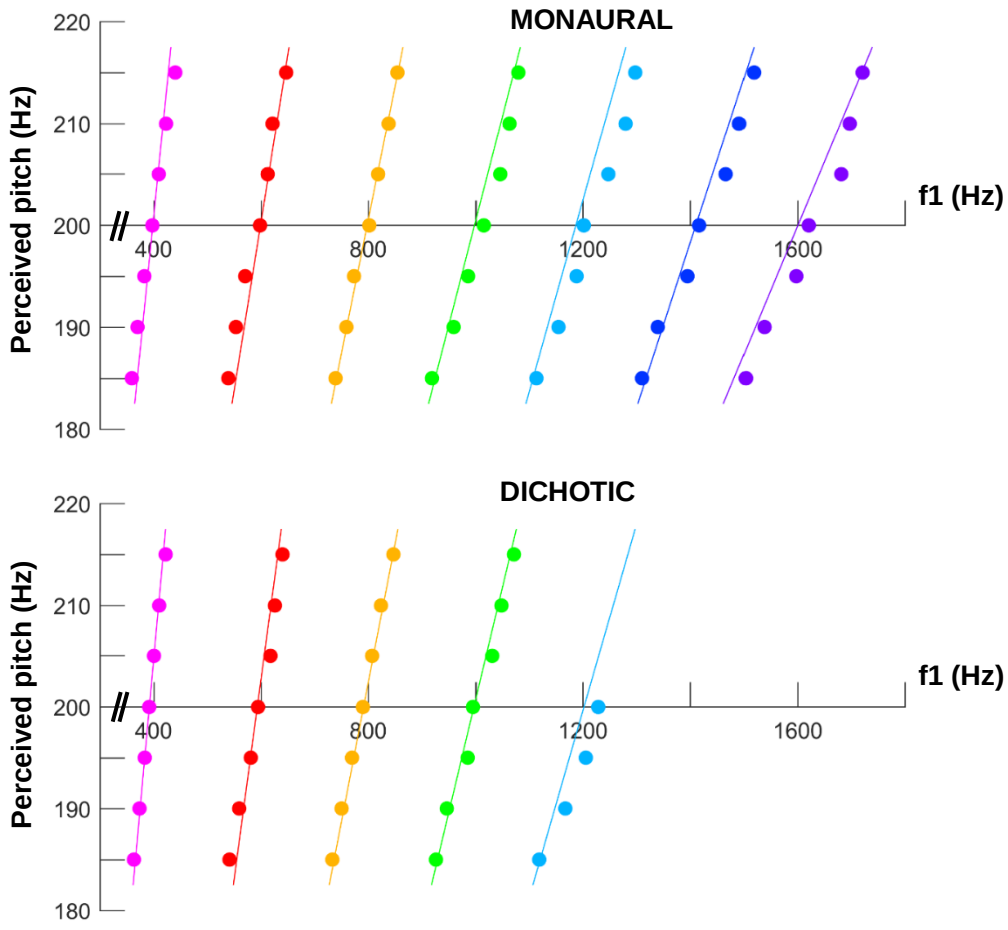

Figure 43 **The first effect of pitch shift is seen in human psychophysics for two-tone complexes, even if the harmonics are presented across ears (dichotically)**. Lines indicate the predicted percepts using de Boer's rule (see **Section 1.3.1, page 13**). Both modes of presentation show that the perceived pitch of the subject shifts in proportion to the linear shift of the entire complex. Replotted from Houtsma & Goldstein, 1971.
f1 – the frequency of the first tone in a given complex; the second is (f1 + F0)-Hz. Each colour represents the data gathered for a different harmonic rank.

Presenting the complex across ears eliminates the possibility of cochlear distortions producing a physical vibration at the place of the F0. Therefore, as both monaural and dichotic presentations lead to the same pitch value, it is theorised that information underlying pitch is combined across ears, termed 'binaural fusion' (see **Section 1.4.3, page 25**) prior to pitch being determined.

Houtsma, 1981 reported that the pitch of DCTs is robust across moderate ILD, and Beerends & Houtsma, 1988 reported that presenting two simultaneous DCTs leads to the perception of two F0s, with not much degradation of performance compared to when complexes were presented monaurally.

4.1.2 Dichotic harmonic tone complexes

Harmonic tone complexes (HTCs) evoke a strong pitch percept at the F0 in monaural (all harmonics to one ear) or diotic (all harmonics to both ears) presentations. In a dichotic HTC, harmonics are distributed alternately across ears, such that one ear receives odd-only harmonics and the other even-only harmonics. Regardless of presentation, humans perceive a single, clear pitch at the F0 of the overall stimulus (Houtsma & Goldstein, 1971, 1972; Bernstein & Oxenham, 2003). As with DCT (see **Section 4.1.1, page 103**), this suggests that binaural fusion occurs before pitch extraction.

Given that dichotic stimulation does not affect the perceived pitch of HTCs, it might be expected that properties of audition that benefit from increased harmonic resolvability (see **Section 1.3.2 page 17**) would improve under dichotic versus diotic stimulation. However, this is not the case; identical F0-difference limens (F0DL), the minimum F0 increment that can be discriminated,

were found in both diotic and dichotic stimulation (Bernstein & Oxenham, 2003). This suggests that the auditory system cannot make use of the increased resolution, implying that information underlying pitch is binaurally-fused prior to reaching the central pitch processor, where pitch is thought to be determined (see **Section 1.4.3.2, page 28**).

4.1.3 Physiological investigations

While a great number of neurophysiological studies have been conducted into the monaural processing of pitch in the auditory system (see **Sections 1.3.3, page 19** and **5.4.1, page 167**), few have investigated pitch arising from binaural stimulation. No prior neurophysiological study has demonstrated correlates of binaural integration of pitch seen in human psychophysics.

In response to dichotic HTCs, clusters of the guinea pig central nucleus of the inferior colliculus (CNIC) do not appear to respond to the F0 of the overall complex. Instead, clusters responded to the monaural temporal envelope (ENV), the 2×F0 spacing between components (Shackleton *et al.*, 2009). Multi-unit responses to diotic and dichotic presentations showed different neural correlates of pitch, contrary to the psychophysical studies.

A follow-up study from the same lab showed that the majority of guinea pig CNIC cells with BFs greater than 2-kHz act to segregate dichotic stimuli (Nakamoto *et al.*, 2014). Competing HTCs of different F0s, 125-Hz and 145-Hz, were presented one to each ear. It was observed that the majority of units responded preferentially at the period of the pitch of one of the F0s. However, some cells showed a response to the common periodicity (20-Hz) across ears. It is possible these cells would show a similar response to the superior olivary complex (SOC) cells shown in **Section 4.3.2 (page 117)**.

Binaural pitch integration requires a positive combination of monaural inputs; the first stage of the auditory system at which this occurs is the medial superior olive (MSO). Cells of the MSO act as coincidence detectors of excitatory spherical bushy cell inputs arising from both ventral cochlear nuclei (VCN) (see **Section 3.1.3.2, page 56**). Monaural-only stimulation gives rise to a markedly weak or sometimes zero response; binaural stimulation, however, gives rise to a strong response (cat – Yin & Chan, 1990). The majority of research carried out on the MSO has focussed on the distribution of best ITDs, responses to ITDs and mechanisms affecting ITD-sensitivity (see **Sections 1.2.5.1, page 9** and **3.1.3.2, page 56**). Consequently, the role of the MSO in the binaural integration of pitch has not previously been explored. The data presented in this Chapter are the first recorded responses of binaurally-responsive cells showing a correlate of binaural pitch integration.

4.2 METHODS

4.2.1 Complex tones

Complex tones consisted of two consecutive harmonics of F0 (50-Hz or 100-Hz) presented in sine-phase. In DCTs, these were presented one to each ear; for monaural complex tones, these were both presented to one ear. The whole complex was then shifted upwards in frequency in $^{F0}/_{9}$-Hz steps. This initially creates an inharmonic stimulus with ambiguous pitch percept, however, after 9 shifts, finishes with a harmonic complex shifted up a harmonic rank. ENV periodicity is maintained throughout at the F0; temporal fine structure (TFS) is altered (see **Section 1.3.1, page 13**).

For a given unit, the harmonics in the DCT were selected so that they spanned the region of the receptive field that elicited the greatest response to single tones. Level was set at the lowest sound level that rate-level functions were maximal for the harmonic nearest BF; this was generally ~60-dB SPL per component.

Unit responses were analysed using shuffled autocorrelograms (SACs, see **Section 4.2.3, page 109**). The 9 SACs in response to the 9 frequency shifts were filtered in two dimensions using a Gaussian kernel and the MATLAB function `imfilt`. Responses were then transformed to polar coordinates using the MATLAB function `cart2pol`.

The population summary of DCT responses was created by adding the filtered 2D plots of all units and all presentations together, linearly, then scaling by the maximum of that plot. This is done to simulate the population response.

4.2.2 Harmonic tone complexes

Binaural pitch responses were assessed using 250-ms HTCs. F0 varied in octave steps between 31.25- and 500-Hz. HTCs consisted of all harmonics below either 3- or 5-kHz. Stimuli

were presented in three configurations: (1) dichotic: odd harmonics presented to one ear, even harmonics to the other; (2) monaural: odd harmonics alone or even harmonics alone to one ear; (3) diotic: odd and even harmonics to both ears. In the SOC, 50 repetitions were used; in the VCN, this was shortened to 25 repetitions. Humans report the pitch of odd-only harmonics to be one of three values: F0, slightly above 2×F0 and slightly below 2×F0, *i.e.* the pitch of this stimulus is ambiguous. Humans report the pitch of even-only harmonics to be an octave above the all-harmonics condition.

Cosine-phase HTCs (cpHTCs) were generated by adding all harmonics in cosine-phase. Level per component varied between 22-dB SPL for a diotic HTC of F0 = 31.25-Hz up to 38-dB SPL for a dichotic HTC of F0 = 500-Hz. Higher-threshold cells (see **Section 3.3.1, page 70**) did not always show a response to these HTCs; in these cells, HTCs were used with levels per component increased by 20-dB. Stimuli were presented at reasonably low sound levels to reduce the possibility of potential interaural crosstalk (IXT) (see **Section 5.1.3, page 144**) affecting results.

Frozen random phase HTCs (frzHTCs) were generated once by adding all harmonics together and randomising the phase of each harmonic. The level per component was kept the same as in cpHTCs. This waveform was presented 50 times.

Random phase HTCs (rpHTCs) were generated in exactly the same manner as frzHTCs, except that 50 versions of each waveform, each with differing randomised phases of harmonics, were generated, each presented once.

Unit responses to cpHTCs and frzHTCs were analysed using SACs. However, rpHTC responses could not be analysed this way as the SAC analysis requires an identical stimulus

waveform for all repetitions; they were instead analysed with all-order interspike interval histograms.

4.2.3 Interspike interval histograms

For all histograms of responses to pitch-evoking stimuli, the first 25-ms of evoked spikes were discarded to avoid onset effects.

First-order interspike interval histograms (FOISIs) were generated by taking each sweep of spikes in response to a particular stimulus, computing intervals between successive spikes, and plotting a histogram of all intervals, using a 0.2-ms bin-width.

All-order interspike interval histograms (AOISIs) were generated by taking each sweep of spikes and computing forward-going intervals from each spike to all other spikes in the same sweep, concatenating across all sweeps and then plotting a histogram. Bin-width was 0.2-ms. The largest overall peak in the AOISI summed across a population of units is considered a neural correlate of pitch (cat AN – Cariani & Delgutte, 1996a, 1996b).

Spike-train SACs were constructed to examine temporal representations of pitch (Louage *et al.*, 2004). Spike intervals were calculated for each spike in each sweep to all other spikes in other sweeps, forwards and backwards in time. A histogram was then calculated using a bin-width of 70-μs.

The summary of SAC responses to HTCs at each F0 was created by adding the SACs of all responses at that F0 in each stimulus configuration (*e.g.* odd-only harmonics presented monaurally) across all units where the spike rate exceeded 5 spikes/s. This was done to obtain a population response.

The normalised interval summary of SAC responses was created by finding the maxima of the SAC around the F0 and 2×F0 periods for each F0, for each unit, for each presentation, for each condition. This data was then normalised and plotted as a function of unit BF and stimulus F0.

4.3 RESULTS

4.3.1 Dichotic complex tones

When stimulated with harmonic DCTs, consecutive harmonics of an F0 presented one to each ear, MSO units respond to the period of the overall F0 (see **Figure 44, page 111**). This can be seen as a peak in the interspike interval (ISI) distributions at the period of the F0: in this case, 10-ms. Although there are differences in the responses to the two DCT configurations (ipsi-lower harmonic, contra-higher harmonic and *vice versa*), the largest peak in the AOISI (**Figure 44C**), the neural correlate of pitch (Cariani & Delgutte, 1996a, 1996b) occurs at the period of the overall F0 across ears, consistent with psychophysical data. This particular unit shows a greater response to the 500-Hz component in both stimulation modes: a large peak in the FOISIs (**Figure 44B**) occurs at 2-ms for both conditions.

Applying a frequency shift to the complex maintains the ENV period equal to the original F0 but alters the TFS of the stimulus (see **Section 1.3.1, page 13**). The pitch percept of humans shifts proportionally to the frequency shift: the first effect of pitch shift. This corresponds to a pseudoperiod in the stimulus waveform. The neural correlate of pitch response of MSO units to frequency-shifted DCTs shifts in a similar manner.

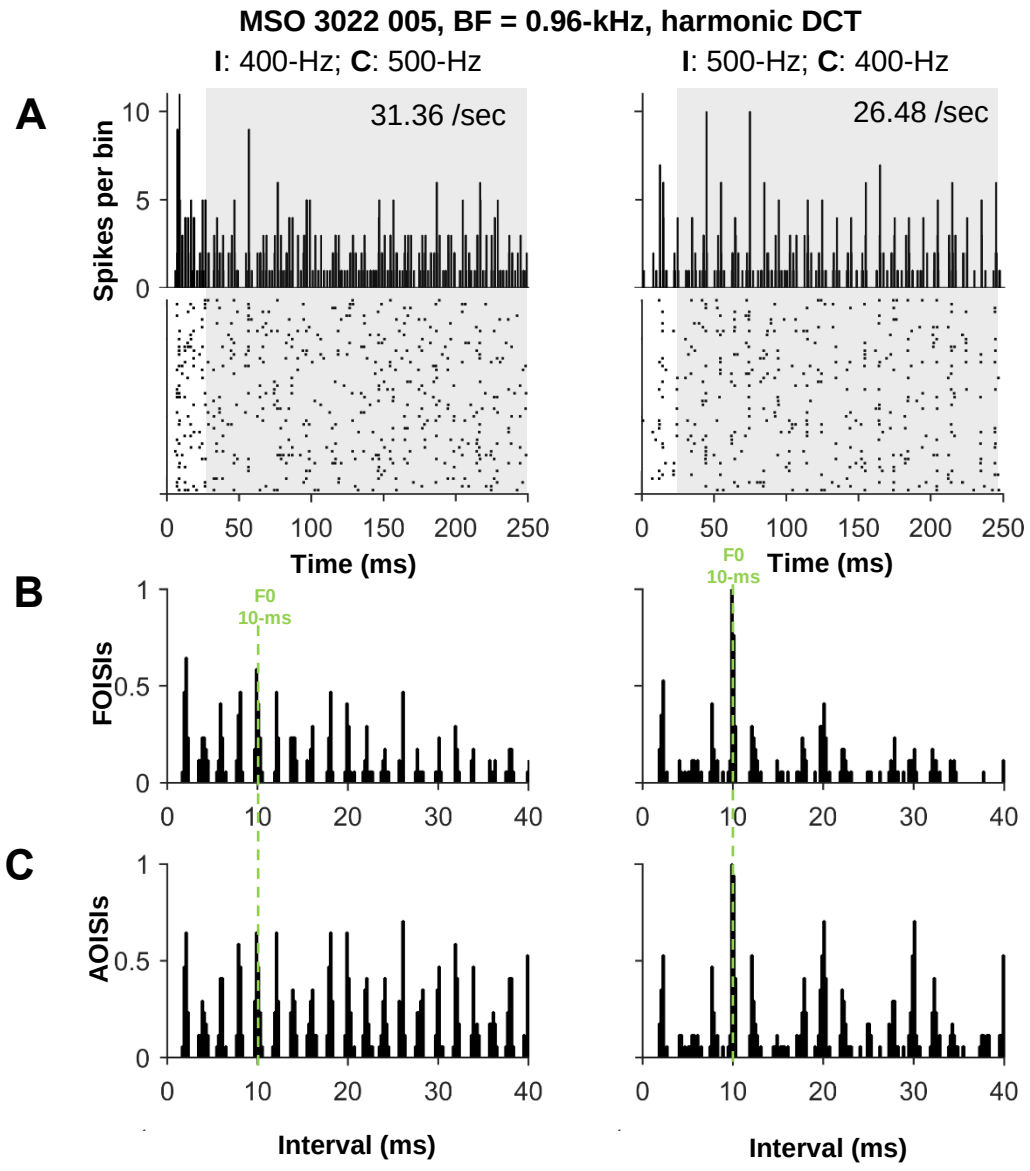

Figure 44 **ISI distributions of MSO unit responses to harmonic DCTs show a peak at the F0-period, 10-ms (shown as the dotted line)**. Stimulus was a DCT consisting of the 4th and 5th harmonics of F0 = 100-Hz, presented in both the ipsi-lower and ipsi-higher configurations. Stimulus level was 30-dB SL. **I** = ipsilateral; **C** = contralateral. Spike rates are inset in the PSTHs. Unit spontaneous rate (SR): 6.2 /sec.

A: PSTHs and raster plots. The ISI analysis window is shown in grey.

B: Normalised FOISI distributions.

C: Normalised AOISI distributions.

Figure 45 (page 112) shows responses of the same unit as **Figure 44** to a DCT shifted by 50-Hz, half the F0. The AOISIs of responses to both stimulus configurations (**Figure 45C**) show

multiple peaks, the three largest of which occur at ½×F0-period and flanking the F0-period. The 'true' F0 of this ambiguous stimulus is 50-Hz (450-Hz/550-Hz = 9th and 11th harmonics). In the AOISIs, the largest non-zero peak indeed occurs at the period of this F0 (20-ms). The three largest peaks occur at similar periods to the ambiguous pitch percepts human subjects report. These originate from pseudoperiods in the stimulus TFS (see **Section 1.3.1, page 13**).

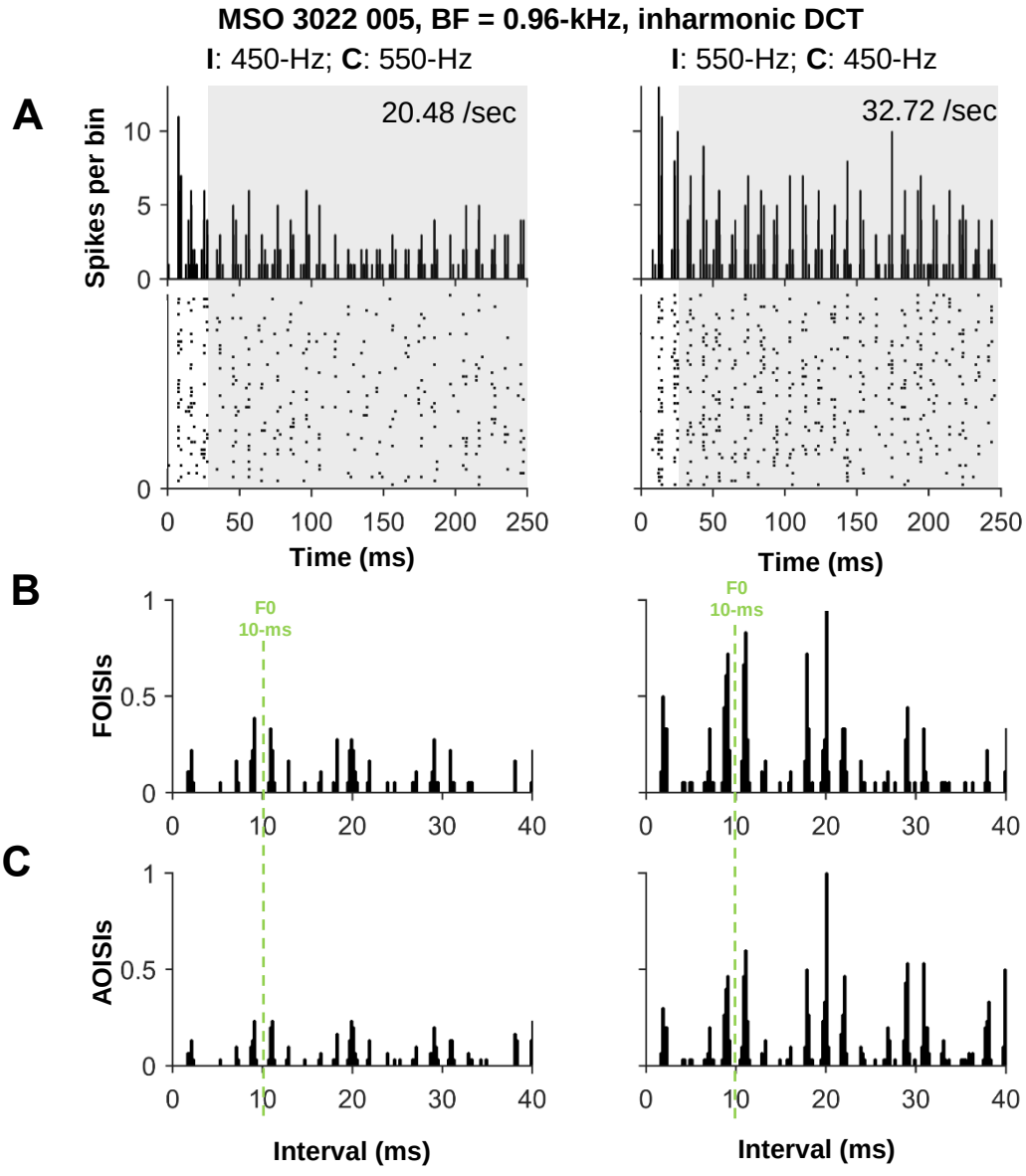

Figure 45 **Applying a frequency shift to the DCT causes a shift in the peaks of interval distributions of MSO responses.** Same unit as **Figure 44 (page 111)**, except that the complex has been frequency-shifted upwards by 50-Hz. Conventions as in **Figure 44 (page 111)**. Unit SR: 6.2 /sec.

DCTs were shifted in 9 steps between harmonic ranks. A TFS-driven change of MSO unit responses can be seen as frequency-shift is increased: the largest peak after the peak at zero lag ('non-zero peak') in SACs, the neural correlate of pitch, shifts to shorter periods. When plotted on polar coordinates (after Sayles & Winter, 2008b), this shift appears as an inward spiral.

Figure 46 (page 114) shows an exemplar MSO unit's responses to frequency-shifted DCTs. In **Figure 46A**, SACs of 5 of the 9 sampled conditions are shown for ipsi-lower presentation. As the DCT increases in frequency, the peaks in the SAC shift to shorter periods. After shifting by F0-Hz (in this case, 100-Hz), the unit again responds to the ENV period of 10-ms: this is a full harmonic rank increase compared to the initial condition, and a harmonic stimulus. **Figure 46B** shows a continuous plot of the 9 SACs, created by filtering the 9 SACs with a Gaussian kernel. The 10-ms ENV period is shown as a white dotted line. The sloped lines flanking the 10-ms dotted line represent a neural correlate of the first effect of pitch shift to DCTs. In **Figure 46C**, the data are transformed onto polar axes; in this representation, the first effect of pitch shift appears as an inward spiral. **Figure 46D** shows the swapped stimulus configuration, where the lower harmonic of the pair is presented to the ipsilateral ear. A similar response profile can be seen to **Figure 46C**, with a similar response to the shifting F0-period. However, the temporal precision of the responses are generally weaker, seen as a difference in colour scale between the plots.

The first effect of pitch shift was seen in all MSO unit responses to frequency-shifted DCTs, regardless of harmonic rank used. In **Figure 47 (page 115)**, responses of another MSO unit to the 2nd and 3rd harmonics of F0 = 100-Hz are shown. This unit also shows a shift in the neural correlate of pitch to shorter periods with increasing DCT frequency shift, and a similar profile across stimulus configurations.

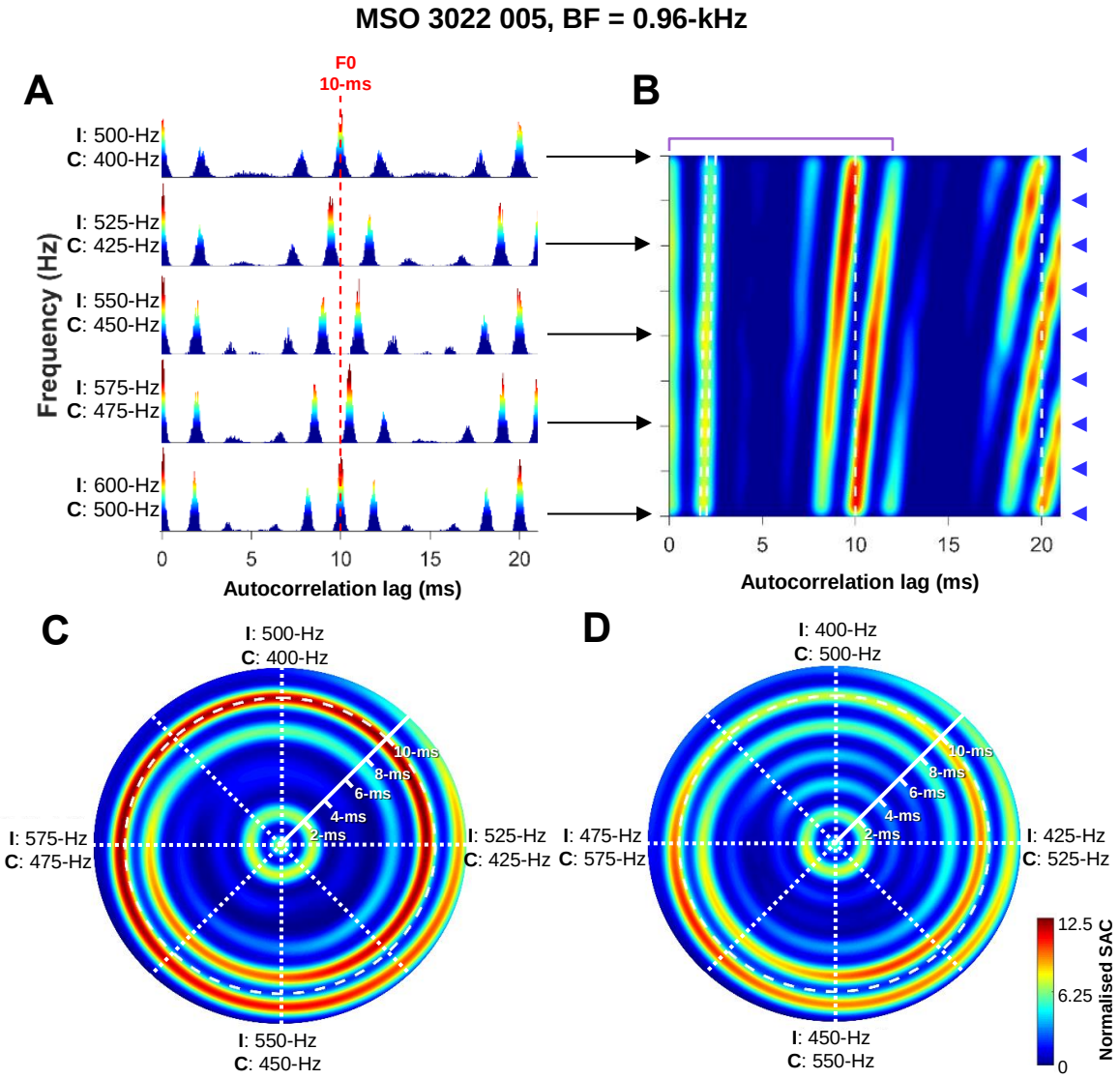

Figure 46 **MSO unit showing a correlate of the first-effect of pitch shift to frequency-shifted DCTs**.
Same unit as **Figure 44 (page 111)** and **Figure 45 (page 112)**. **I** = ipsilateral; **C** = contralateral.
A: SACs of 5 of the 9 sampled conditions are shown for ipsi-lower presentation. 'Normalised coincidences' are represented as a colour scale. The ENV period of 10-ms is shown as a red dotted line.
B: Continuous plot of the 9 SACs, filtered with a Gaussian kernel. Blue arrowheads indicate the sampled frequency pairs. The 10-ms period is shown as a white dotted line. The purple brace denotes the timespan plotted in **C**.
C: Polar plot of B. The period of the 100-Hz spacing is shown as a dotted circle.
D: A polar plot of the ipsilateral-lower presentation.

Across the population of MSO units (*n* = 20), the neural correlate of pitch to DCTs shifts with increasing frequency shift (see **Figure 49A, page 116**). This is in line with the first effect of

pitch shift described both in human psychophysics and in monaural physiological data from the VCN (see **Section 5.3.5, page 156**).

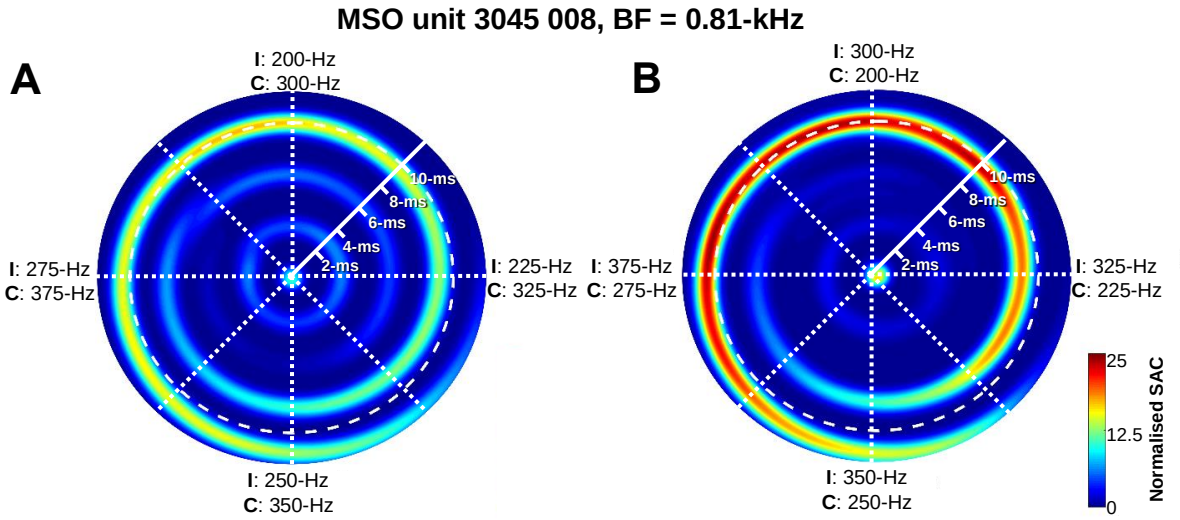

Figure 47 **MSO unit responses to DCTs with different harmonic ranks also show a correlate of the first effect of pitch shift**. Tones were presented at 30-dB SL. **A** and **B** show the ipsi-lower and ipsi-higher presentations of the DCT; the same pitch shift is seen in both cases. Conventions as **Figure 46 (page 114)**. Unit synchrony to pure tones and receptive fields are shown in **Figure 40 (page 92)**.

LSO units (n = 2) showed stronger evidence of a response to individual tone frequency in their raw responses, namely the ipsilateral carrier, which is expected given their response profile of [E·I] (see **Section 3.3.4, page 83**). However, they also responded to the overall F0 across ears, showing the first-effect of pitch shift in response to frequency-shifted DCTs. In **Figure 48 (page 116)**, I show DCT responses for an exemplar LSO unit. In **Figure 48A**, the initial condition shows a peak in the SAC at 5-ms, the period of the ipsilateral carrier (200-Hz). In **Figure 48B**, the presentation is swapped, and the initial condition shows a peak at 3.33-ms, the period of the ipsilateral tone in this presentation (300-Hz). However, the neural correlate of

pitch occurs at the F0-period in both cases, and it shifts similarly to MSO units. **Figure 49B** below shows the averaged DCT response across the LSO population (*n* = 2).

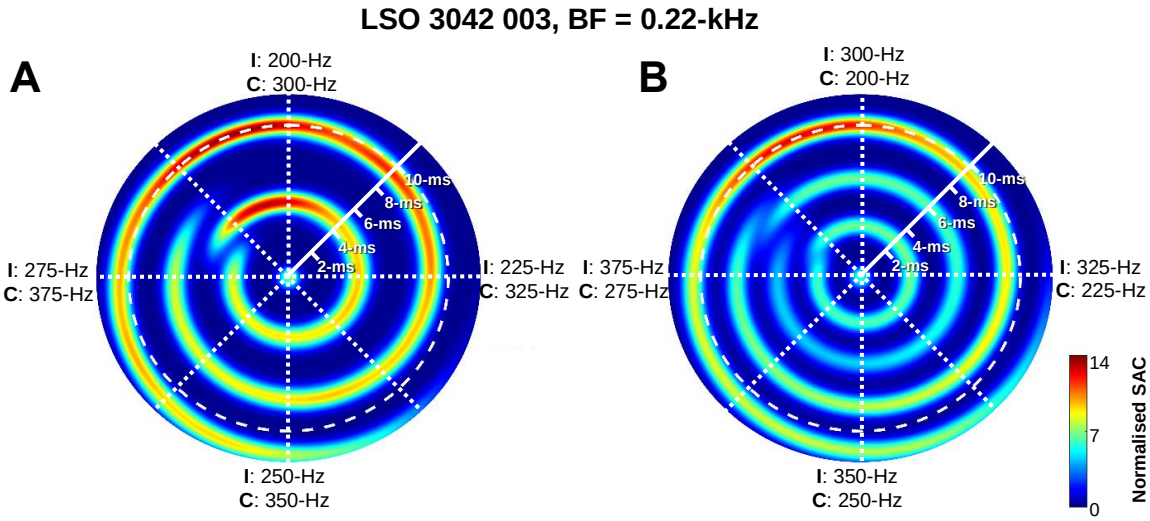

Figure 48 **LSO units also show the first effect of pitch shift, irrespective of harmonic rank and stimulus presentation**. Note the clear response to the ipsilateral carrier in both presentations. Conventions as **Figure 46 (page 114)**. Unit receptive fields are shown in **Figure 34 (page 85)**.

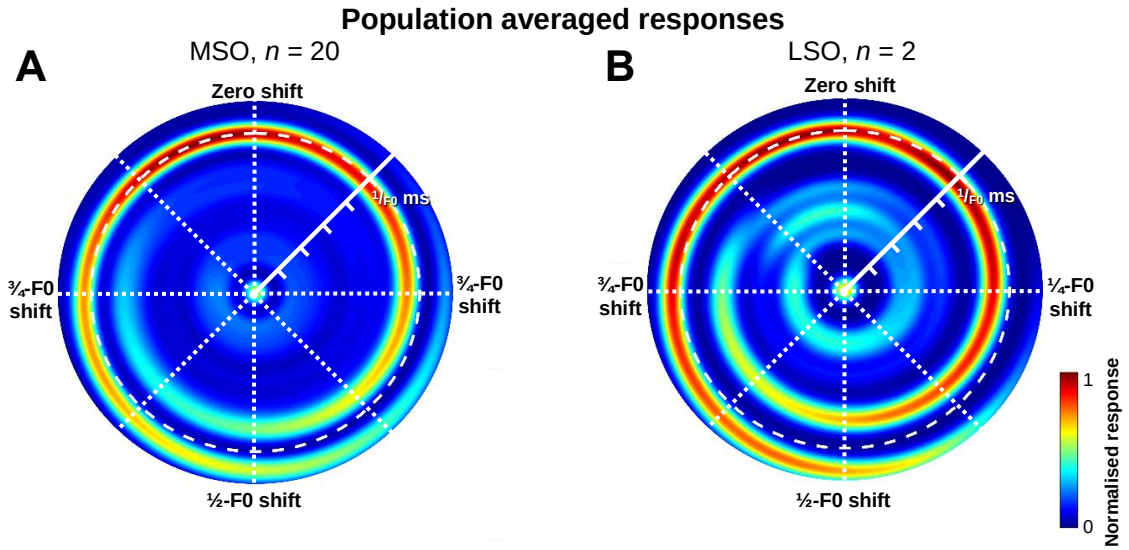

Figure 49 **Averaged SACs of responses to DCTs across all units, presentations, F0s and harmonic ranks for the MSO (**A**) and LSO (**B**) populations**. The first effect of pitch shift can be seen as an inward spiral in both the MSO and LSO. Conventions as **Figure 46 (page 114)**.

4.3.2 Dichotic harmonic tone complexes

In this section, I show MSO and LSO unit responses to dichotic HTCs similar to those used previously in human psychophysics (see **Section 4.1.2, page 104**) and neurophysiological studies of the guinea pig CNIC (see **Section 4.1.3, page 105**). These stimuli, contrary to the DCTs, are unresolved at low-F0s. In contrast to the finding of an equivalence of perception of the pitch of diotic and dichotic HTCs in human psychophysics, the neurophysiological studies found that cells of the CNIC responded differently in diotic and dichotic stimulation.

Figure 50 below shows example electrode recordings that contributed to the histogram in **Figure 51 (page 118)**. This unit tended to fire two spikes in close succession, termed 'doublets', and showed variable spike amplitude. The red crosses demarcate triggered spikes, and the average spike waveform is shown inset. Evidence of the neurophonic (see **Section 3.1.3.3, page 58**) can be seen as the sinusoid-like waveform centred at 0-mV.

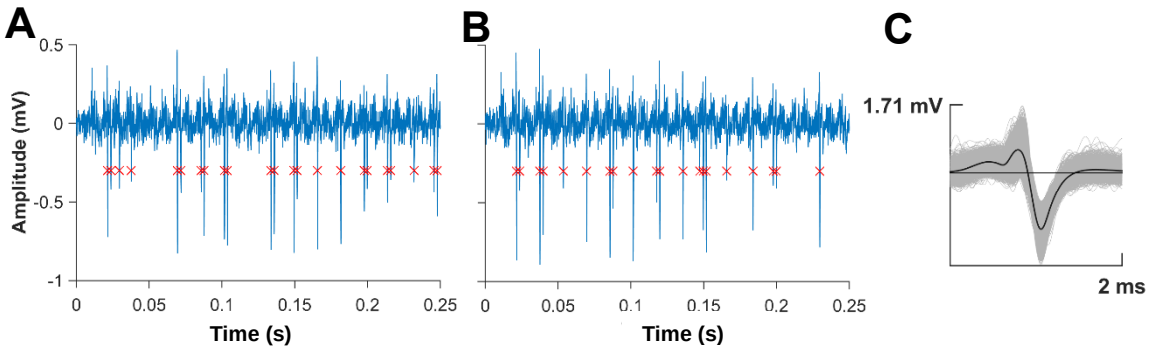

Figure 50 **Electrode recordings (**A **and** B**) of an MSO unit's responses to 62.5-Hz dichotic cpHTCs, and the average spike waveform (**C**) of all triggered waveforms**. This unit exhibited doublet spiking, however no events occurred in the refractory period and no second spike waveform shape is seen, indicating this was an isolated single unit. The action potential is upward-going and biphasic.

Figure 51 (page 118) shows an MSO unit's responses to 62.5-Hz HTCs. This unit shows a weak response to monaural-only conditions, expected from its weak monaural receptive fields

(see **Figure 36, page 87**). In dichotic presentation, the largest peak in the FOISI histogram (**Figure 51D**) occurs at ~2.5-ms, which corresponds to the harmonic (0.50-kHz) nearest unit BF (0.46-kHz). However, in the AOISI histogram (**Figure 51E**), the largest peak, the neural correlate of pitch, occurs at 16-ms, the F0 of the HTC across ears.

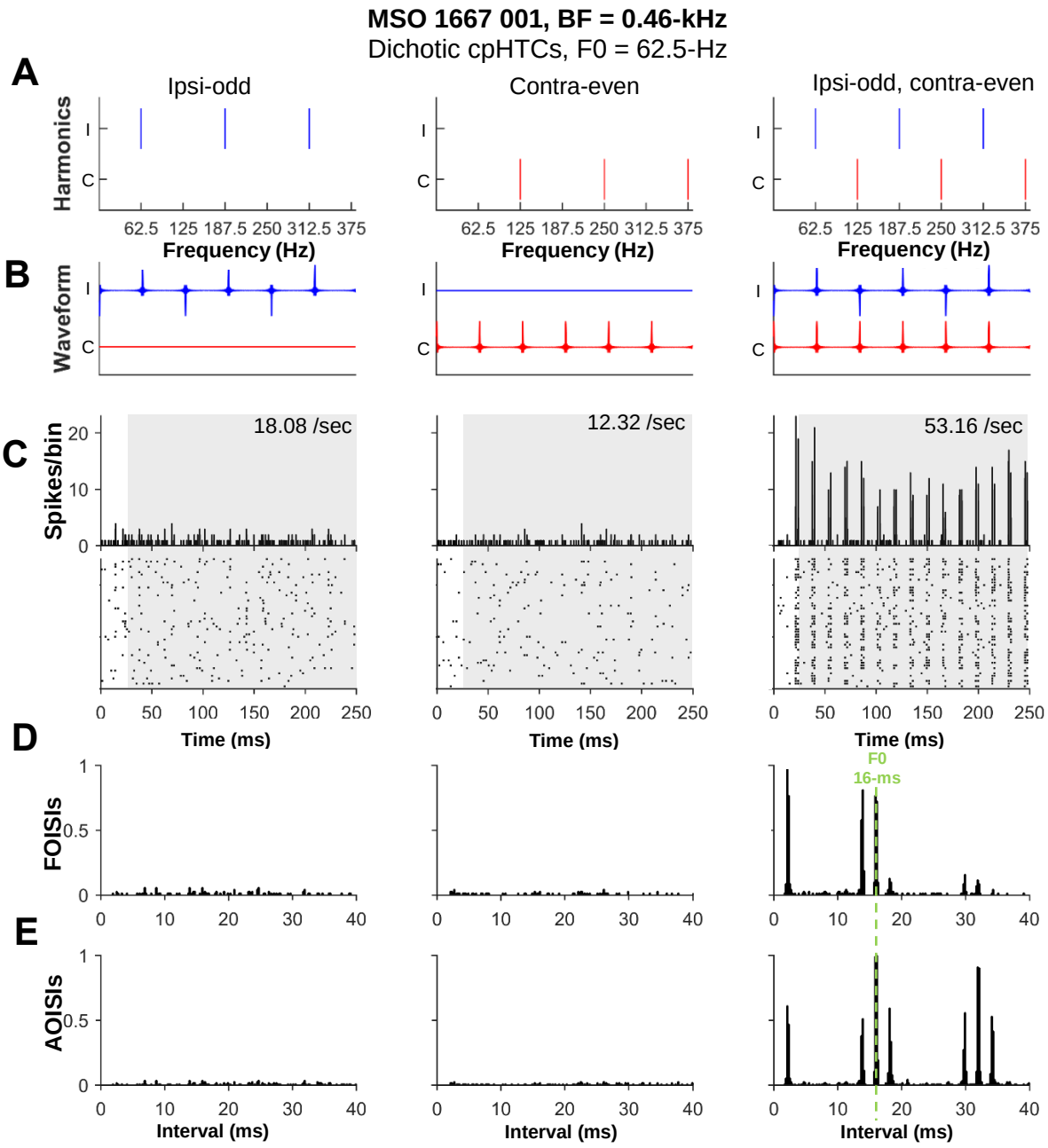

Figure 51 **MSO units respond weakly to monaural odd-only and even-only HTCs, but robustly to dichotic HTCs at the period of the overall F0**. In this case, this is 16-ms, marked by the dotted line. Conventions as in **Figure 44 (page 111)**; receptive fields for this unit are shown in **Figure 36 (page 87)**. Unit SR: 7.9 /sec.
A: schematic harmonic ladders; complexes extended to 5-kHz. **B**: stimulus waveforms. **C**: PSTHs; spike rates are inset. **D** and **E**: normalised FOISI and AOISI histograms of the grey analysis window in **C**.

Figure 52 (page 120) shows SACs of an MSO unit's responses to cpHTCs presented in both dichotic configurations: ipsi-odd, contra-even and ipsi-even, contra-odd. This unit showed weak responses and low spike rates to monaural stimulation. In odd-only presentation, MSO unit responses show three peaks in the SAC (**Figure 52C**): two surrounding the 2×F0 (red dotted line) period and one at the F0 (green dotted line). The periods at which these peaks occur correlate with the ambiguous pitches reported by human subjects. In even-only presentation, MSO units show responses at the 2×F0 period – a response to all the harmonics of double the F0. In dichotic presentations, responses are stronger and the SAC shows a large peak at the period of the overall F0 across ears.

MSO units show responses to the F0, irrespective of whether harmonics are presented diotically or dichotically, and irrespective of whether the complex was presented with odd-only harmonics to the ipsilateral ear or *vice versa*. In this way, MSO unit responses to dichotic cpHTCs resemble responses to diotic cpHTCs. In **Figure 53 (page 121)**, I have plotted an MSO unit's responses to dichotic and diotic presentations for cpHTCs of three F0s. The response profile is very similar across both dichotic and diotic presentation, across all F0s: the neural correlate of pitch is consistently to the F0 of the entire complex across both ears, echoing the human psychophysical data.

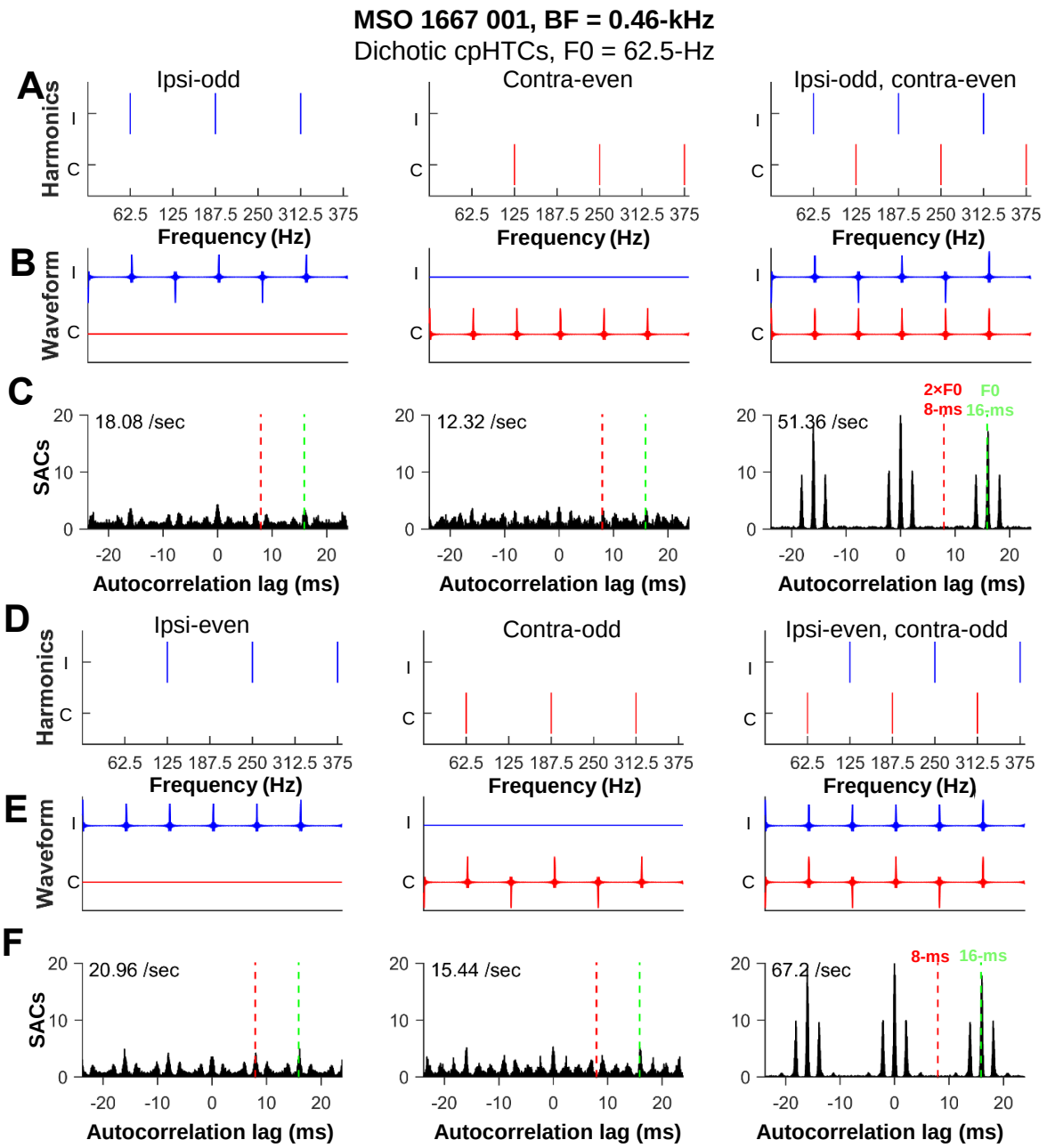

Figure 52 **MSO unit responding to the overall F0, 62.5-Hz, of dichotic cpHTCs, irrespective of stimulus configuration (ipsi-odd, contra-even or ipsi-even, contra odd)**. Red line: 2×F0-period. Green line: F0-period. Unit SR: 7.9 /sec.

A and **D**: schematic harmonic ladders. **B** and **E**: stimulus waveforms. **C** and **F**: SACs of unit responses, normalised according to Louage et al., 2004.

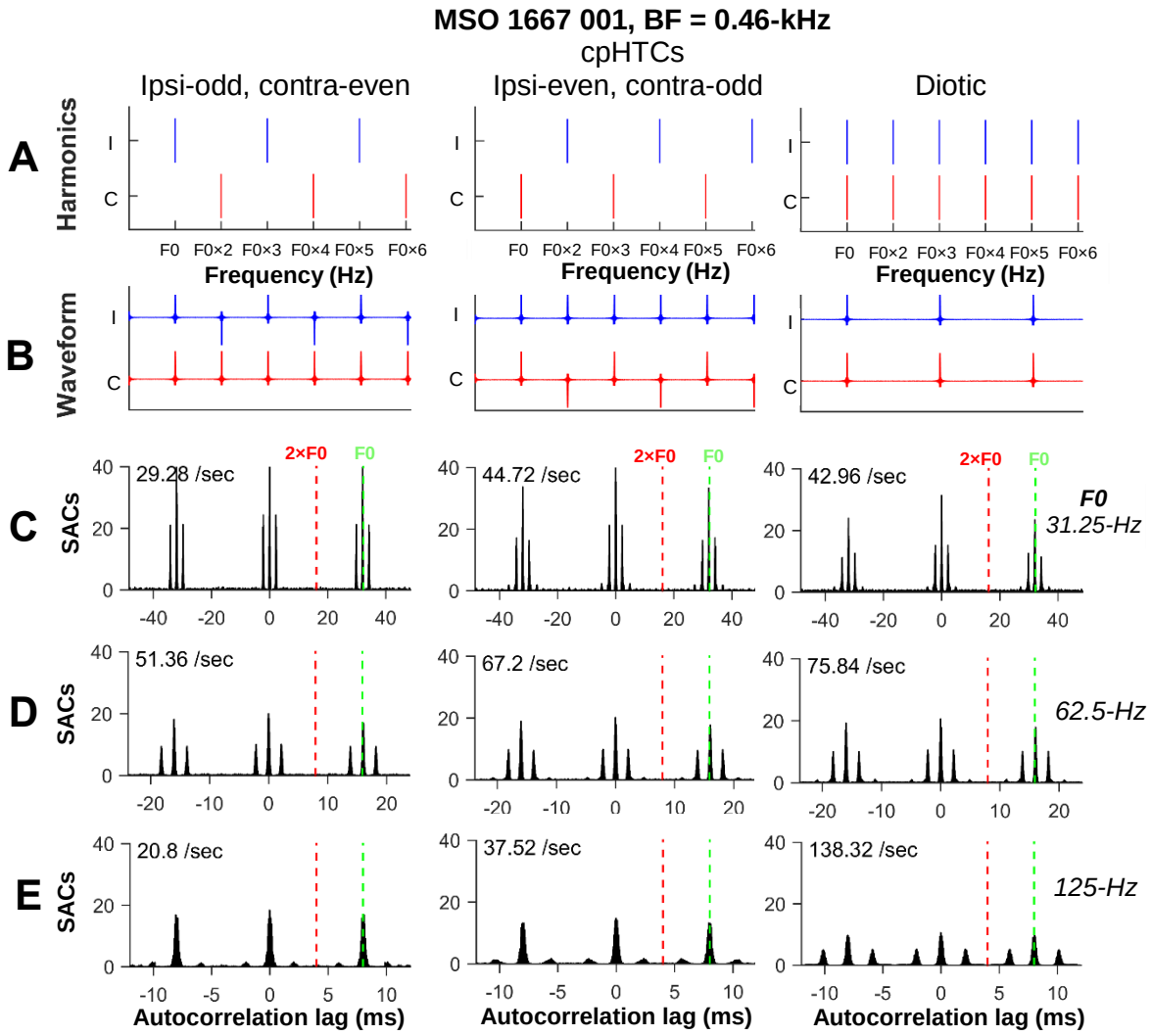

Figure 53 **MSO unit responses to dichotic stimulation are similar to diotic stimulation, however there are differences in evoked spike rates**. Across all F0s and presentations, the MSO unit responds to the F0-period across ears. Conventions as in ***Figure 52*** **(page 120);** same unit as **Figure 52 (page 120)**. *Note that the x-axis scale changes between* ***C****,* ***D*** *and* ***E***. Unit SR: 7.9 /sec.

In later experiments, responses of MSO units to HTCs with randomised phase were used to assess the impact of decreased crest factor. Additionally, the Bernstein & Oxenham, 2003 study used random-phase HTCs, so it was desirable to have data gathered in the most similar paradigm possible. Frozen-random-phase HTCs (frzHTCs) (*n* = 15) and random-phase HTCs

(rpHTCs) (*n* = 3) were used. Responses to these stimuli could show lower spike rates, however MSO unit responses still show a peak in the SAC or AOISI to the F0 across ears. Two MSO units did not respond at all to frzHTCs, even at the maximum level used. No functional difference was observed between frzHTCs and rpHTCs. ISI distributions of an exemplar MSO unit response to frzHTCs are shown in **Figure 54** below; ISI distributions of an exemplar response to rpHTCs can be seen in **Figure 55 (page 123)**.

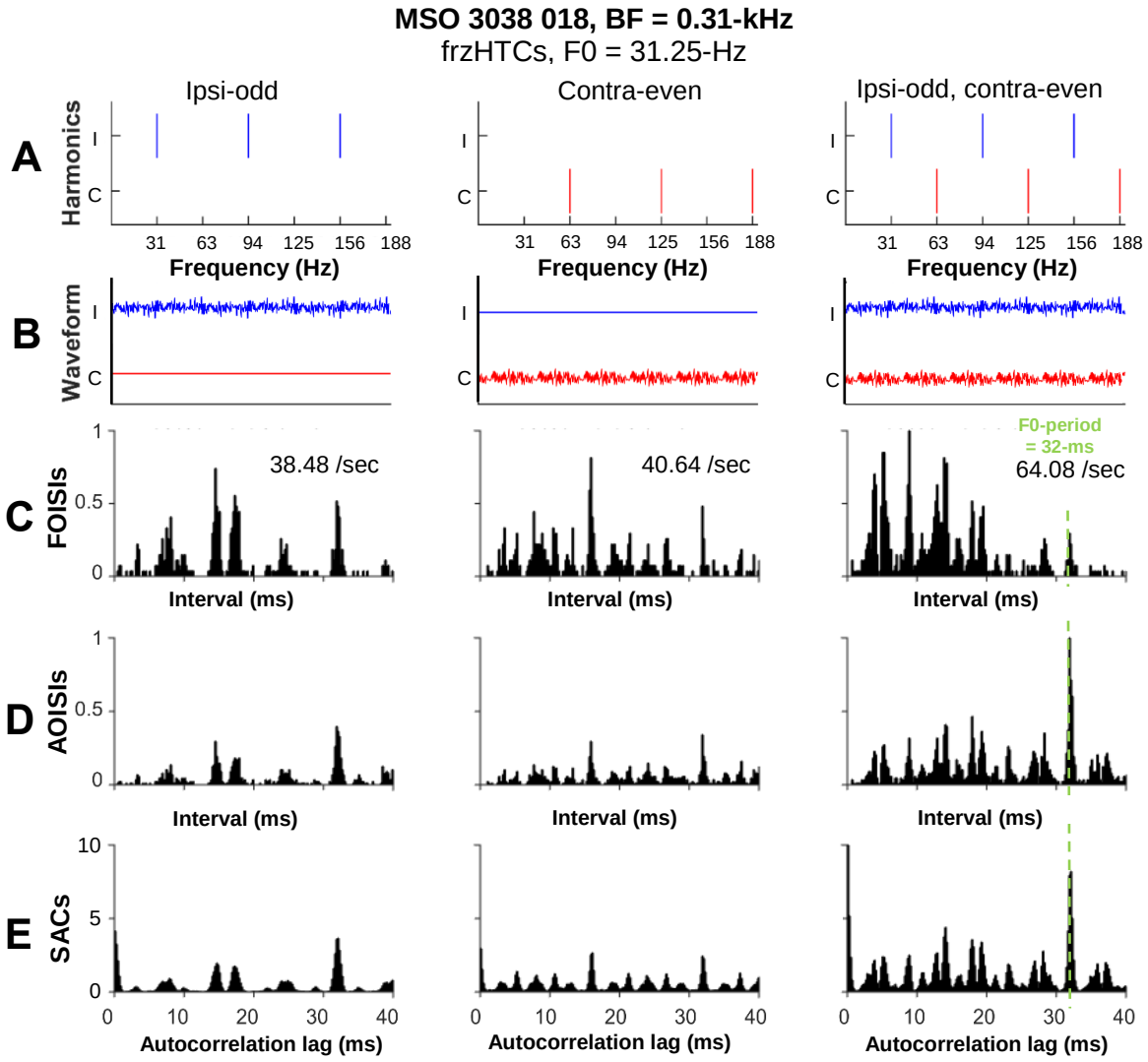

Figure 54 **In response to frzHTCs, MSO units still show a response to the period of the F0 across ears**. F0 = 31.25-Hz (period = 32-ms), shown at the green line. Unit SR: 11.7 /sec.

A: schematic harmonic ladders; note that these extended to 3-kHz. **B**: stimulus waveforms. **C** and **D**: FOISI and AOISI histograms, normalised to the greatest spikes/bin across configurations. **E**: SACs of unit responses, normalised according to Louage et al., 2004.

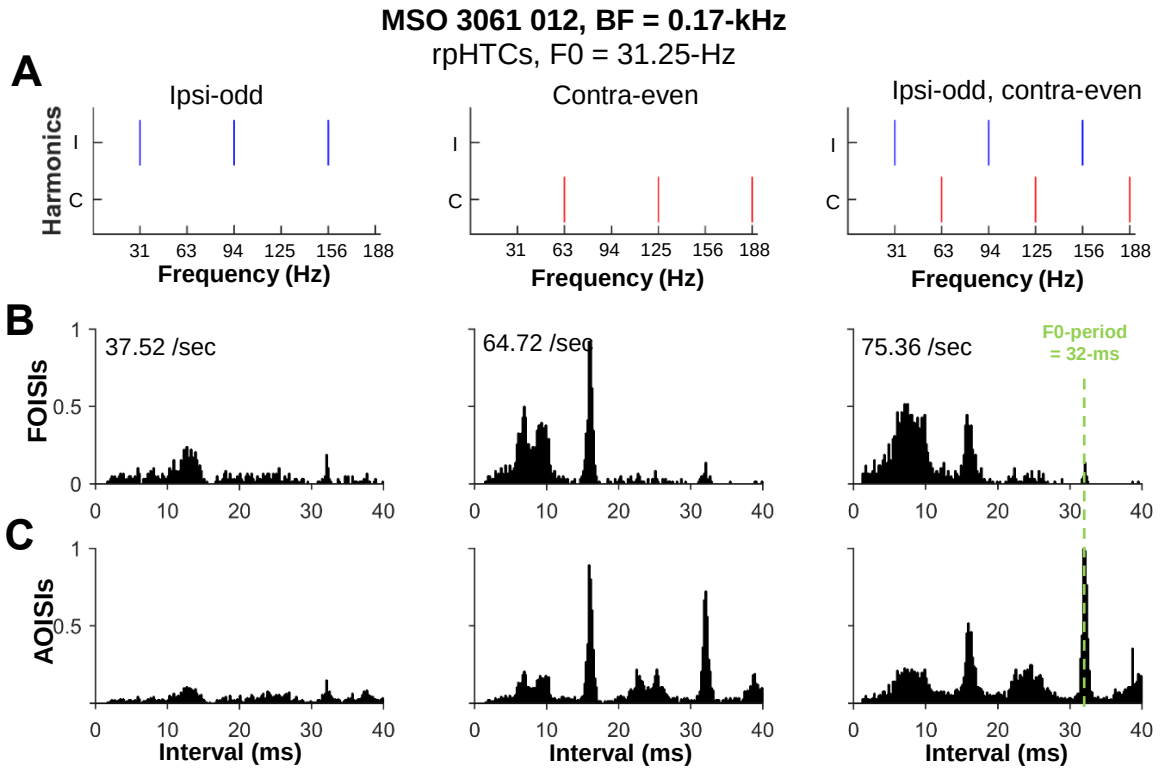

Figure 55 **MSO unit responses to rpHTCs resemble those to frzHTCs and cpHTCs, with the neural correlates of pitch occurring at the period of the overall F0 across ears**. Conventions as in **Figure 54 (page 122)**, however, no waveform has been shown as each presentation used a different waveform; waveforms would resemble those shown in **Figure 54 (page 122)**. Unit SR: 39.6 /sec.

Across the population of MSO units, across F0s and presentations, units responded to the overall F0 across ears, reflecting binaural integration of pitch. This is seen in **Figure 56 (page 124)**, the summed SACs of all MSO units to cpHTCs, and **Figure 57 (page 125)**, the summed SACs of MSO responses to frzHTCs. In dichotic presentation, the largest peak in the SAC occurs at the period of the overall F0 across ears, except at 500-Hz where the odd only and dichotic responses are almost identical. This is likely due to low stimulus energy in the MSO filters, with single low sound level components presented to each ear, compared to the many components at lower F0s.

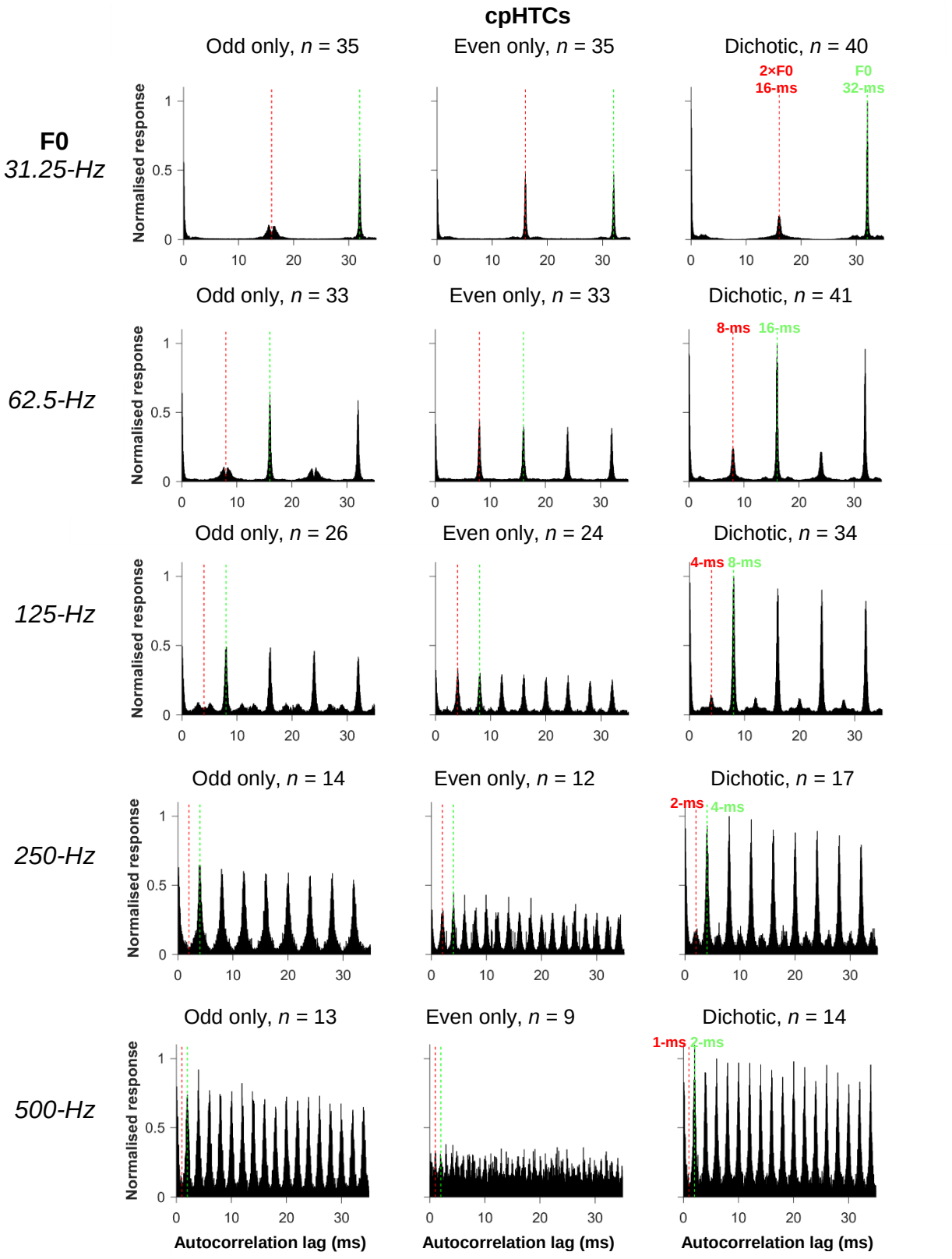

Figure 56 **Averaged SACs of MSO unit responses show clear neural correlates of pitch to the overall F0 across ears in response to dichotic cpHTCs**. Responses were combined across both dichotic configurations, and only rejected if spike rates were very low (<5 spikes/s). F0-period: green line; 2×F0-period: red line.

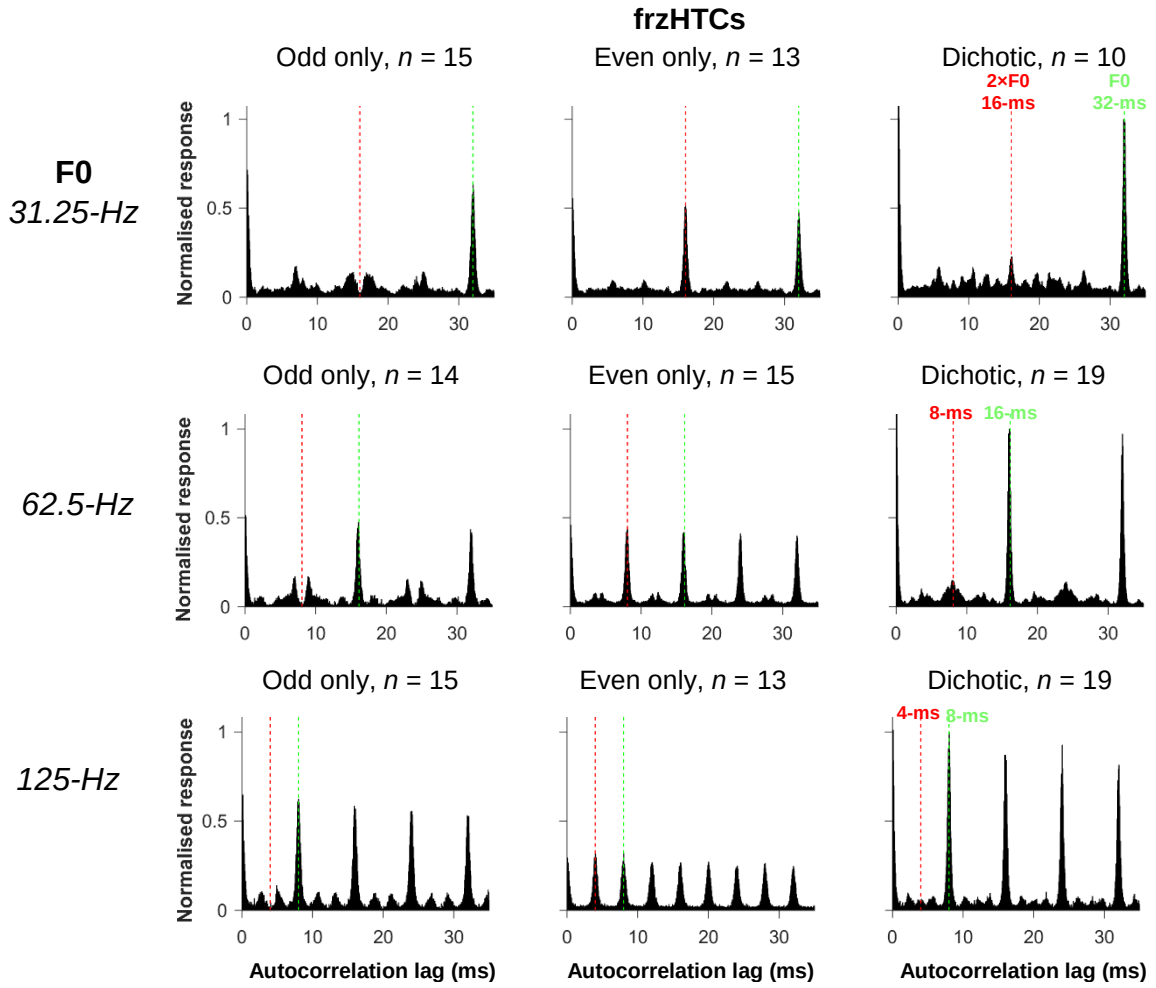

Figure 57 **Averaged SACs of MSO unit responses show clear neural correlates of pitch to the overall frzHTCs**. Conventions as **Figure 56 (page 124)**.

Units of the LSO show similar responses to the MSO, however the spike rates responses to dichotic stimulation are weaker than ipsilateral-only stimulation. In response to dichotic cpHTCs, LSO units also respond to the F0 across ears, independent of presentation. See **Figure 58 (page 126)** for ipsilateral-odd presentation and **Figure 59 (page 127)** for ipsilateral-even presentation.

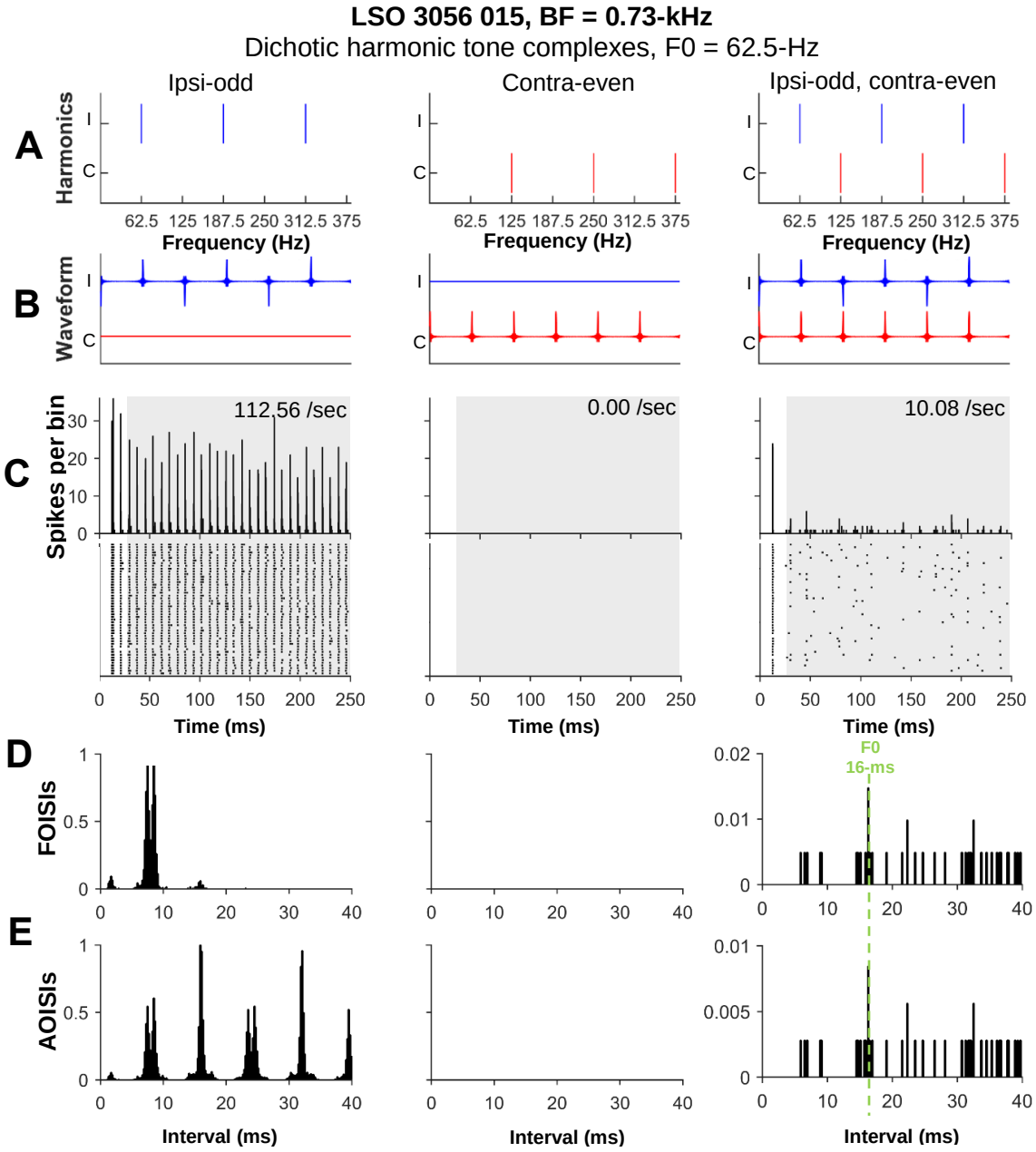

Figure 58 **LSO responses to dichotic cpHTCs show evidence of binaural integration of pitch**. However, strong contralaterally-derived inhibition decreases the response magnitude greatly compared to the strong responses of ipsilateral-only stimulation. Conventions as in **Figure 44 (page 111)**, but note that the normalised ISI distribution y-axes have been scaled up in the dichotic configuration. Receptive fields for this unit can be found in **Figure 39 (page 90)**. Unit SR = 0.3 /sec.

Despite the asymmetry in monaural responses of LSO units (see **Section 3.3.4, page 83**), and the fact that stimulation rates were considerably lower in dichotic than in ipsilateral-only presentations, LSO units also responded to the overall F0 across ears.

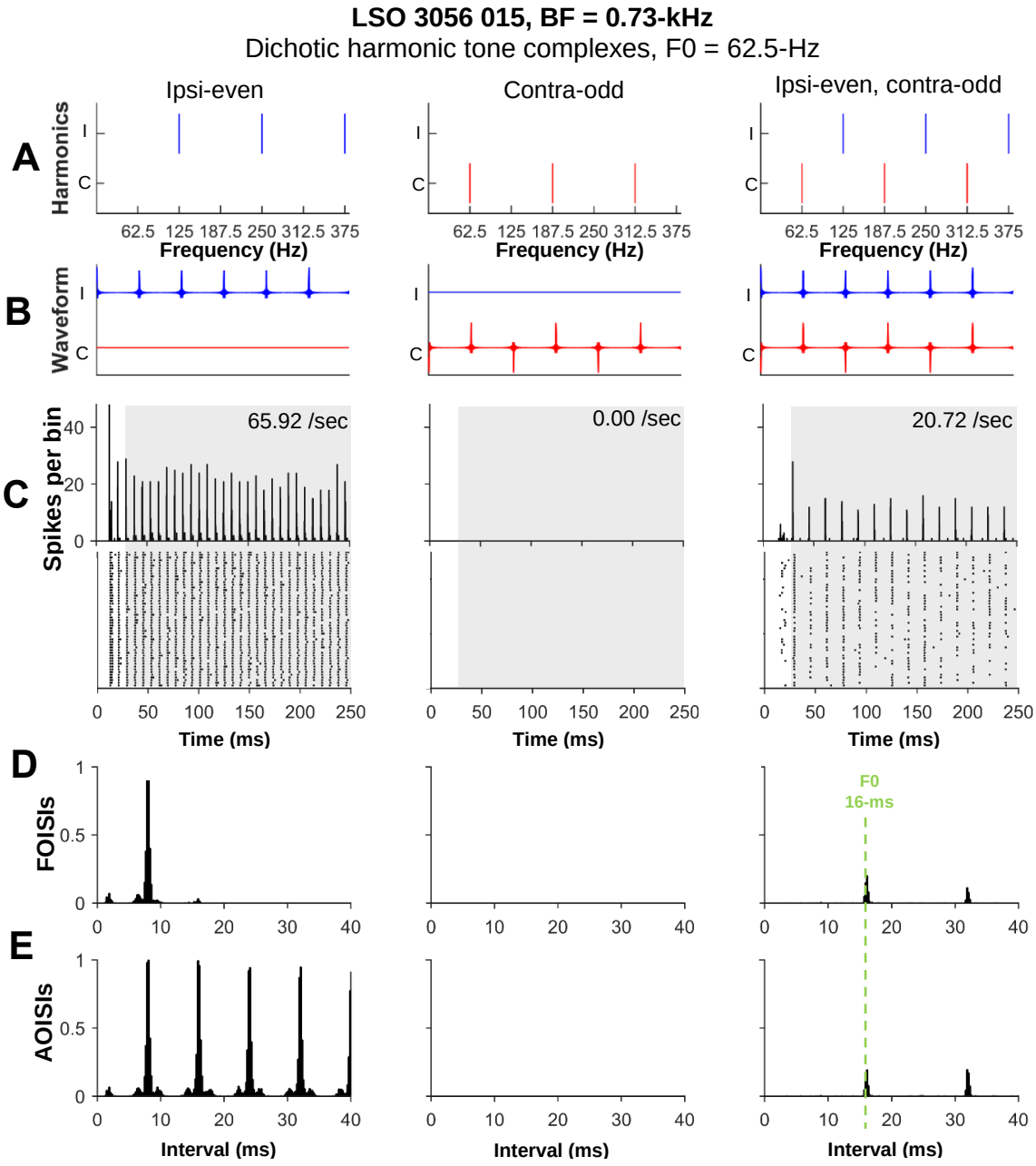

Figure 59 **When the dichotic presentation is reversed is swapped, the same LSO unit as Figure 58 (page 126) shows stronger evidence of binaural integration of pitch**. As previously, spike rates are lower in the dichotic configuration compared to ipsilateral-only stimulation. Conventions as in **Figure 44 (page 111)**. Unit SR = 0.3 /sec.

Averaged SACs of LSO responses are shown in **Figure 60 (page 128)**. The peak in the SAC, the neural correlate of pitch, of responses to dichotic stimulation occurs at the F0-period in all cases.

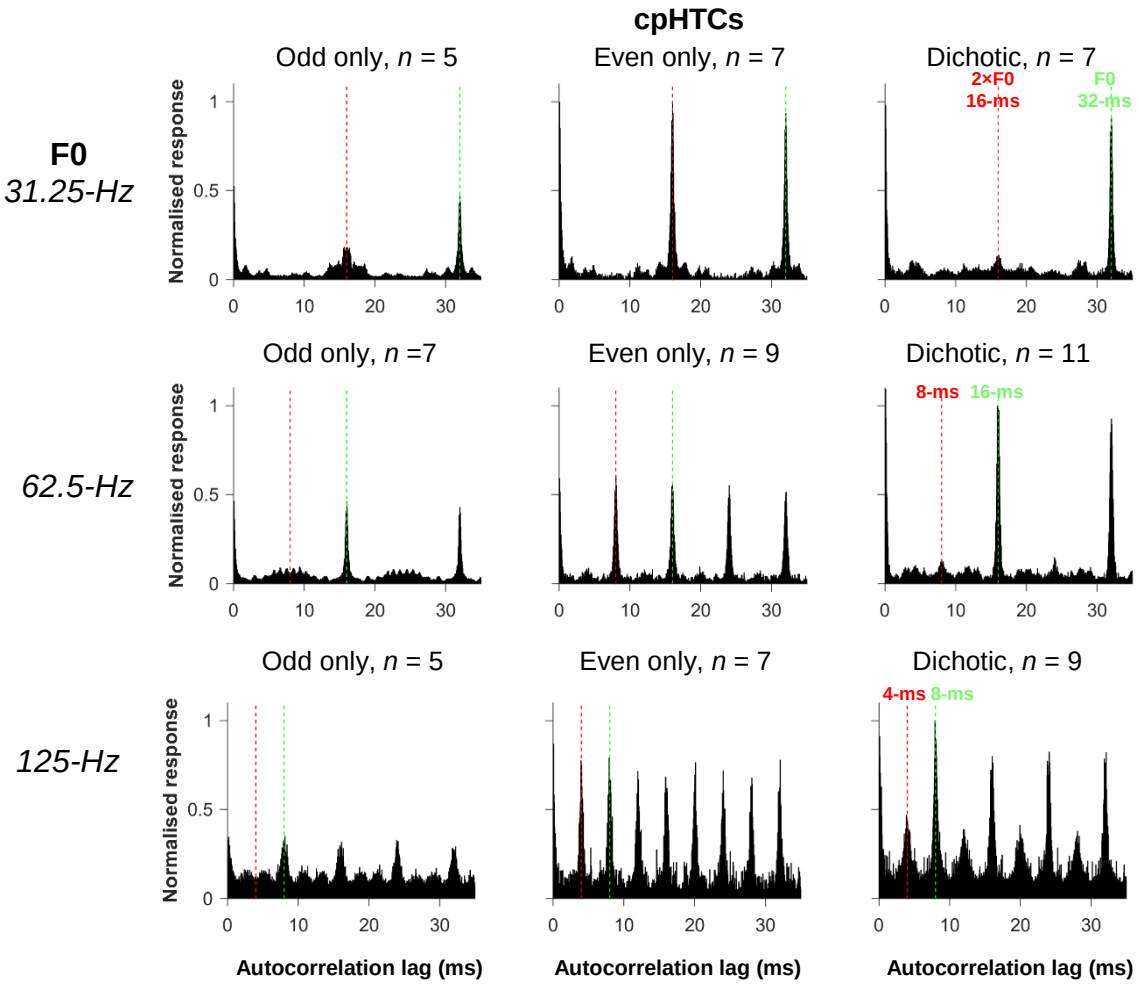

Figure 60 **Averaged SACs of LSO unit responses to dichotic cpHTC show clear binaural integration of pitch**. In dichotic presentation, the largest peak in the averaged SAC is to the overall F0 across ears. Conventions as **Figure 56 (page 124)**. Averaged SACs of MSO unit responses can be found in **Figure 56 (page 124)**.

The normalised period at which peaks occur in the SACs of MSO and LSO unit responses to cpHTCs is shown in **Figure 61 (page 129)** as a function of BF and F0. Each column of points corresponds to the largest peaks in the SAC of one unit's responses to one stimulus configuration of one F0. When normalised in this way, the peaks surrounding the 2×F0 period in odd-only presentation can be seen to vary in a consistent manner, fitted by lines indicating predicted autocorrelation peaks using the equation: $0.5 \pm \frac{\mathrm{BF}}{2\times\mathrm{F0}}$, where BF and F0 are in kHz (Bilsen & Ritsma, 1969; Sayles & Winter, 2007, 2008b). In dichotic presentation, the largest peak in the SAC is at the overall F0 across ears. Compare this to the responses from VCN

(see **Figure 70, page 155**) where the dichotic presentation elicits near-identical responses to the ipsilateral-only presentation and there is no binaural integration of pitch.

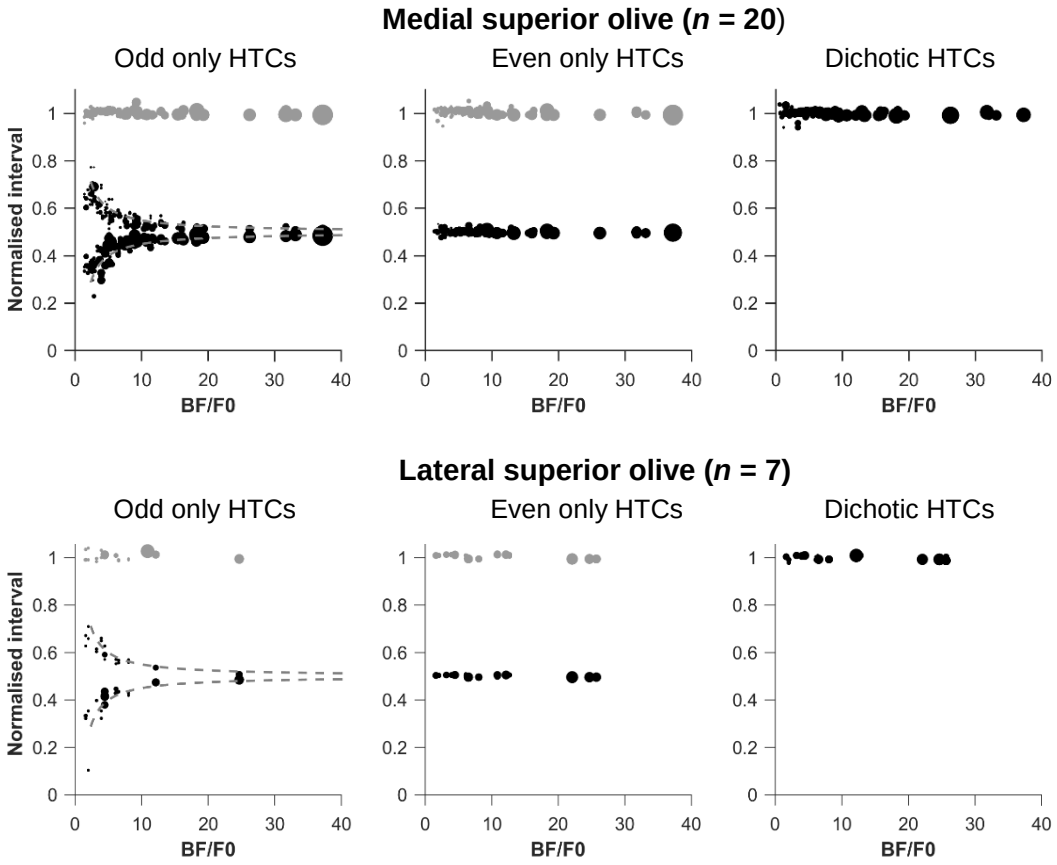

Figure 61 **Summary of cpHTC responses for MSO and LSO, showing binaural integration of pitch**. SAC peaks around 2×F0-period and F0-period have been taken from each SAC for each condition in each presentation across all units, then replotted as normalised interval (F0 scaled to 1). The size of the scatter point is proportional to the normalised coincidence value. Intervals are plotted against BF/F0.

Ipsi-odd: in each case, the three largest peaks occur at $\mathbf{0.5} \pm \frac{BF}{2\times F0}$ and 1.

Contra-even: peaks occur at 0.5 (2×F0-period) and 1 (F0).

Dichotic: peaks occur at 1 (F0-period) in all cases.

4.4 DISCUSSION

4.4.1 Comparison with psychophysical literature

In response to cpHTCs, frzHTCs and rpHTCs, all sampled units in the low-BF (<2-kHz) population of MSO and LSO showed a neural correlate of pitch at the period of the F0 across

ears. This is in agreement with the human psychophysics for dichotic HTCs (Houtsma & Goldstein, 1971, 1972; Bernstein & Oxenham, 2003),

A correlate of the first-effect of pitch shift in response to DCTs was found in all MSO and LSO units sampled, in agreement with the human psychophysical data (Houtsma & Goldstein, 1971). In **Figure 62 (page 131)**, I have overlaid the averaged MSO data from **Figure 49 (page 116)**, in black, onto the original psychophysical data; the data fit between the 3rd and 4th harmonic datasets and are well-fitted by a line plotted according to de Boer's rule (see **Section 1.3.1, page 13**). In other words, the MSO response average is dominated by the 3rd and 4th harmonics, in line with psychophysical data showing that, for F0s up to around 400-Hz, a spectral region spanning harmonics 3, 4 and 5 appears to dominate pitch perception across a variety of pitch-evoking stimuli (Ritsma, 1967; Bilsen & Ritsma, 1969; Ritsma & Bilsen, 1970). Additionally, the pitch shift obtained in psychophysics to monaural frequency-shifted HTCs shows an influence of component spacing (*i.e.* ENV), but only when the lowest harmonic is ranked 4 or lower (Patterson & Wightman, 1976).

The data presented in this Chapter provide possible evidence that the MSO, and, to some extent, the LSO, are the origins of the binaurally-fused inputs that feed into the central pitch processor (see **Section 1.4.3.2, page 28**). The same code, the neural correlate of pitch, seen in AN and CN responses to monaural pitch-evoking stimuli (see **Section 1.3.3, page 19**) was found to correspond to binaural pitch-evoking stimuli in the MSO and LSO.

In **Figure 63 (page 131)**, single MSO unit responses to DCTs are compared with previously published CN unit responses to monaural frequency-shifted HTCs; in **Figure 75 (page 160)**, the population average MSO unit responses are compared with population average VCN unit responses; both show the same F0-period shift correlating with the first effect of pitch shift.

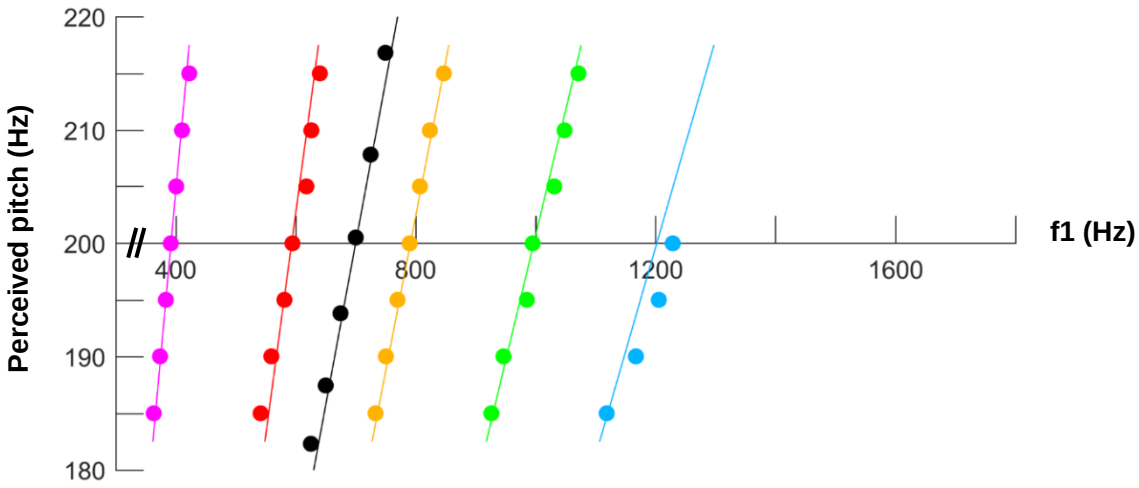

Figure 62 **Averaged data from the MSO, in black, overlaid onto human psychophysical data from Houtsma & Goldstein, 1971**. Data were excerpted from **Figure 49 (page 116)**, then scaled such that F0 = 200-Hz. Lines are projected first-effect of pitch shift estimates from de Boer's rule. MSO data are fitted closely by a line with harmonic rank 3.5 (black line). f1 – the frequency of the first tone in a given complex; the second is (f1 + F0)-Hz. Each colour represents the data gathered for a different harmonic rank.

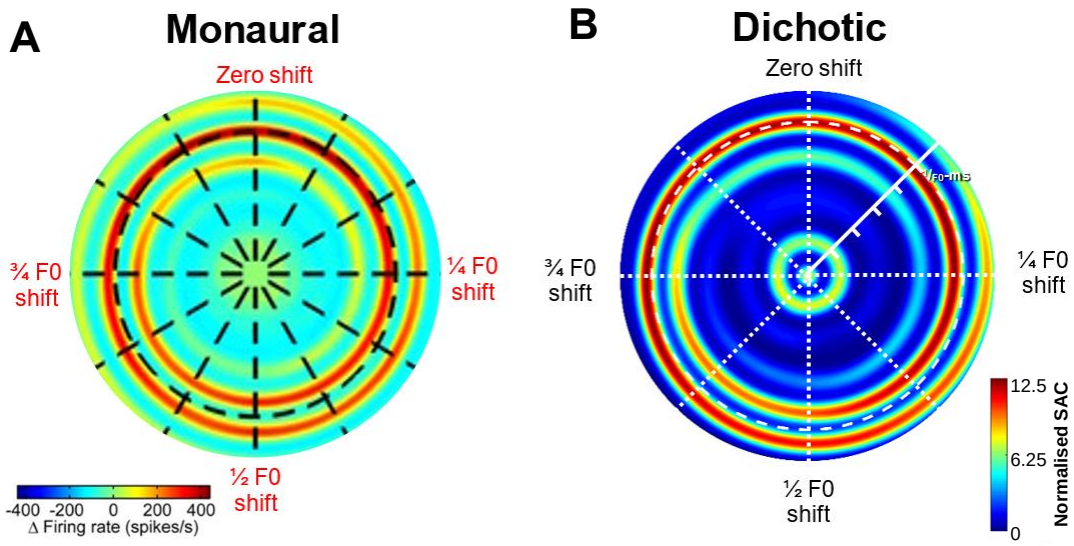

Figure 63 **Both PL units in guinea pig VCN and MSO units of the guinea pig SOC show the first-effect of pitch shift to frequency-shifted complex tones**. Note that the data from VCN were analysed using AOISIs and on the right with SACs, however the neural correlate of pitch is clear in both cases.

A: VCN PL unit; stimulus was an F0 = 500-Hz cpHTC; dotted circle shows the ENV-period of 2-ms. Excerpted from Sayles & Winter, 2008b.

B: MSO unit **3022 005**; stimulus was a DCT, F0 = 100-Hz, presented **I**: 500-Hz, **C**: 400-Hz, and then frequency shifted; dotted white circle is at the ENV period of 10-ms.

4.4.2 Comparison with previous binaural neurophysiological studies

MSO unit responses to dichotic and diotic HTCs were very similar. There was no dominance of one ear over the other in terms of the neural correlate of pitch, however the spike rates did vary. In the LSO sample, there were very low rates to dichotic stimulation, however the neural correlate of pitch was always to the overall F0 across ears.

In contrast, previous neurophysiological studies investigating binaural integration of pitch using HTCs found that CNIC cells acted to separate information between ears. In one study, HTCs of 125-Hz and 145-Hz were presented one to each ear. The majority of CNIC units responded preferentially at the F0-period of a HTC in one ear instead of the periodicity across ears (guinea pig – Nakamoto *et al.*, 2014). However, this study only used CNIC units with BFs greater than 2-kHz, past the phase-locking limit of the guinea pig IC of just over 1-kHz (Liu *et al.*, 2006). Given that responses to dichotic cpHTCs in this Thesis were only seen at F0s where harmonics were at least partially-unresolved, binaural integration of pitch at the level of the MSO or IC may require unresolved harmonics, with cells phase-locking to modulated TFS resulting from interactions of multiple harmonics in a filter.

This explanation, while reasonable, does not account for the lack of binaural integration of pitch in Shackleton *et al.*, 2009. This study used sine-phase dichotic HTCs, similar to the stimuli used in this Chapter. One third of CNIC clusters had BFs<1-kHz, and additionally, 98% of clusters phase-locked significantly to a diotically-presented F0 of 50-Hz, with 40% still phase-locking significantly to a diotically-presented F0 of 400-Hz, the highest they tested. However, in dichotic presentation, CNIC clusters responded instead at the 2×F0-period, with no differences seen in the responses to ipsilateral-odd, contralateral-even and ipsilateral-even, contralateral-odd configurations. Whether this study did not adequately sample the optimal subpopulation of CNIC cells (for instance, more low-BF units), or whether the experimental protocol they used, or indeed any other factor, may have impacted their non-finding of binaural

integration of pitch in the CNIC is not clear, and remains an open question. Interestingly, in the Nakamoto *et al.*, 2014 study, [E·E] cells were found in the CNIC that responded to the F0 difference across ears, 20-Hz. These cells might respond to dichotic HTCs similarly to the MSO cells shown in this Thesis.

To date, no other study has examined responses of binaural neurons to dichotic complex tones. In **Section 5.3.5 (page 156)**, I present data of monaural two-tone complexes recorded from low-frequency units of the VCN.

4.4.3 Binaural integration of pitch in the LSO

Given that the responses of LSO cells show binaural suppression in response to diotic stimulation, it may seem odd that low-BF LSO units respond, albeit with weak spike rates, to the periodicity of binaural complex tones, *i.e.* the overall F0. Low-BF LSO units show trough-type responses (see **Section 3.3.4, page 83**), with maximal response to binaural beat stimulation at close to 0.5 cycles and minimal response at close to 0 cycles. This reveals their action as 'anticoincidence detectors': the contralaterally-derived MNTB inhibition is strongly phase-locked (see **Section 3.3.6, page 91**).

For dichotic cpHTCs, in ipsi-even, contra-odd presentation, one could reason that every other EPSP evoked by the 2×F0 complex in the ipsilateral ear would be cancelled out by contralaterally-evoked IPSPs mostly occurring at the F0-period. This is indeed what was observed (see **Figure 59, page 127**). However, a weak response to the F0-period is still seen upon swapping the configuration of the cpHTCs. IPSPs occurring at the 2×F0-period do not seem to completely cancel the EPSPs occurring mostly at the F0-period (see **Figure 58, page 126**). Normalising the SACs, which accounts for the low spike rate, shows that these 'residual'

responses have reasonable temporal precision, as seen in the population SACs (see **Figure 60, page 128**) and DCT responses (see **Figure 48, page 116**).

The responses could be accounted for by BF-disparities in the inhibitory and excitatory filters. It appears that some LSO units in this Thesis show an incomplete overlap of ipsilateral (excitatory) and contralateral (inhibitory) filters; additionally, some low-BF LSO units showed residual, weak responses to diotic tonal stimulation (see **Figure 34, page 85**). Alternately, some LSO units may show binaural facilitation (see **Section 3.1.4.2, page 62**), although no definitive evidence of this was seen in the LSO population of this Thesis.

4.4.4 Responses to 'high' F0s

In this study, binaural integration of pitch was not seen in individual units for F0 = 250-Hz, or over the population for higher F0s. However, binaural pitch perception occurs at F0s above this. It is possible that MSO units do not binaurally-integrate pitch at these F0s, however I will consider some alternate explanations below.

The dataset gathered was from MSO and LSO units with BFs<2-kHz. It is likely that the responses shown in this Thesis arise from the interactions of unresolved harmonics within filters. However, at 'high' F0s (250-Hz and above), harmonics are fully resolved at the periphery. Therefore, sampling higher-BF units of the MSO and LSO might have led to the finding of binaural integration of pitch to these F0s.

The levels at which stimuli were presented were quite low to minimise the possibility of acoustic crosstalk, with levels per component generally not exceeding 30 dB-SPL for the HTCs. As thresholds at unit BF of MSO units are higher than this (mean±sd: 32 ± 13 dB-SPL), and most harmonics are not close to unit BF, it is possible that there was not enough energy in the filter

to induce a response at some F0s. Subsequent studies could investigate this by presenting high-F0 HTCs at higher sound levels.

It is possible the MSO may not be able to respond to unresolved harmonics in a phase-locked manner for 'high'-F0 complexes – this would require high BFs, which would not phase-lock as strongly as low-BF MSO units. Additionally, high BF MSO units are less prevalent: the population BF distribution of the MSO has a low-frequency bias. However, the existence of harmonic template neurons (HTNs) in auditory cortex that respond exclusively to 'high'-F0s (≥400-Hz) has been recently shown (macaque – Feng & Wang, 2017). The harmonics of these F0s are resolved at the periphery, yet HTNs responded only when stimulated by multiple harmonics of a particular F0 near their BF. Whether HTNs would respond similarly to harmonics in dichotic presentation remains an open question.

4.4.5 Summary

The findings of this chapter can be summarised as follows:

- Both MSO and LSO units show the first-effect of pitch shift in response to DCTs, irrespective of harmonic rank, F0 or stimulus configuration, resembling monaural first-effect of pitch shift data and in line with human psychophysical responses.
- Both MSO and LSO units show binaural integration of pitch in response to cpHTCs, with the peak of their ISI distributions, a neural correlate of pitch, occurring at the period of the overall F0 across ears.
- MSO units also respond to frzHTCs and rpHTCs in the same manner.
- Across the population of MSO and LSO cells, responses consistent with binaural integration can be seen for HTCs with F0s between 31.25-Hz and 250-Hz.

Chapter 5

Responses to complex tones in the VCN input to the SOC

Abstract

Units of the VCN provide the principal excitatory input giving rise to the monaural and binaural response properties of nuclei of the SOC. Spherical bushy cells provide ipsilateral and contralateral excitatory inputs to the MSO and ipsilateral excitatory inputs to the LSO. Globular bushy cells provide contralateral excitatory input to the MNTB, which in turn provides inhibitory inputs underpinning the functionality of the LSO and SPN. Units of the VCN themselves, however, are mostly monaural.

Responses of low-BF VCN units to monaural complex tones were compared to SOC unit responses elicited from binaural equivalents. In addition, the binaural pitch stimuli used in **Chapter 4 (page 101)** were presented to low-BF (<2-kHz) units of the VCN. No binaural interactions resembling those found in the MSO and LSO was observed in these units. System crosstalk was also measured using BF tone rate-level functions of VCN units. The threshold difference in ipsilateral-only and contralateral-only presentations was used to determine interaural crosstalk attenuation.

The results of this Chapter confirm that the VCN input to the SOC is monaural. There is evidence of temporal sharpening in low-BF cells that provide inputs to the SOC. The high level of interaural crosstalk attenuation, combined with the relatively low sound levels of stimulation used in binaural paradigms, suggest that interaural crosstalk does not affect the binaural responses observed in the SOC.

5.1 INTRODUCTION

Auditory nerve fibres (ANFs) arising from the cochlea bifurcate and terminate on cells of the ipsilateral cochlear nucleus (CN) (cat – Ramón y Cajal, 1909; guinea pig – Brown, 1987; cat – Tsuji & Liberman, 1997). Neurotransmission is mediated using the excitatory neurotransmitter glutamate (rat – Wang *et al.*, 1998). The ascending branch innervates the anteroventral CN (AVCN) (cat – Fekete *et al.*, 1982; cat – Liberman, 1991), which contains spherical bushy cells (SBCs) and stellate cells (guinea pig – Moore, 1986; guinea pig – Hackney *et al.*, 1990). The descending branch innervates the posteroventral CN (PVCN), which contains globular bushy cells (GBCs) and stellate cells (guinea pig – Hackney *et al.*, 1990). Other CN cell types are beyond the scope of this Thesis. **Figure 64** below shows the anatomical distribution of guinea pig CN cell types and their output pathways.

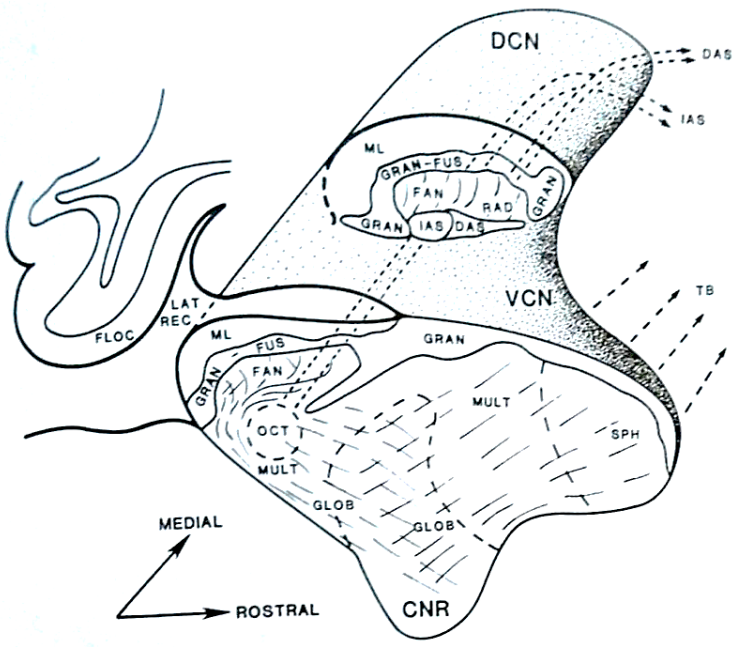

Figure 64 **Diagram showing the guinea pig cochlear nucleus, viewed from the lateral side**. Areas with concentrations of a certain cell type have been labelled as follows: SPH: spherical bushy cells; GLOB: globular bushy cells GBCs; MULT: multipolar (stellate) cells; other cell types are beyond the scope of this Thesis. Broken lines show the course of branches of the auditory nerve through the CN. Arrows indicate output pathways: TB: trapezoid body, output from the VCN going to the SOC. Other output pathways are beyond the scope of this Thesis. Figure excerpted from Moore, 1986.

5.1.1 Cell types

5.1.1.1 Bushy cells

SBCs receive input from between two and four ANFs (cat – Ryugo & Sento, 1991; rat – Nicol & Walmsley, 2002). The synapses between ANFs and SBCs are highly specialised, comprising highly-arborised axosomatic terminals termed 'endbulbs of Held' (cat – Brawer & Morest, 1975; guinea pig – Tsuji & Liberman, 1997). Synaptic activity between ANFs and SBCs can sometimes be observed in extracellular spike waveforms. Such waveforms show three components: 'PP' component – prepotential, the synaptic activity; 'A' component – the excitatory postsynaptic potential (EPSP) evoked by the PP, and; 'B' component – the action potential itself (cat – Pfeiffer, 1966a; guinea pig – Winter & Palmer, 1990; gerbil – Kopp-Scheinpflug *et al.*, 2002; gerbil – Typlt *et al.*, 2010).

Though the endbulbs of Held were long thought to optimise synaptic transmission (cat – Ryugo *et al.*, 1996), a large proportion of EPSPs fail to evoke action potentials (cat – Pfeiffer, 1966a; guinea pig – Winter & Palmer, 1990; gerbil – Kuenzel *et al.*, 2011). The cause of these 'spike failures' is thought to be glycinergic inhibitory inputs (gerbil – Kopp-Scheinpflug *et al.*, 2002; gerbil – Keine *et al.*, 2016). Possible benefits of spike failure include gain control and signal enhancement of the inputs underpinning interaural time difference sensitivity of MSO cells (gerbil – Keine *et al.*, 2016). Inhibition has also been shown to increases the strength of phase-locking of SBCs (gerbil – Dehmel *et al.*, 2010). Therefore, failures likely act to minimise spike timing jitter.

Post-stimulus time histograms (PSTHs) of SBC responses to short tone bursts at their best frequency (BF) are 'primary-like' (PL)[13] (cat – Pfeiffer, 1966b; cat – Rhode *et al.*, 1983; cat – Rouiller & Ryugo, 1984; cat – Smith *et al.*, 1993). This term was coined due to the similarity of the responses of PL units to those of ANFs (cat – Kiang *et al.*, 1965).

SBCs provide the main excitatory input bilaterally to the medial superior olive (MSO) and ipsilaterally to the lateral superior olive (LSO) (cat – Smith *et al.*, 1993).

In contrast to the few inputs to SBCs, GBCs receive convergent input from as many as 64 ANFs (cat – Liberman, 1993). The synapses between ANFs and GBCs are also highly specialised and are termed 'modified endbulbs of Held' (rat – Harrison & Irving, 1965; cat – Rouiller *et al.*, 1986).

PSTHs of GBC responses to BF short tone bursts are similar to PL unit responses with one key difference: the timing of the first spike is highly precise. This leads to a pronounced short-term decrease in the probability of firing immediately following the first spike (cat – Bourk, 1976). This response is termed 'primary-like with notch' (PN) (cat – Smith & Rhode, 1987).

GBCs comprise the main excitatory input to the medial nucleus of the trapezoid body (MNTB) (cat – Smith *et al.*, 1991) via another highly-specialised synapse, the calyx of Held; this is discussed in **Section 3.1.1 (page 48)**.

Phase-locking to low-frequency pure tones in SBCs and GBCs has been found to be stronger than observed in ANFs, *i.e.* these units show enhanced temporal precision of firing (cat – Joris *et al.*, 1994a, 1994b; cat – Joris & Yin, 1998; rat – Paolini *et al.*, 2001; macaque – Joris & van der Heijden, 2004; reviewed in Joris & Smith, 2008; chinchilla and cat – Recio-Spinoso, 2012).

[13] The convention throughout this Chapter will be to use unit types rather than anatomical cell types, as no anatomical verification of unit type was carried out in the AVCN experiments of this Thesis.

In GBCs, this is thought to be due to the high degree of convergent input from many ANFs – an average of 23 (cat – Spirou *et al.*, 2005) – combined with a short membrane time constant (guinea pig – Manis & Marx, 1991). However, SBCs can also display similar levels of temporal precision to GBCs in spite of receiving inputs from very few ANFs. This may result from axodendritic synapses from ANFs (cat – Liberman, 1991) that give rise to tens of subthreshold inputs to a given SBC. This, along with one or two major endbulb of Held inputs, gives rise to temporal enhancement in simulations (Rothman *et al.*, 1993; Rothman & Young, 1996). Guinea pig SBCs, however, seem to lack this temporal enhancement, with vector strengths (VSs) to pure tones not exceeding the range found in ANFs (Winter & Palmer, 1990). Additionally, some cat studies also did not see enhanced temporal precision in SBCs, *e.g.* Blackburn & Sachs, 1989.

5.1.1.2 Stellate cells

Two main types of stellate (or 'multipolar') cell are present in the VCN: T-stellate cells, so-called as their axons project predominately through the trapezoid body, and D-stellate cells, so-called as their axons project dorsal-ward (mouse – Oertel & Wu, 1990). D-stellate cells receive few somatic synaptic contacts, with synaptic terminals being widely distributed throughout proximal dendrites; T-stellate cells receive numerous somatic and dendritic contacts (cat – Cant, 1981).

Stellate cell responses to BF tone bursts show periodic firing ('chopping') at a frequency independent of unit BF; as such, they are termed 'choppers' (reviewed in Oertel *et al.*, 2011). The firing rates of D-stellate cells decrease to near zero after the first few chopping periods; these are termed 'onset choppers' (OC). T-stellate cells, however, respond throughout stimulus presentation (cat – Smith & Rhode, 1989; guinea pig – Winter & Palmer, 1990; guinea

pig – Palmer *et al.*, 2003). T-stellate cells form two classes based on the shape of PSTHs of their responses to BF tones: sustained choppers (CS) and transient choppers (CT). CS units show lower levels of spike timing jitter, leading to a longer duration of chopping than CT units (cat – Rouiller & Ryugo, 1984; cat – Blackburn & Sachs, 1989; guinea pig – Winter & Palmer, 1990).

To differentiate between CS and CT units, steady-state (15- to 20-ms after onset) responses to BF-tone bursts are analysed using the coefficient of variation (CV), the ratio of the standard deviation of the interspike interval (ISI) to the mean ISI. This forms a bimodal distribution (cat – Blackburn & Sachs, 1989); a criterion of CV≤0.30 characterises CS units, whereas a CV of greater than 0.30 characterises CT units. There is no evidence of an anatomical difference between T-stellate cells with CS or CT responses (reviewed in Oertel *et al.*, 2011).

5.1.2 Contralateral inputs

Tracer injections have shown CN commissural projections to all subdivisions of the CN in guinea pig (Shore *et al.*, 1992). As such, binaural responses could arise in the CN.

In response to stimulation with contralateral broadband noise, spontaneous rates of VCN units either decreased (30% of units) or did not change (65.5% of units) (guinea pig – Shore *et al.*, 2003). However, 4.5% of all recorded units from VCN showed contralateral excitation at high sound levels[14]. Another study using pure tones did not find any contralateral excitatory responses in VCN (guinea pig – Ingham *et al.*, 2006). *In vitro* studies have shown that inhibitory postsynaptic potentials (IPSPs) originating from inhibitory inputs are found in up to 70% of

[14] This was not broken down by unit type and may not include any PL or PN units.

principal VCN cells (guinea pig – Babalian *et al.*, 1999); these inputs are glycinergic (guinea pig – Babalian *et al.*, 2002).

A study using anatomical tracer injections in guinea pig CN showed that SBCs and GBCs do not project to the contralateral VCN (Cant & Benson, 2003). As neither SBCs nor GBCs receive convergent excitatory input from both sides of the auditory system, binaural integration of pitch in the manner described in **Chapter 4 (page 101)** is unlikely to occur in these cells.

5.1.3 Interaural crosstalk estimation

One of the principal challenges in binaural stimulation paradigms is ensuring that sound energy does not transfer from the stimulating apparatus at one ear to the other ear. There are multiple potential sources of this 'crosstalk'; the most difficult source to reduce or eliminate is acoustic crosstalk. This occurs due to signal leakage from the ear canal of the stimulated ear, or from conduction of vibrations through the speculae or even through the skull.

One method of determining crosstalk is to measure gross potentials of the cochlea. Thresholds of either the compound action potential (CAP) for high frequencies (*e.g.* guinea pig – Robertson & Irvine, 1989) or cochlear microphonic for low frequencies (*e.g.* cat – Gibson, 1982) are obtained from stimulation of either ear with pure tones. The difference in thresholds at a given frequency gives the interaural crosstalk (IXT) attenuation at that frequency.

Another method is to measure unit thresholds to BF pure tones in ipsilateral and contralateral stimulation, usually using ANFs (*e.g.* cat – Caird *et al.*, 1980). A similar paradigm was used in this study, making recordings instead from the VCN. Whether or not crosstalk can be seen at the level of the AN, crosstalk-evoked responses might not be transmitted to the SOC via the VCN. For instance, the weak contralateral inhibition seen in 44% of PL and PN units (guinea

pig – Ingham *et al.*, 2006) could cancel out responses that would otherwise arise due to crosstalk.

Identical binaural pitch paradigms to those used in the MSO (see **Sections 5.3.4, page 153** and **5.3.6, page 157**) supplemented the tonal crosstalk estimates. Absence of VCN responses indicative of binaural integration of pitch provide evidence that the SOC responses to binaural pitch stimuli in **Chapter 4 (page 101)** occur as a result of *de novo* processing in the SOC.

5.2 METHODS

5.2.1 Unit classification

VCN units were characterised using a scheme building on that delineated in Blackburn & Sachs, 1989. After audiovisually determining unit BF and threshold, 250 repetitions of a random-phase 50-ms pure tone were presented at unit BF, 20-dB above unit threshold (20-dB sensation level (SL)). A PSTH was plotted using the resultant response. Most VCN units can be characterised fully on the basis of responses to BF tone bursts at one level, however the PSTH can vary between levels. In these cases, responses to BF tones at 50-dB SL were also used to classify units.

Spike waveform shape was used to distinguish between fibre (axonal) recordings and cellular recordings, and also for detection of prepotentials in PL units (see **Section 2.4.2, page 41**).

The electrode was positioned under visual guidance into the CN (see **Section 2.1, page 32**). No histological confirmation was carried out at recording sites: the criteria for classifying CN units based on basic physiological response properties are well established.

5.2.2 Interaural crosstalk

IXT was measured using rate-level functions to ipsilateral and contralateral presentation. Ten repeats of a 50-ms BF-tone were presented in 2-dB increments; the spike rate was averaged across these repeats. A sigmoid was fitted in MATLAB to the ipsilateral rate-level function, then, using the function `nlinfit`, a sigmoid of the same shape was shifted in level to optimally fit the contralateral rate-level function. IXT was defined as the difference in level between the midpoints of the two fitted sigmoids.

The maximum stimulation level used was 116 dB-SPL. In some units, contralateral stimulation at this level did not elicit a response above spontaneous rate, leaving IXT undefined. The difference of the maximum sampled level in contralateral presentation and the ipsilateral threshold comprised a lower-bound of IXT attenuation in these cases.

5.3 RESULTS

5.3.1 Unit population

A total of 100 units were recorded from the left CN of 14 guinea pigs. Their BFs varied from 0.16-kHz to 26.50-kHz, mean±sd: 5.66 ± 6.63-kHz; 49 units had BFs<2-kHz ('low-BF'). Thresholds varied from 7 to 67 dB-SPL, mean±sd: 31 ± 13 dB-SPL. Across a given unit type, the threshold range does not exceed 40-dB at a given BF. See **Figure 65** below for the threshold distribution of the VCN population as a function of unit BF.

Spontaneous rates (SRs) varied from 0.0 to 157.0 spikes/s, mean±sd: 26.8 ± 33.6 spikes/s. $Q_{10\text{-}dB}$ varied in a BF-dependant manner, being lower at lower BFs, mean±sd: 2.37 ± 1.77.

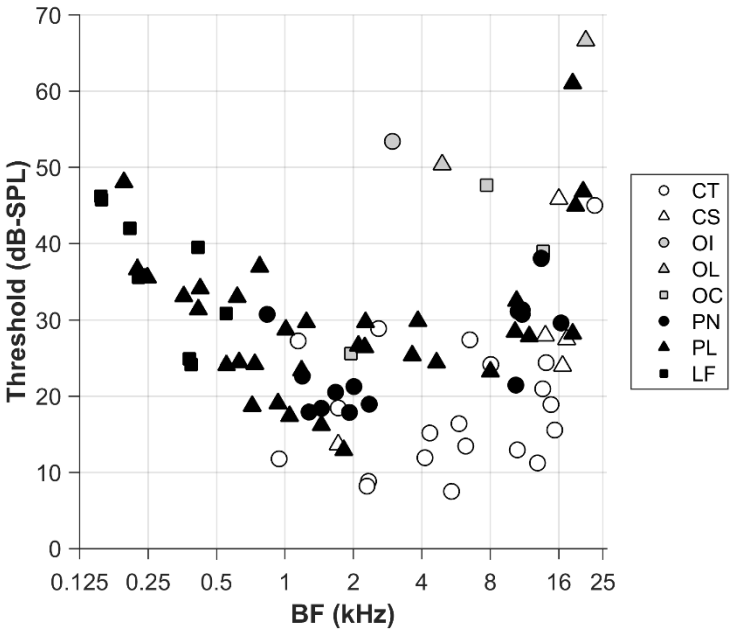

Figure 65 **CN unit threshold against unit BF**. PL and LF units show the lowest thresholds at low-BFs; at mid-to-high BFs, the lowest thresholds are shown by CTs. Onset units show the highest thresholds.

5.3.2 Basic response types

Obtaining low-BF PL units was an experimental priority, as these units provide the input to low-BF SOC units. PSTHs from example PL and PN units are shown in **Figure 66 (page 149)**. PL unit responses to BF pure tones are similar to ANF responses, with a relatively large response at onset declining to a steady state. PN units show a PL response with an additional notch in-between the onset peak and the steady-state response. PL units sometimes show a three-component extracellular spike waveform shape, with activity in the ANF>SBC synapse manifesting as the prepotential ('PP') component.

Recordings were also made from CT and CS units (see **Figure 67, page 150**). PSTHs (**Figure 67A**) of CT and CS unit responses to BF pure tones show periodicity. **Figure 67B** shows FOISI histograms of unit responses; these are unimodal, meaning that the ISIs cluster around a particular 'chopping frequency'. CT units have a greater degree of jitter in ISIs; CV≥0.30 averaged at 15-20 ms after stimulus onset separates the CT and CS populations (**Figure 67C**).

An additional classification, low-frequency (LF), was used for units with BFs below around 500-Hz, due to the difficulty of classifying units at these frequencies. Onset units treat every cycle of a low-frequency pure tone as a new stimulus rather than merely responding at stimulus onset. Additionally, low-BF choppers might have similar chopping frequencies to their BFs; in these cases, one is not able to differentiate chopping from phase-locking. However, PL units with BFs below 500-Hz can sometimes be distinguished when using BF tones with randomised starting phase. **Figure 68 (page 151)** shows an example of this. In **Figure 68A**, the PSTH of the PL unit doesn't show chopping, unlike the LF wide-mode-chopper (CW) unit. In **Figure 68B**, phase-locking is apparent in the multi-peaked FOISI distribution of the PL unit, but not the LF unit. This is not apparent in the PSTH due to the randomised starting phase of the BF

tones. **Figure 68C** shows the VSs of unit responses to pure tones. In the PL unit, these responses are enhanced above the range seen in guinea pig ANFs, plotted as grey lines (data from Palmer & Russell, 1986).

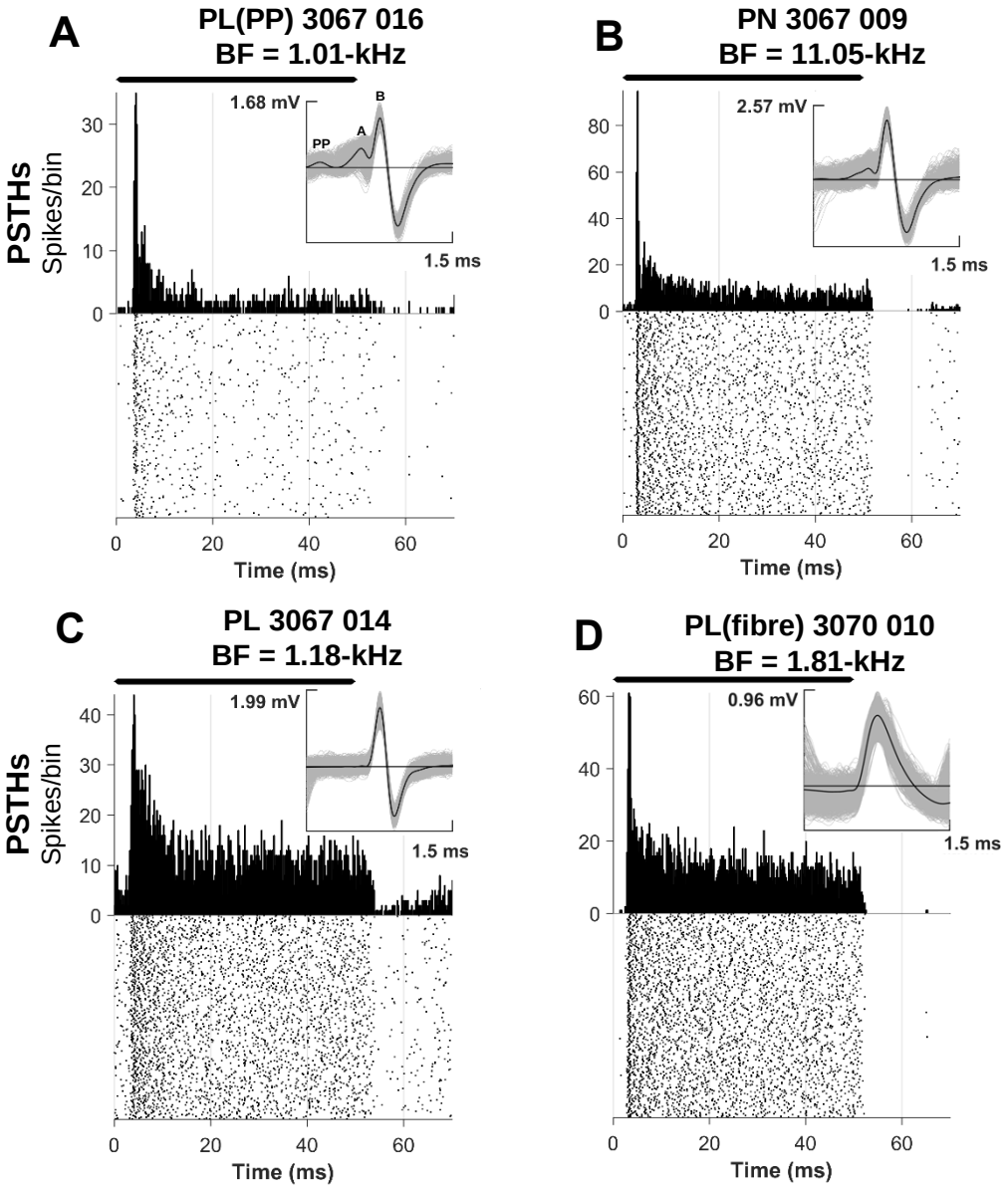

Figure 66 **PSTHs of PL and PN unit responses to BF-tones are similar, however PN units** (B) **can be distinguished by the presence of a notch in-between the onset peak and the steady-state response**.

A: PL unit showing a three-component extracellular spike waveform, with presynaptic activity from ANF observed in the form of prepotentials. **C**: PL unit showing no evidence of prepotentials.

D: PL fibres are also present in the CN, with a monopolar waveform typical of an axonal recording.

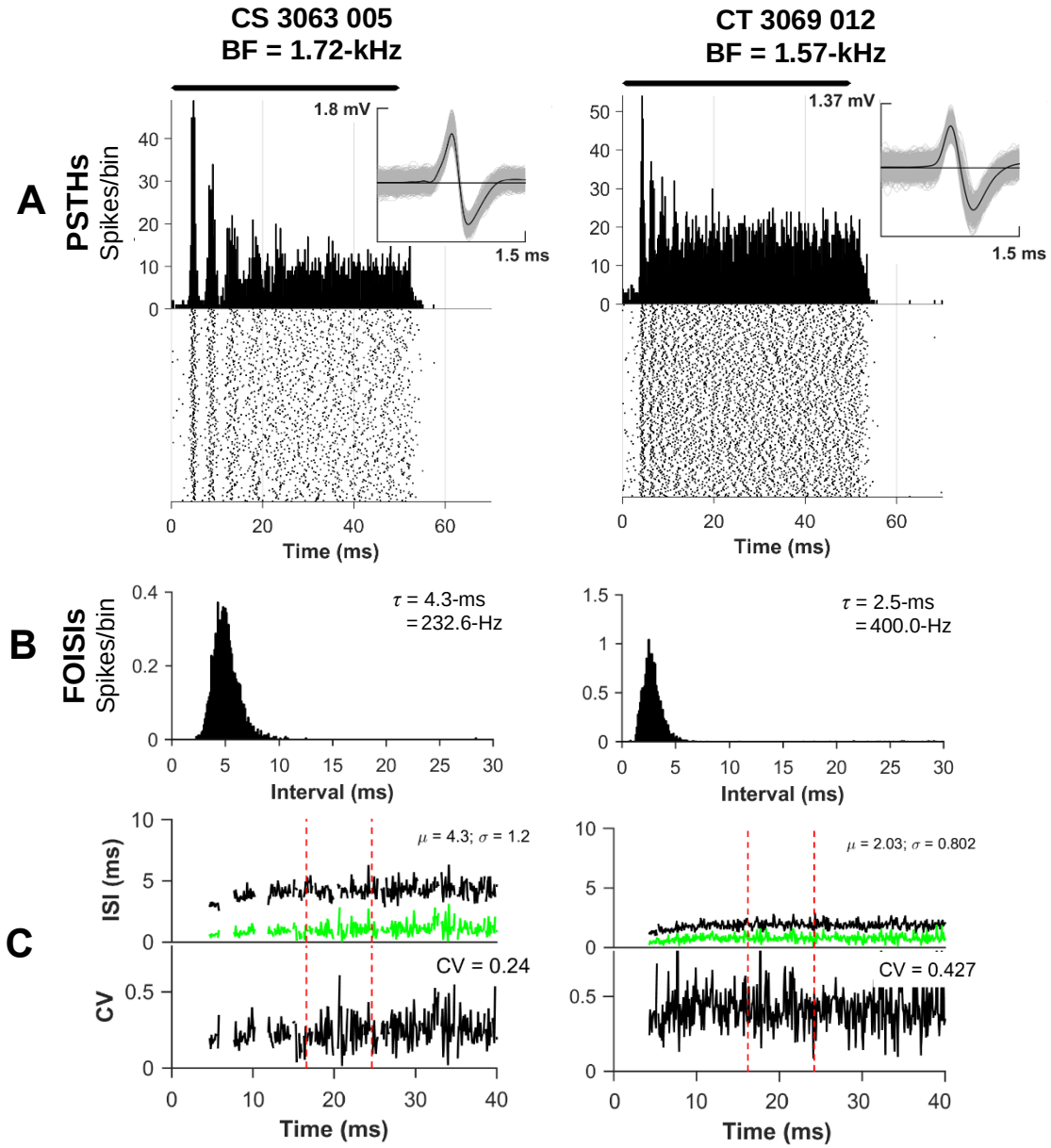

Figure 67 **Choppers show a pronounced bursting in response to pure tones at a 'chopping frequency' at a frequency unrelated to the unit BF**.

A: PSTHs and rasters of unit responses to BF pure tones. Average spike waveforms are inset.

B: FOISI histograms of unit responses. τ = chopping period; chopping frequency is also shown.

C: Top: mean ISI (black) and standard deviation of ISI (green) as a function of time; bottom: CV analysis, the standard deviation of ISIs (σ in the top plot) over the mean (μ in the top plot) during the analysis window (red lines).

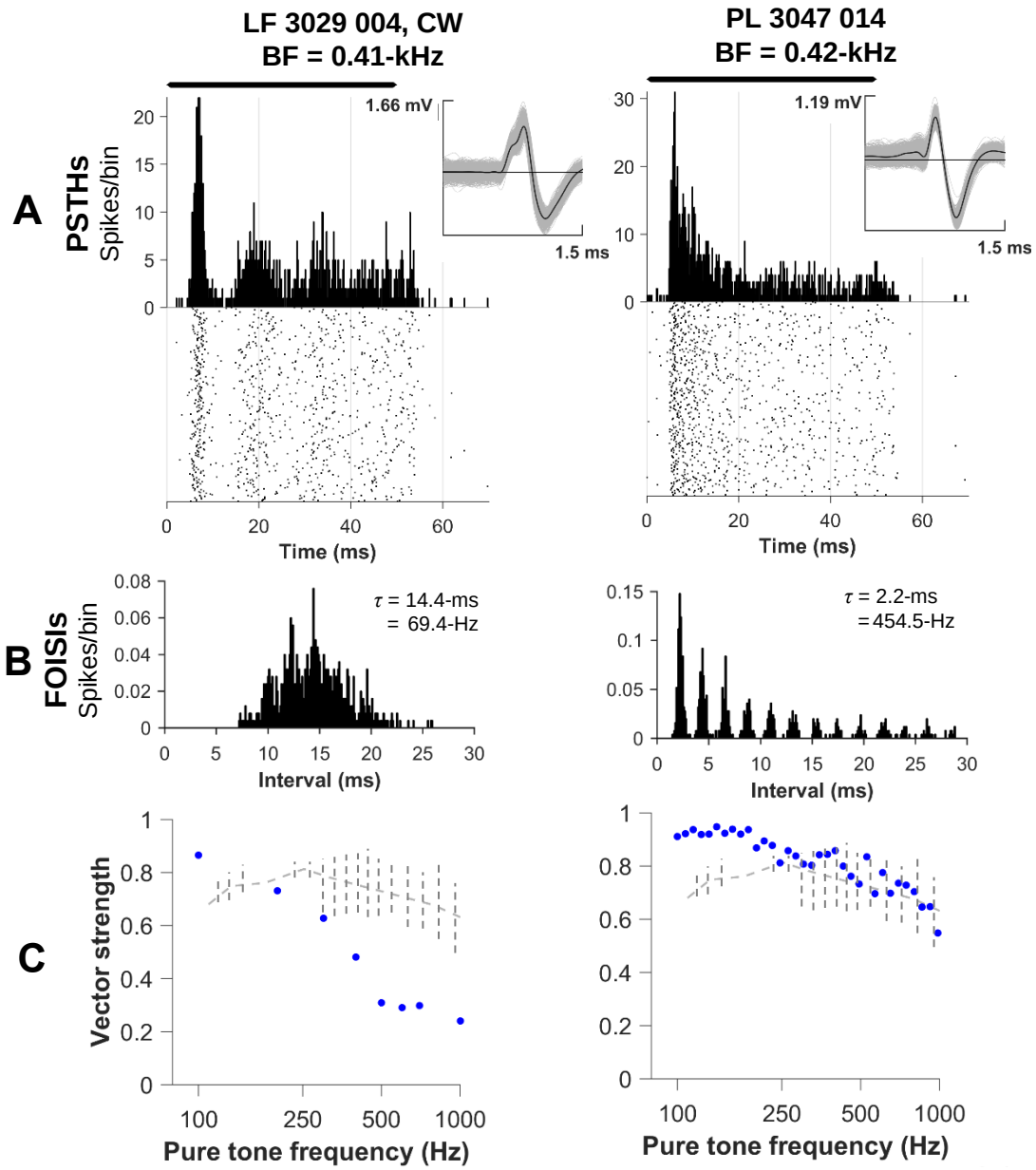

Figure 68 **Distinguishing between low-BF PL and LF units is sometimes possible**.

A: PSTHs of unit responses to random-phase pure tone bursts at BF. Average spike waveforms inset.

B: FOISI histograms. τ = chopping period; chopping frequency is also shown.

C: VS of pure tone responses. Receptive field of the PL unit is shown in **Figure 11 (page 38)**.

Recordings were also made of onset ideal (OI), onset locker (OL) and OC unit types. OI and OL are characterised with broad receptive fields and a strong response at the onset of short tone bursts; OL units show weak continued activity after onset whereas OI do not. DCN units and unusual VCN units were discarded.

5.3.3 Monaural harmonic tone complexes

Monaural equivalents of the cosine-phase harmonic tone complexes (cpHTCs) and dichotic complex tones (DCTs) used in **Chapter 4 (page 101)** were presented to low-BF VCN units; the responses of these units were compared to binaural responses of the SOC. Dichotic cpHTCs and DCTs were also presented to these units in order to test for binaural integration of pitch in the VCN input to the SOC, either genuine or arising due to crosstalk. See **Section 4.2.2 (page 107)** for a description of the cpHTC and DCT stimuli and analyses performed on unit responses to these stimuli.

Figure 69 (page 153) shows the responses of a low-BF PL unit to ipsilateral (*i.e.* monaural) cpHTCs in three different configurations: odd-only, even-only and all harmonics. PLs form the principal excitatory input to units of the MSO, so these monaural responses are similar to those signals that MSO cells receive. In other words, in the ipsi-odd, contra-even presentation, PL unit input to the MSO from the ipsilateral CN is similar to the ipsi-odd response shown here and PL unit input to the MSO from the contralateral CN is similar to the ipsi-even response.

Shuffled autocorrelograms (SACs, see **Section 4.2.3, page 109**) of responses to odd-only cpHTCs resemble those seen in the MSO (see **Figure 52, page 120**). Two peaks surround the 2×F0-period (red line), and a larger peak occurring at the F0-period (green line). These three peaks are correlates of the ambiguous pitches that human subjects report, and their positions vary as a function of BF and F0 (see **Section 4.3.2, page 117** for an explanation).

The responses to even-only cpHTCs show a clear peak at the 2×F0 period. The responses of this unit to the ipsi-all presentation are similar to response to even-only responses, except that the peak at the 2×F0 period disappears, leaving the largest non-zero peak at the F0-period. The maximum normalised coincidences value is greater in this condition than the ipsi-odd and ipsi-even conditions, reflecting greater temporal precision of firing.

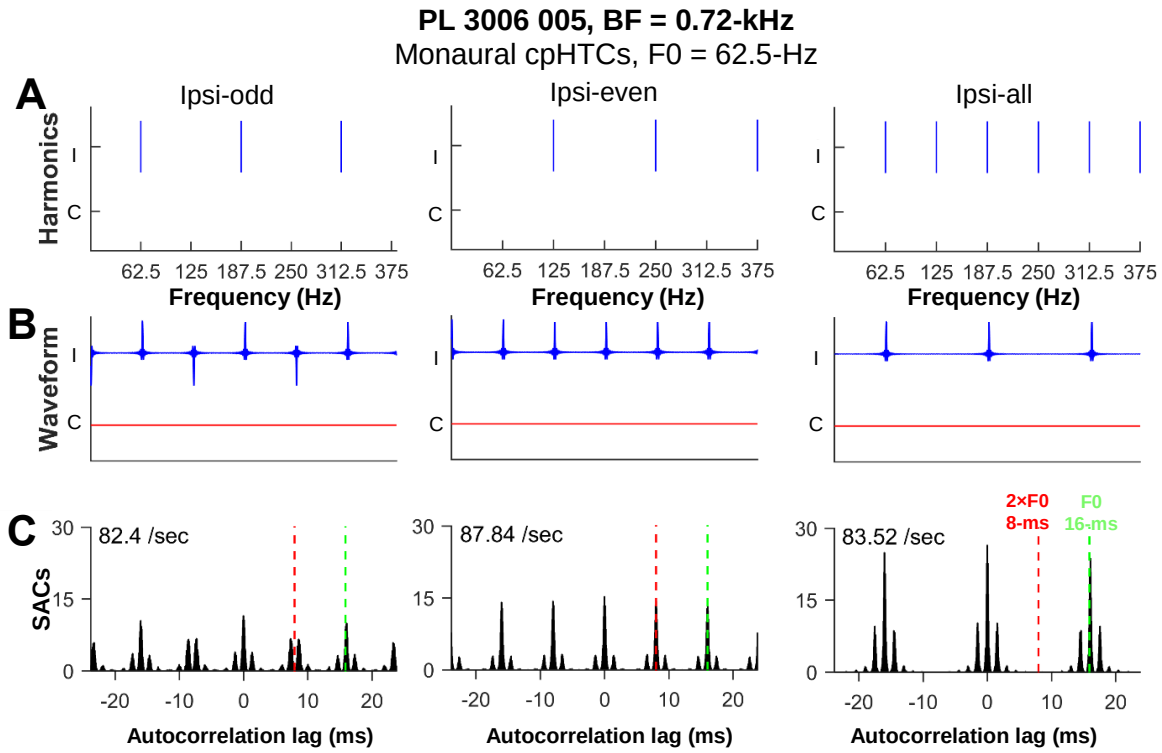

Figure 69 **PL unit responses to monaural cpHTCs, F0 = 62.5-Hz, presented in three different configurations**.

A: Schematic harmonic ladders (note that the complexes extend to 5-kHz).

B: Stimulus waveforms.

C: SACs, normalised according to Louage et al., 2004. The red line is centred at the period 2×F0 (8-ms); the green line at the period of F0 (16-ms). Spike rates are inset. Unit SR = 6.0 /sec.

5.3.4 Dichotic harmonic tone complexes

Responses of low-BF VCN units to dichotic cpHTCs were assessed for binaural integration of pitch. Units likely forming the input to the SOC (in particular, the MSO) did not show evidence

of binaural integration in dichotic stimulation, instead responding to the waveform at the ipsilateral ear. This, together with estimates of crosstalk using the threshold difference of ipsilateral and contralateral rate-level functions (discussed in **Section 5.3.9, page 164**), indicates that the results of **Chapter 4 (page 101)** are very unlikely to be due to crosstalk.

Figure 70 (page 155) shows responses of a low-BF PL unit to dichotic cpHTCs. The response to dichotic stimulation is almost identical to that resulting from ipsilateral-only stimulation. Both the monaural ipsi-odd and the dichotic ipsi-odd, contra-even response show the ambiguous pitch response discussed in **Section 5.3.3 (page 152)**. Swapping the dichotic stimulus configuration changes the response profile: the dichotic response changes to reflect the ipsi-even input. The difference in response profile seen in these cells stands in contrast to cells of the SOC, which respond to the overall F0 across ears in both dichotic presentations (see **Figure 52, page 120**).

The largest non-zero peaks in the SACs of low-BF unit (n = 26) responses to cpHTCs were obtained, then the autocorrelation lags at which peaks occurred were normalised so that the F0-period corresponded to 1. Scatter plots of these data are shown in **Figure 71 (page 156)**, separated by stimulus configuration. Units do not show responses to contralateral-only stimulation. In the ipsi-odd and ipsi-odd, contra-even cases, the three largest peaks occur at $0.5 \pm \frac{\mathrm{BF}}{2\times\mathrm{F0}}$ and 1 (see **Section 4.3.2, page 117**), as was seen in SOC cell responses (see **Figure 61, page 129**). In the ipsi-even and ipsi-even, contra-odd cases, there are two peaks of approximately equal amplitude at F0-period (1) and 2×F0-period (0.5). A two-sample t-test was conducted on the distributions of peak normalised coincidence values obtained from ipsilateral-only and dichotic presentations. No significant difference was found between the two populations ($p>0.05$). In conclusion, unlike units of the SOC, there is no evidence of binaural integration of pitch of cpHTCs in low-BF units of the VCN.

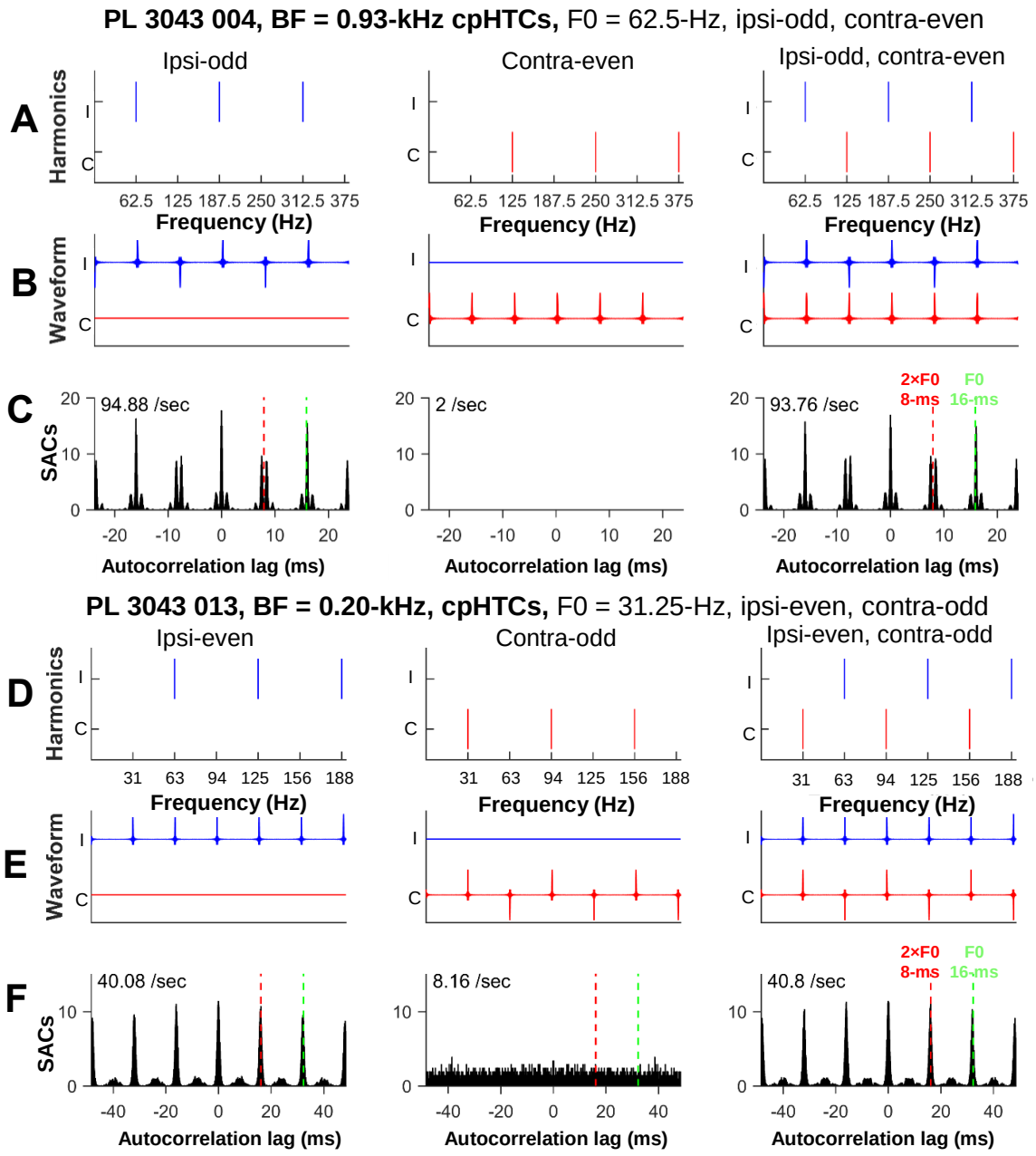

Figure 70 **No evidence of binaural integration of pitch in low-BF CN units**. Instead, near-identical responses to cpHTCs occur in dichotic presentation (3rd column) to those resulting from ipsilateral-only presentation (1st column). Contralateral stimulation does not result in a response above unit SR. Conventions as in **Figure 69 (page 153)**. PL 3043 004 SR: 0.9 /sec; PL 3043 013 SR: 13.9 /sec.

A and **D**: schematic harmonic ladders. **B** and **E**: stimulus waveforms. **C** and **F**: SACs of unit responses, normalised according to Louage et al., 2004.

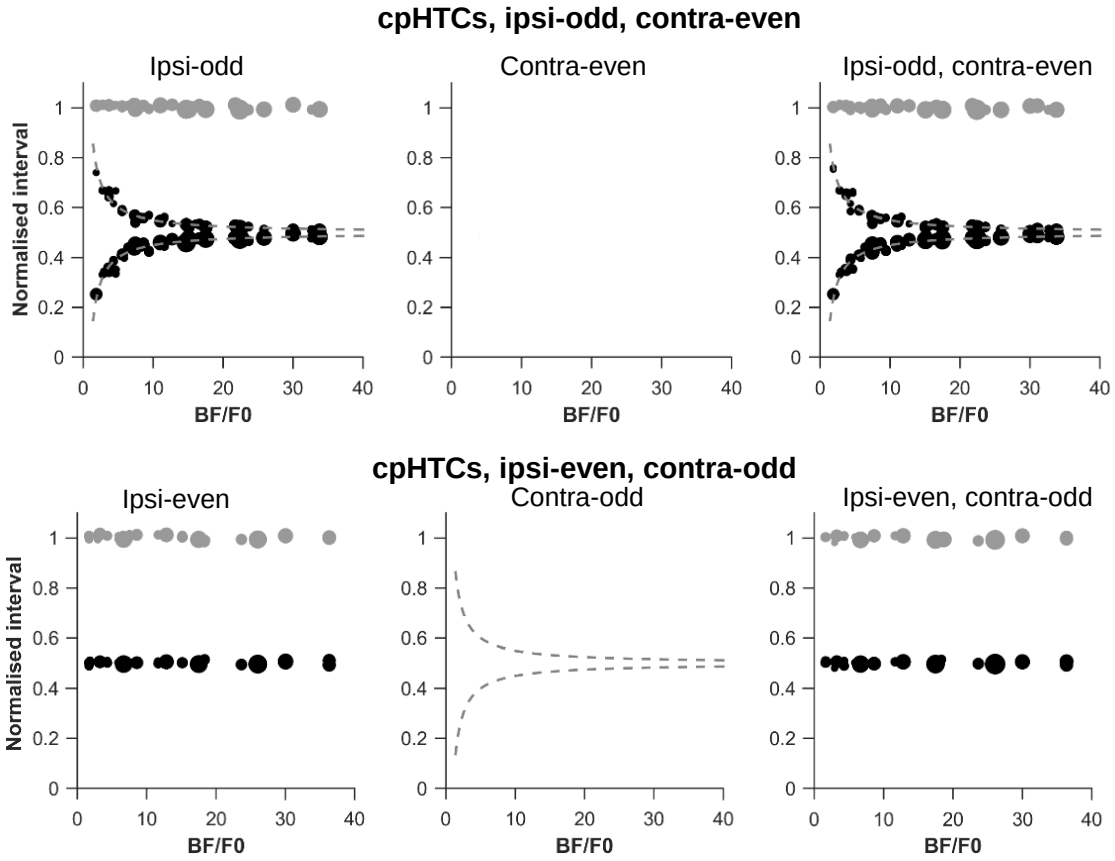

Figure 71 **Low-BF VCN unit responses to dichotic presentation of HTCs do not show binaural integration of pitch; instead, units respond exclusively to the ipsilateral stimulus**. *n* = 25. The size of scatter point being proportional to the normalised coincidence value (larger represents a stronger response). Responses were rejected if spike rate did not exceed 5 /sec. Normalised intervals are plotted against BF/F0 (see **Section 4.3.2, page 117**).

5.3.5 Monaural complex tones

As in **Section 5.3.3 (page 152)**, complex tones of two consecutive harmonics, similar to the DCTs used in **Section 4.3.1 (page 110)**, were presented monaurally to low-BF PL units. This allowed comparison of responses between monaural presentation in the VCN and binaural presentation in the SOC. Responses to monaural complex tones were measured in 13 low-BF units. An example PL unit's responses is shown in **Figure 72 (page 158)**. In **Figure 72A**, SACs of 5 of the 9 sampled conditions are shown. As the monaural complex is frequency-shifted, the

peaks in the SAC shift to shorter periods. After shifting by F0-Hz (in this case, 100-Hz), the unit responds again responds at the ENV period of 10-ms. **Figure 72B** shows a continuous plot of the 9 SACs. The slope of the line represents a neural correlate of the first effect of pitch shift to monaural complex tones. In **Figure 72C**, the data are transformed onto polar axes. The F0-period shift appears as an inward spiral.

The population average of PL unit responses to monaural complex tones are compared with the population average of MSO unit responses to DCTs in **Figure 75 (page 160)** – both show the first effect of pitch shift. Note that the response is more temporally precise in the MSO compared to the VCN, reflected in a greater normalised coincidences value across the population averaged SACs; the averaged normalised coincidence value peaks at 12.5 in the MSO population, but only 4.2 in the VCN.

5.3.6 Dichotic complex tones

The same DCTs used in the SOC were presented to low-BF units of the VCN to determine if a binaural integrative response occurred in the VCN input to the SOC.

Figure 73 (page 159) shows responses of the same low-BF PL unit as **Figure 72 (page 158)** to DCTs. The response is markedly different here than the responses to monaural complex tone stimulation. The largest non-zero peak in the SAC occurs at the period of the ipsilateral carrier rather than the period of the F0 across ears. The effect was similar when the stimulus configuration was reversed, *i.e.* the higher harmonic was instead presented to the ipsilateral ear. This was seen in all sampled low-BF VCN units (n = 13); population average data are plotted in **Figure 74 (page 160)**. Thus, the input to MSO cells in response to DCTs from any

given PL unit, as also shown in **Section 5.3.4 (page 153)**, does not itself show binaural integration of pitch.

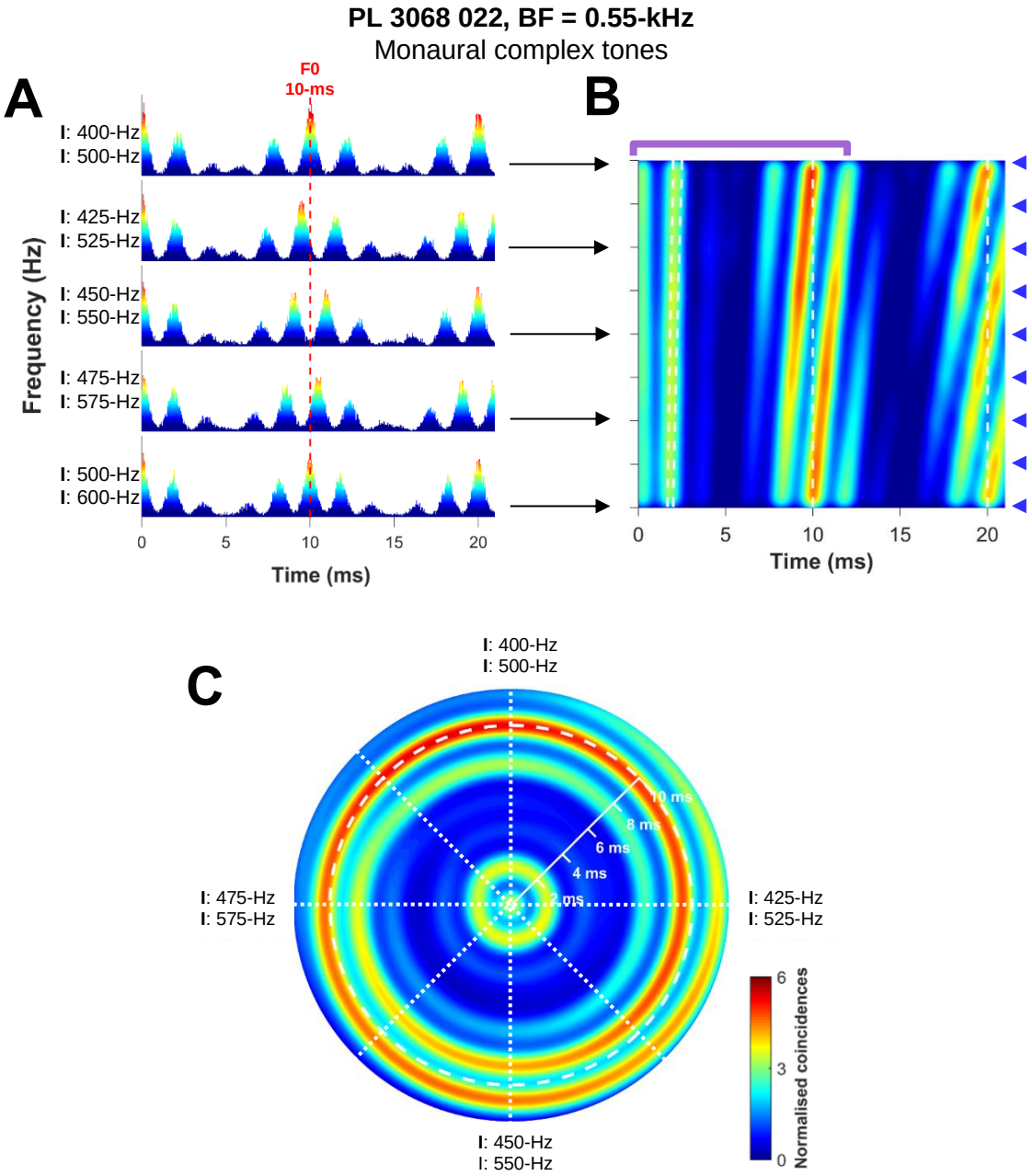

Figure 72 **F0-shift in the responses of a CN unit to monaural complex tones presented ipsilaterally**.

A: SACs of responses to complex tones. 'Normalised coincidences' are represented as a colour scale. The 10-ms period is shown by the dotted line.

B: Continuous plot of the 9 SACs, filtered with a Gaussian kernel. Arrowheads represent the sampled frequency pairs. The purple brace denotes the region plotted in **C**.

C: Polar plot of the linear interpolation in **B**. The period of the 100-Hz spacing, 10-ms, is shown as a dotted circle.

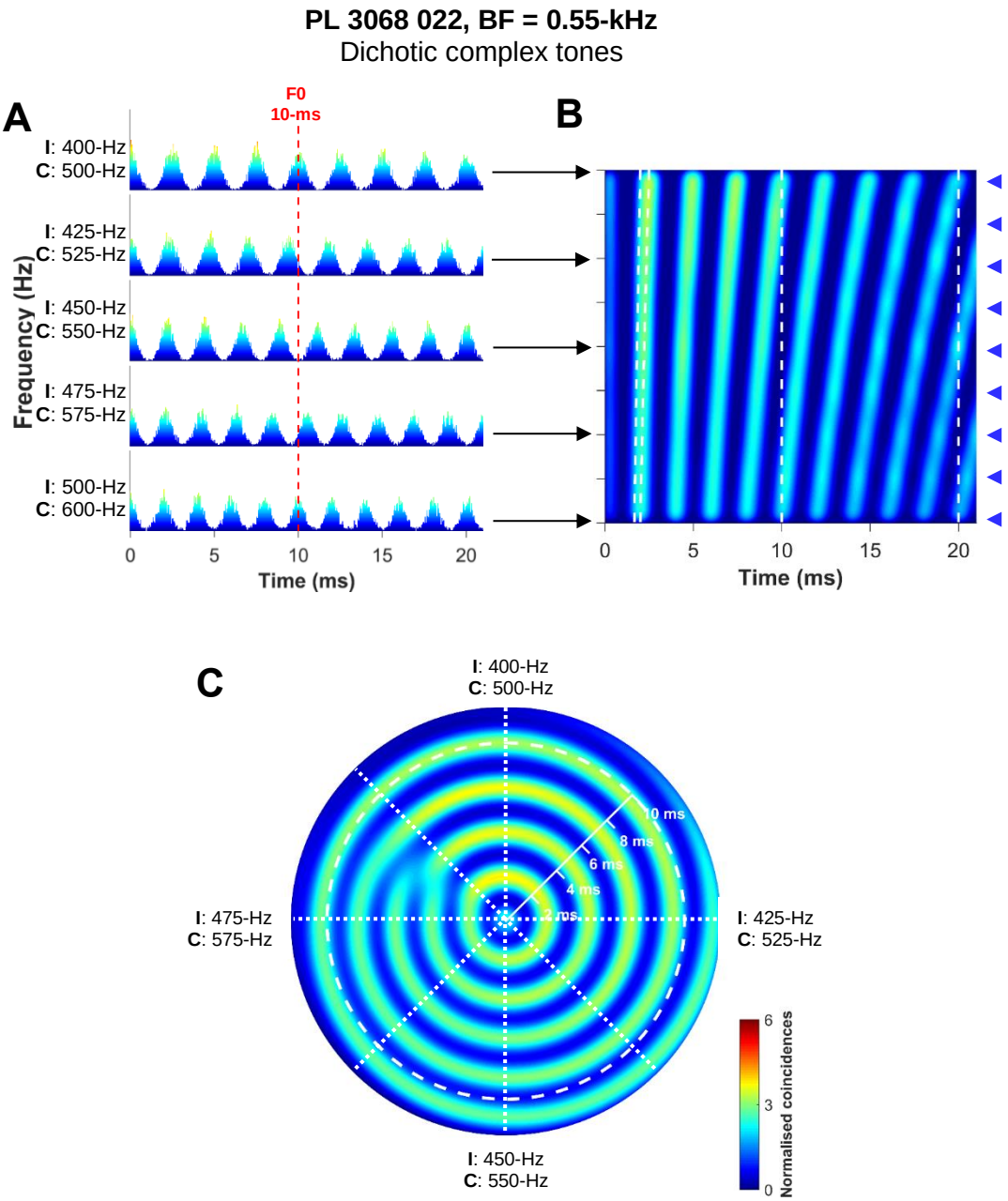

Figure 73 **When complex tones are presented dichotically, low-BF PL units respond to the ipsilateral carrier alone**. Conventions as in **Figure 72 (page 158)**.

A: SACs for 5 of the 9 DCT conditions. Unlike **Figure 72** (same unit), the largest non-zero peak is no longer at 10-ms; rather, it is at the period of the carrier in the ipsilateral ear; in the 400-Hz/500-Hz case, this is 2.5-ms.

B: The largest non-zero peak occurs at the period of the contralateral carrier.

C: Polar plot of **B**. The first-effect of pitch shift is not seen and the response is locked to the ipsilateral carrier throughout.

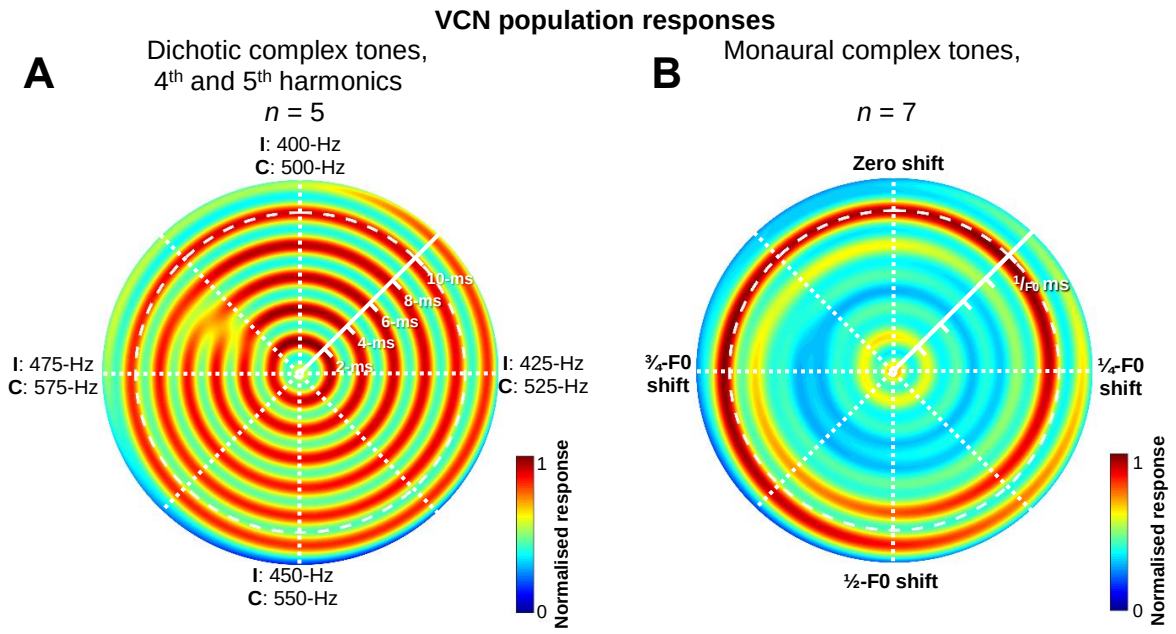

Figure 74 **The first-effect of pitch shift is seen in VCN population responses to monaural complex tones but not DCTs**. Conventions as in **Figure 72 (page 158)**.

A: Low-BF VCN unit responses to DCTs show locking to the carrier in the ipsilateral ear. Compare this with the MSO population average response to DCTs in **Figure 75B** below, where the response is to the overall F0.

B: In contrast, responses to monaural presentation of the complex tone show a neural correlate of pitch at the F0-period which shift, in line with the first effect of pitch shift.

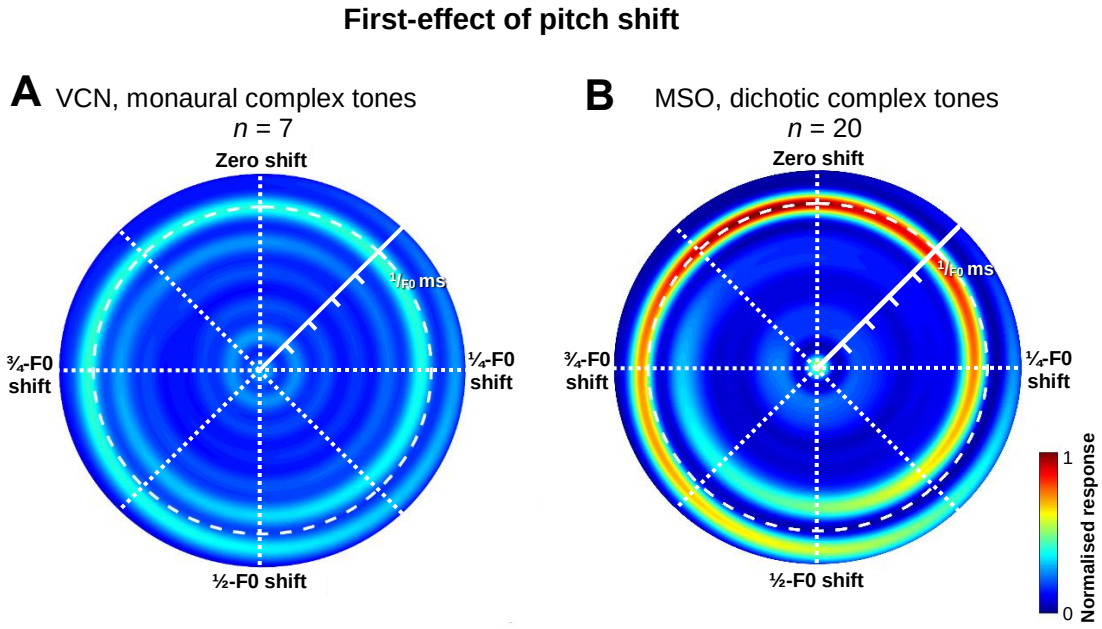

Figure 75 **First-effect of pitch shift shown monaurally in VCN (**A**) and binaurally in MSO (**B**)**. In both cases, the units respond in a manner analogous to the shift in pitch percept reported by humans (see **Section 4.1.1, page 103**). Conventions as in **Figure 72 (page 158)**.

5.3.7 Spike failure

Spike failures were encountered in 8 PL units (23% of all PLs). Two units had EPSPs of sufficiently large amplitudes to enable reliable triggering online of either the EPSPs or the spikes; this allowed for comparison of the resultant ANF input and spike output. As can be seen in **Figure 76 (page 162)**, a lower response rate was observed both to tones across the receptive field and to BF pure tones.

Figure 76A shows PSTHs of responses to 250 presentations of a 50-ms, 20-dB SL BF pure tone. There are substantially more EPSPs (9816) than spikes (1251), meaning that 87.3% of PP-A events failed to evoke spikes. **Figure 76B** shows receptive fields using a common colour scale; a substantially weaker response is seen in the output (spikes) compared to the input (EPSPs arising from ANF inputs). The lack of response from units such as these made measuring the IXT of these units, which form inputs to MSO, particularly problematic.

In these two units, IXT attenuation and responses to monaural and dichotic HTCs were measured triggering off either the input or output. As discussed in **Section 5.1.1.1 (page 140)**, it was hypothesised that spike failure leads to greater signal-to-noise ratios and temporal enhancement of the PL input to the MSO. **Figure 77 (page 163)** shows responses of a PL(PP) unit to a monaural HTC presented to the ipsilateral ear, F0 = 62.5-Hz (**Figure 77A** and **B**). In the left column, the trigger position was set to capture all PP-EPSP events including those that did not result in spikes; this can be used as a proxy for the ANF inputs to this unit. In the right column, the trigger position was set to capture only evoked spikes. In both cases, the unit locks to the F0-period of the HTC: that there is a peak in the SACs at the period of the F0 (16-ms), demarcated by the green line. The spike rate is lower than the EPSP rate (**Figure 77C, inset**),

reflecting spike failures. However, the peak normalised coincidence value[15] is greater in the output compared to the input of this unit, *i.e.* the temporal coding is more precise in the output than the input. Future studies will be needed to verify this finding across a larger population of low-BF PL units.

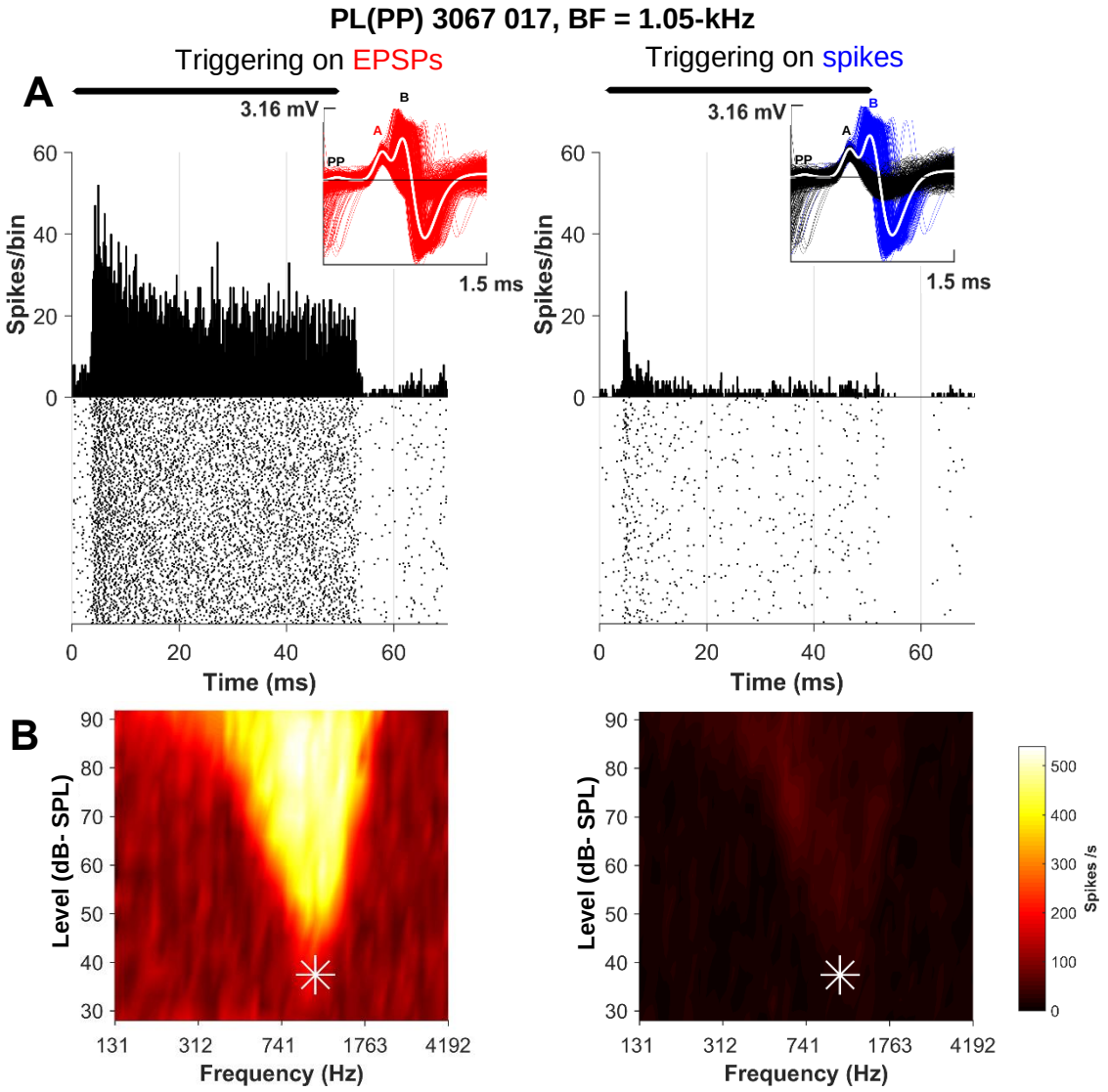

Figure 76 **Low-BF PL(PP) unit showing spike failures**. The left column shows responses triggered off the EPSPs, demarcated as the 'A' component in the average spike waveform; in this way, PP-A and PP-A-B events are triggered. The right column shows responses triggered from only spikes, *i.e.* only PP-A-B events.

A: PSTHs of responses to 250 presentations of a 50-ms, 20-dB SL BF pure tone.

B: Receptive fields are plotted using a common colour scale. Fitted BF and threshold are marked by the white asterisks.

Spike waveform analysis of this waveforms recorded from this unit can be seen in **Figure 15 (page 44)**.

[15] The normalisation used in calculating SACs (Louage *et al.*, 2004) takes spike rates into account.

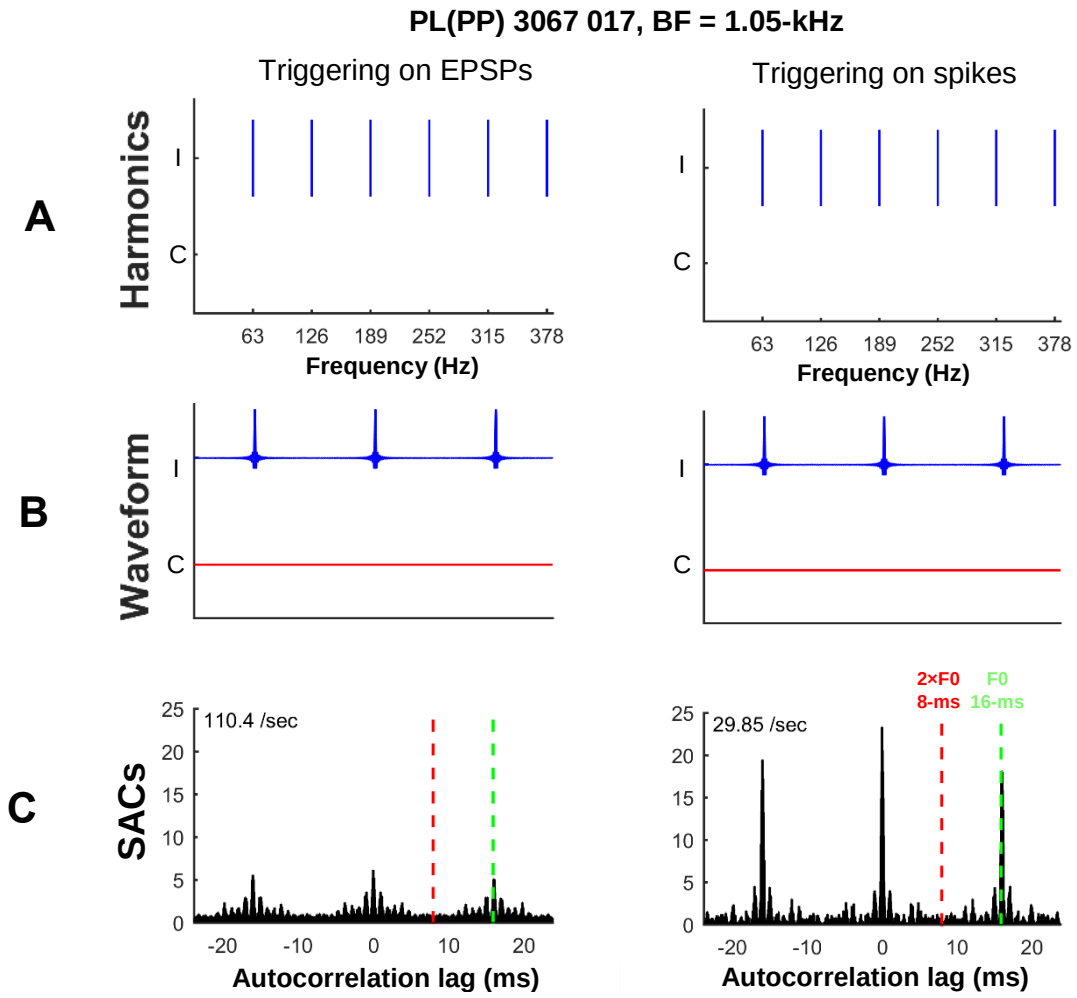

Figure 77 **Enhancement in the temporal precision in the output of a PL unit (same unit as Figure 76, page 162) compared to the input, analysed using SACs**. In the left column, the trigger position was set to capture all PP-EPSP events including those that did not result in spikes; this can be used as a proxy for the ANF inputs to this unit. In the right column, the trigger position was set to capture only evoked spikes. In both cases, the unit locks to the F0-period of the HTC. Unit SR, triggered off spikes, was 5.2 /sec; triggered off EPSPs, SR was 106.7 /sec.
A: Schematic harmonic ladders; note that harmonics extended to 5-kHz.
B: Stimulus waveforms.
C: Normalised SACs.

5.3.8 Synchrony

The strength of phase-locking to pure tones was measured in 11 low-BF PL units. Previous studies in guinea pig did not find enhanced temporal precision in PL units compared to the ANF inputs, in contrast to other species (see **Section 5.1.1.1, page 140**). This was in spite of

PL units of the guinea pig showing spike failures, found in other species and thought to enhance temporal coding.

Fixed phase short tone bursts were presented at sound levels that guaranteed that the stimuli exceeded 20-dB SL across all sampled frequencies. Responses were analysed using VS (Goldberg & Brown, 1969); data are plotted in **Figure 78** below. Each point shows the VS of a unit's response to a pure tone of a particular frequency. The majority of points are above the VS average of guinea pig ANFs (horizontal grey dotted line), and some points exceed the range seen in ANFs (vertical grey lines).

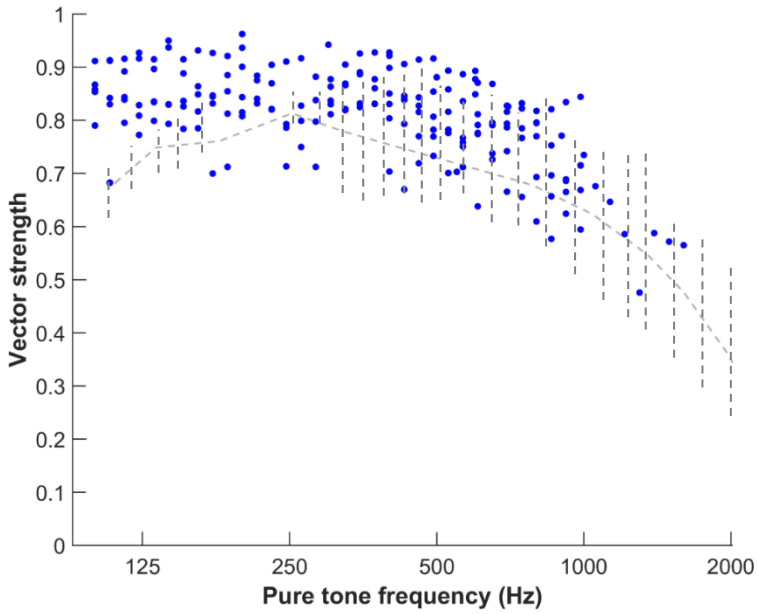

Figure 78 **Vector strengths of PL unit responses to pure tones show enhancement of phase-locking compared to ANFs**. *n* = 11. The dotted lines show the range of VSs found in guinea pig ANFs (replotted from Palmer & Russell, 1986).

5.3.9 Interaural crosstalk

IXT was measured subtracting the fitted midpoints of rate-level functions at unit BF obtained for ipsilateral and contralateral presentations (*n* = 66). In many cases (*n* = 34), it was not

possible to stimulate VCN units with contralateral tones even at the highest stimulation level. This left IXT undefined, however, a lower bound of IXT was found by subtracting the ipsilateral threshold from the maximum level used in the contralateral ear. Some examples of rate-level functions to ipsilateral and contralateral stimulation, and an example of the IXT calculation methodology, are plotted in **Figure 79** below.

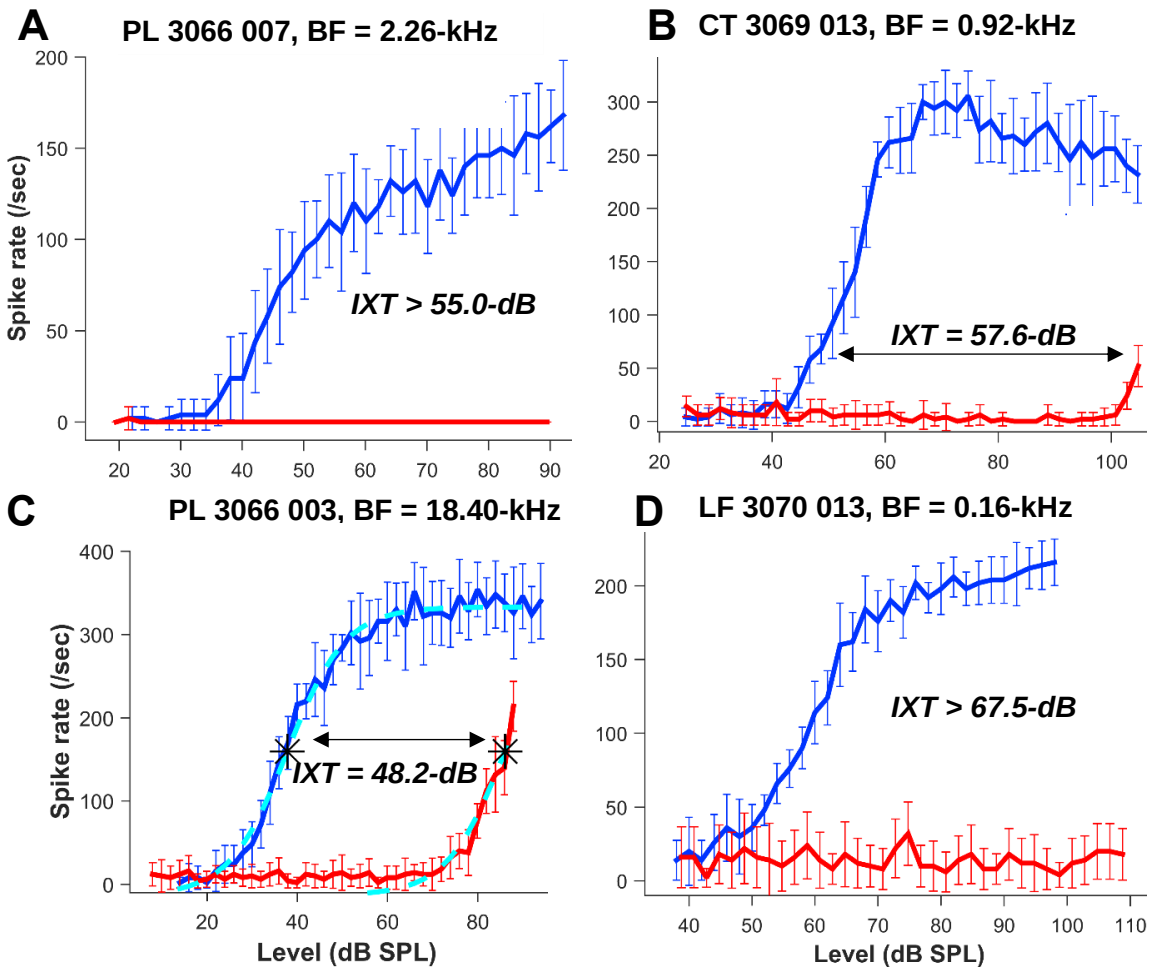

Figure 79 **IXT estimation for ipsilateral (blue) and contralateral (red) BF-tone rate-level functions**.

A: PL unit showing no evidence of crosstalk even at high contralateral stimulation levels. This leaves IXT undefined, with a lower bound of 55.0-dB.

B: Low-BF CT unit showing evidence of crosstalk at high contralateral stimulation levels. This is the only unit with BF<1-kHz that had a defined IXT.

C: The IXT calculation methodology is shown: a sigmoid was fitted to the ipsilateral response, then shifted in level (*i.e.* rightwards on the x-axis) to optimally fit the contralateral. IXT attenuation was calculated by subtraction of fitted midpoints (black asterisks).

D: A low-BF PL unit showing no evidence of crosstalk at high contralateral stimulation levels. IXT exceeds 67.5-dB.

Across all frequencies with defined IXT, the average IXT was 59.1-dB SPL. The entire dataset has been plotted in **Figure 80** below.

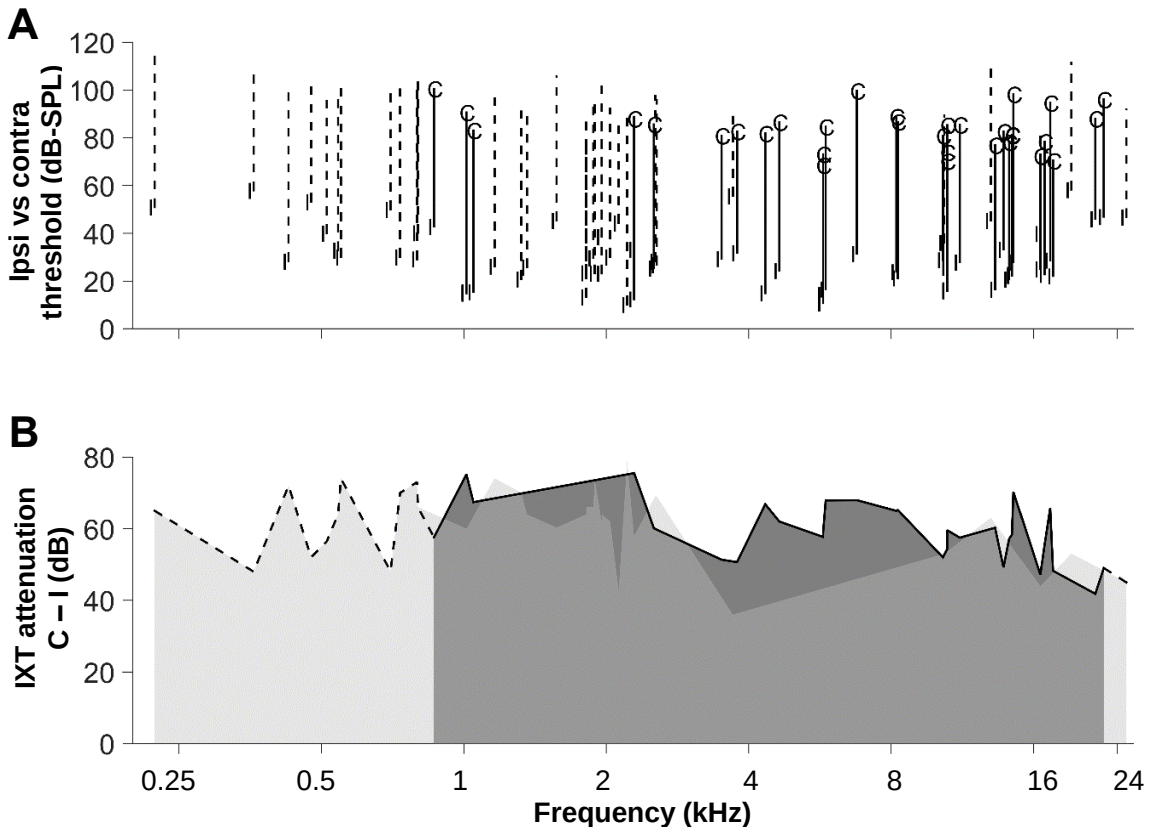

Figure 80 **IXT attenuation is at least 50-dB across the entire frequency range**. *n* = 66.
A: Ipsilateral versus contralateral threshold in dB-SPL. Dotted lines reflect undefined IXT, *i.e.* where no contralateral threshold could be found.
B: IXT attenuation as a function of frequency. Dark region: defined IXT. Light region: lower-bound of IXT, undefined at sampled points.

The majority of units of the MSO population in this Thesis (see **Section 3.2.2, page 67**), 80%, have BFs below the lowest frequency with a defined IXT (57.6-dB), 0.86-kHz. Additionally, all MSO units had BFs below 2-kHz; the average lower-bound IXT estimate below 2-kHz was

64.2-dB. In order for crosstalk to affect the binaural responses of low-BF MSO and LSO cells, contralateral stimulation levels would have to exceed these levels. The signal, greatly attenuated, would likely be far below unit threshold and, even if energy entered the filter, the response would be dominated by the ipsilateral stimulus. These results suggest that IXT is highly unlikely to underpin the binaural responses throughout this Thesis.

5.4 DISCUSSION

5.4.1 Monaural pitch responses

Correlates of the first effect of pitch shift have previously been shown for frequency-shifted monaural HTCs in guinea pig VCN (Sayles & Winter, 2008b) and to sinusoidal amplitude-modulated tones in cat ANFs and VCN (cat CN – Rhode, 1995; cat AN – Cariani & Delgutte, 1996b). **Section 5.3.5 (page 156)** shows low-BF VCN unit responses to frequency-shifted monaural complex tones consisting of only two consecutive harmonics. The largest non-zero peak in SACs shifts to lower periods – reflecting an increased F0 – proportionally to the frequency shift. This is a neural correlate of the first effect of pitch shift.

VCN PL units have been shown to encode the F0 of cpHTCs (guinea pig – Palmer & Winter, 1993; chinchilla – Sinex, 2008). This was also found in this study (see **Section 5.3.3**, **page 152**). The periods at which peaks in SACs of responses to odd-only cpHTCs occur are fitted by lines indicating predicted autocorrelation peaks (Bilsen & Ritsma, 1969). These peaks correlate with the ambiguous pitch percepts reported in human subjects. The periods at which peaks in the SAC occur vary as a function of BF and F0: ½F0 ± ½BF (in kHz); this has been shown previously (guinea pig – Sayles & Winter, 2007, 2008b).

5.4.2 Temporal coding in inputs to MSO

PL units of the guinea pig VCN show enhancement of phase-locking to low-frequency tones compared to ANFs (**Section 5.3.8, page 163**). This was previously a species-specific difference that could have made comparing the MSO responses of guinea pigs to other species problematic.

The question arises: why might previous studies of the guinea pig (*e.g.* Winter & Palmer, 1990) not have found enhanced phase-locking in PL units? The object of this Chapter was to characterise the properties of low-BF units of the VCN; it is possible that the low-BF oversampling led to this study standing a greater chance of finding PL units with enhanced phase-locking compared to previous studies. Another possibility is that VS is not necessarily maximal for tones 20-dB above the threshold at a given frequency. The paradigm used by Winter & Palmer, 1990 consisted of determining unit thresholds at each frequency using a threshold-tracking algorithm, then presenting tones at 20-dB SL. In this Thesis, a constant stimulus level was used across frequency: tones around unit BF could be in excess of 50-dB above threshold.

It has been proposed that spike failures in SBCs increase the signal-to-noise ratio and the temporal precision of responses (gerbil – Keine *et al.*, 2016). Some evidence of this was found in low-BF PL units showing spike failure: the temporal coding of HTCs was sharper in the SAC of the output spikes of the unit than in the EPSPs of the ANF input (**Section 5.3.7, page 161**.

5.4.3 Ruling out crosstalk as a factor in binaural responses

No evidence for interaural crosstalk underpinning the results of **Section 4.3 (page 110)** was found in the VCN unit population. Comparing pure-tone thresholds ipsilaterally and

contralaterally (**Section 5.3.9, page 164**) gave an estimate of average IXT in the sound system of 59.1-dB. At most stimulation levels in binaural paradigms, there was likely no crosstalk at all. IXT was undefined in most units with BFs<1-kHz, making it less likely that there could be any crosstalk-evoked contamination in the neural inputs to MSO. The lack of responses to ipsilateral tones in contralateral-only units of the SOC (**Section 3.3, page 70**) provide additional support to this interpretation.

Responses of low-BF VCN units to the same binaural pitch stimuli as in **Section 4.3 (page 110)** show no evidence of binaural integration (**Sections 5.3.4, page 153** and **5.3.6, page 157**). Taken with the rate-level function evidence, it is therefore unlikely that the binaural integration seen in the MSO and LSO (see **Chapter 4** (**page 101**) is influenced by crosstalk.

5.4.4 Summary

The findings of this chapter can be summarised as follows:

- No evidence for binaural integration of pitch was seen in response to dichotic cpHTCs in low-BF VCN units.
- Low-BF VCN units responded to only the ipsilateral carrier of DCTs.
- The first effect of pitch shift was observed in low-BF PL unit responses to frequency-shifted monaural complex tones.
- Synchrony is enhanced in guinea pig SBCs over their ANF inputs.
- Spike failure may strengthen the temporal precision in SBCs; further studies are needed to confirm this occurs throughout the low-BF population.
- Interaural crosstalk is unlikely to influence binaural responses seen in the SOC.

Chapter 6

General discussion

6.1 MAIN FINDINGS

The results presented in this Thesis are divided into three main sections. For each, I will describe and discuss the results, then outline potential avenues for future investigation.

6.1.1 Basic response properties of single units in the superior olivary complex

Chapter 3 (page 45) showed that the basic properties of single units of the guinea pig SOC are consistent with those observed in other species. SOC unit best frequency (BF) ranges, thresholds, receptive field bandwidths, spontaneous rates, responses to monaural and diotic pure tones, binaural beat responses and responses to BF-tones were consistent with those found in other species. The offset-only responses of the guinea pig superior paraolivary nucleus (SPN) and characteristic three-component spike waveform shape of the guinea pig medial nucleus of the trapezoid body (MNTB) are also observed in other mammals. The interaural level difference (ILD)-sensitivity of the LSO observed in other species was also seen in guinea pig LSO units. Finally, low-BF units of the MSO and LSO were shown to be interaural phase difference (IPD)-sensitive. These findings demonstrate the suitability of the guinea pig SOC as a model for studying binaural interactions, such as those underlying the responses in **Chapter 4 (page 101)**.

6.1.2 Binaural pitch

The results of **Chapter 4 (page 101)** showed that MSO and LSO cells respond to the fundamental frequency (F0) period of binaural complex tones. Interspike interval (ISI) analyses of MSO and LSO cell responses to frequency-shifted dichotic complex tones (DCTs) showed a neural correlate of the first-effect of pitch shift. This response was seen regardless of harmonic rank and stimulus configuration, *i.e.* if the lower harmonic was presented to the

ipsilateral ear, or *vice versa*. When presented with dichotic harmonic tone complexes (HTCs), with odd harmonics being presented to one ear and even to the other, MSO and LSO cells responded robustly to the overall F0 across ears. Similar responses were seen regardless of configuration, F0 and when randomising the phases of harmonics in the HTC. These results are the first to demonstrate a neural correlate of pitch integration of binaural tone complexes.

6.1.3 Responses to complex tones in the VCN input to the SOC

The results presented in **Chapter 5 (page 137)** showed that there is no evidence of binaural interaction in the responses of low-BF VCN units that form the inputs to the MSO and LSO. This confirms that the SOC is the first stage in the ascending auditory pathway that binaural integration of pitch occurs. Evidence of enhanced phase-locking to low-frequency pure tones in guinea pig primary-like (PL) units was also shown. Finally, by comparing the BF tone rate-level functions obtained from ipsilateral and contralateral presentation, I have shown that interaural crosstalk was unlikely to have played a role in the binaural stimulation paradigms used in **Chapters 3 (page 45)** and **4 (page 101)**.

Taken together, I have shown that the coding of temporal fine structure (TFS) information underlying periodicity pitch perception is enhanced in the inputs to the SOC and preserved in the output of the SOC. Cells of the MSO and LSO also respond similarly to dichotic and diotic stimuli that evoke binaural pitch perception in humans. The output of these cells to these stimuli, spike timing patterns locked to the overall stimulus F0 in harmonic stimuli or to pseudoperiods in inharmonic stimuli, is passed to cells of the IC and may underpin binaural pitch perception.

6.2 COMPARING RESPONSES WITH BINAURAL MODEL PREDICTIONS

The low-F0 dichotic HTCs used in this Thesis have unresolved harmonics. Determining the pitch of such signals requires some form of autocorrelation, not present in the modified equalisation-cancellation (m-EC) and central spectrum models (see **Section 1.4.3.1, page 27**). The responses of the MSO to unresolved HTCs shown in this Thesis likely result from binaural coincidence detection of the monaural, TFS-locked inputs arising from PL units. As such, it seems unlikely that MSO responses are accounted for by the m-EC or central spectrum models. However, the autocorrelation of binaural inputs in the triplex theory (Licklider, 1959) could account for the results seen in this Thesis.

The responses of the MSO to unresolved dichotic HTCs and DCTs show a combination of harmonics across ears and a response to the F0-period across ears. Such results are accounted for by the m-EC, central spectrum and optimum processor (see **Section 1.4.3.1, page 27**) models.

However, the weak response of the MSO to monaural-only stimulation precludes its role as a central pitch processor. It is likely that the actual central pitch processor receives direct monaural inputs in addition to inputs from the MSO. This could place it at the level of the inferior colliculus (IC).

6.3 FUTURE DIRECTIONS

6.3.1 SOC characterisation

The main experimental priority of this Thesis was to find MSO units and assess their temporal coding of binaural pitch-evoking stimuli. Unfortunately, this means there are limitations in the SOC characterisation shown in **Chapter 3 (page 45)**. A major issue is that no anatomical verification was carried out in recordings from nuclei other than the MSO and LSO; this was

done in order to maximise the experimental yield of MSO and LSO cells. Presumed MNTB and SPN units were identified by their similarity to characteristic responses seen in other species, however, there may be atypical types that were rejected as a result of the online classification scheme (see **Section 3.2.1, page 65**). In addition, responses of other periolivary nuclei were discarded. An obvious future direction is to carry out a basic response characterisation of the guinea pig SOC, similar to that of Guinan *et al.*, 1972a, 1972b in cat.

Binaural beat data were gathered for MSO and low-BF LSO units, however as no *in situ* interaural phase calibrations were performed, the data have not been used to assess the best interaural time difference (bITD) distributions and characteristic phase (CP)/characteristic delay (CD) distributions of the guinea pig MSO and LSO. Data from the guinea pig IC have previously been used as a proxy for the guinea pig MSO. For instance, frequency-dependency of ITD indicative of the cochlear delay model of the formation of bITD variations is not seen in the IC (McAlpine *et al.*, 2001). Having developed the skill to reliably find and record from guinea pig MSO units, a possible future direction would be to carry out a full investigation of the ITD properties of the guinea pig MSO and compare them with the guinea pig IC and the MSO of other species.

Finally, the frequency sampling of the receptive fields of MSO units in this Thesis was not fine enough to allow comparison of the BFs of ipsilateral and contralateral filters. Analysis of this could be used to assess the likelihood of cochlear delay (see **Section 1.2.5.1, page 9**) underpinning the ITD responses of MSO. Future studies of the guinea pig MSO could examine this by more precisely measuring frequency tuning curves, however there were units where monaural tonal stimulation did not elicit spikes above unit spontaneous rates, themselves close to 0 spikes/s. **Figure 36 (page 87)** shows examples of such units. A better method might be to reverse correlate MSO unit responses to dichotic uncorrelated broadband noise (BBN), as used by Sayles *et al.*, 2016 in chinchilla LL.

6.3.2 Effects of ITD and IPD on binaural pitch responses

All binaural pitch signals presented to the MSO and LSO were approximately matched in level between the two ears and close to zero ITD. The non-zero bITD distribution of low-frequency MSO and LSO cells means that these cells were likely not stimulated at their bITDs. However, there is evidence that ITD curves to Huggins pitch (HP) are demodulated when the phase transition of HP is centred on unit BF (see **Section 6.3.3, page 176**). Future studies could investigate whether this is the case for harmonic stimuli.

The pitch of DCTs has been shown to be robust across moderate ILD (Houtsma, 1981); an obvious next step would be to introduce ILDs to see if they affected the F0-locked responses seen in this Thesis.

6.3.3 Dichotic pitch

Having demonstrated that the MSO and LSO respond to dichotic HTCs and DCTs with the same neural correlate of pitch demonstrated in studies of the auditory nerve (AN), cochlear nucleus (CN) and IC, the most obvious question to ask is: do the MSO and LSO, which are IPD-sensitive, respond in the same manner to dichotic white-noise-based pitch (see **Section 1.4.3, page 25**)?

To date, the only published work looking at the neurophysiology of white-noise-based dichotic pitch is that undertaken by Hancock & Delgutte, 2000 in rabbit. This study used two variants of HP: HP+, where the interaural phase was π-rads except over a narrow frequency region, where it was 0-rads, and HP–, the opposite case (see **Figure 81, page 177**). Recordings were made from low-frequency neurons in the IC of cats of responses to the two categories of stimuli with multiple band frequencies. When the band frequency (F_b) was close to unit BF, units

responded with an increased rate for HP+ stimuli at the best ITD or with a decreased rate for HP– stimuli, and *vice versa*. ITD curves measured to HP with F_b near unit BF were unmodulated, *i.e.* the effect of ITD on the unit response was negligible in these cases. The authors drew comparison with studies showing that demodulation of ITD sensitivity occurs when broadband noise inputs are decorrelated between the ears (cat IC – Yin *et al.*, 1987). Whether the MSO underpins this response and its output is relayed to the IC or the response is generated in the IC remains an open question.

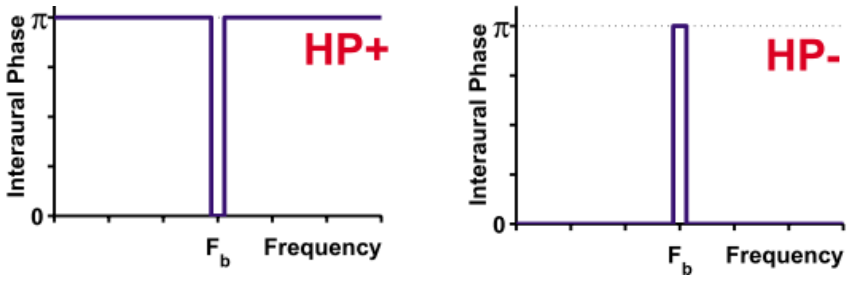

Figure 81 **HP+ and HP– as delineated in Hancock & Delgutte, 2000**. In these stimuli, the pitch perception occurs at F_b for both types; the smooth transition as is typically used, as in **Figure 9 (page 26)**, is not required.

When presented with BBN, unit responses show a dominant frequency ('DF') correlating with BF in ANFs (cat – Louage *et al.*, 2004), CN (cat – Louage *et al.*, 2005; cat – van der Heijden *et al.*, 2011) and IC (cat – Joris *et al.*, 2005). In other words, interspike interval distributions of responses to BBN show a peak around the period of the BF.

In response to the HP stimulus shown in **Figure 9 (page 26)**, the majority of MSO cells, with BFs away from the interaural phase transition region, would be responding strongly as the signals in each ear are identical. Over the interaural phase transition region, MSO cells would show varying responses: the relatively narrow transition region is likely to fall within an MSO

cell's receptive field. Some of these cells would show reduced responses due to the interaural decorrelation, however some cells might show increased responses due to the interaural phase at BF being close to the equivalent best ITD of the cell. Cells of the LSO respond maximally when stimuli are antiphasic between ears: LSO cells with BFs within the transition region would show a strong response, with the strongest response where IPD = π-rads, *i.e.* in the centre of the transition region, the frequency at which subjects perceive the pitch of these stimuli. However, the temporal responses of both MSO and LSO units are likely to be similar to that to BBN; MSO and LSO cells would likely not show a 'neural correlate of pitch' in response to HP.

In the central spectrum model, filtered DRP waveforms are approximately sinusoidal; at filters with centre frequencies that are integer multiples of the delay (F0-period), the phase-difference between these sinusoids is minimal. In these cells, the signals from each ear are delayed by an exact number of cycles and so reinforce. I would therefore expect MSO cells with BFs around integer multiples of F0 to show the greatest response in terms of spike rates. It is difficult to imagine a temporal neural correlate of pitch response manifesting to this stimulus as this would necessitate MSO unit bITDs of up to 32-ms for low-F0s, far outside the range found in studies of the MSO. However, MSO cells may respond monaurally with a few spikes locked to a specific stimulus feature; interspike intervals of such responses would show a peak at the delay, similar to responses of the CN to monaural iterated rippled noise (IRN) (guinea pig – Winter *et al.*, 2001; guinea pig – Sayles & Winter, 2008b).

In summary, I have presented the first responses of single units of the guinea pig SOC recorded *in vivo* and validated its use as a model for studying binaural interactions. I have demonstrated that binaural integration of pitch is not present in the VCN input to the SOC, and

that interaural crosstalk is unlikely to account for my results. Finally, I have shown a neural correlate of binaural integration of pitch to binaural complex tones in line with results from human psychophysics but previously not demonstrated in binaural neurons of the auditory system.

Having shown that cells of the guinea pig MSO respond to the period of the overall F0 of binaural complex tones, the race should be on for MSO researchers specialising in other mammals to confirm the same codes are used there as well. Additionally, a great many experiments await using dichotic white-noise-based pitch, in particular HP and DRP, to examine whether similar temporal encoding of binaural pitch in the MSO or IC occurs to these stimuli. I see the field continuing to have many avenues to investigate and enjoying a great deal of success.

Appendix

7.1 POPULATION SUMMARIES

LSO

An	Un	BF (kHz)	Thr (dB SPL)	SR (sp/s)	Lesion	ILD	Binbeat	Binbeat battery	RFs			ITD	Plock	Dichotic pitch		cpHTC		rdmHTC		DCT
									I	C	D			NB	WB	IoCe	IeCo	IoCe	IeCo	
1580	001	13.34	13	0.1					✓	✓	✓									
1598	013	0.25	40	45.4	**1**		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
1599	019	0.81	37	0.0	**2**	✓		✓	✓	✓										
1645	005	0.53	32	20.5			✓	✓	✓	✓										
1659	017	12.54	46	23.5					✓	✓	✓									
	018	11.17	32	6.1					✓	✓										
1660	003	1.15	25	11.8		✓			✓	✓	✓		✓							
	010	0.69	28	2.9	**3**		✓	✓	✓	✓	✓						✓			
1672	001	0.39	28	9.2	**4**	✓	✓	✓	✓	✓	✓		✓			✓	✓			
3010	022	6.41	32	0.6		✓			✓	✓	✓									
3021	003	4.34	12	0.0		✓			✓	✓	✓									
3036	005	0.22	47	17.0			✓	✓	✓	✓	✓									
3038	006	1.46	22	0.1					✓	✓	✓									
3042	003	0.22	33	7.6			✓	✓	✓	✓	✓		✓				✓			✓
	004	23.93	45	47.3		✓			✓	✓	✓									
	009	2.95	16	0.1				✓	✓	✓	✓									
	012	0.80	18	0.0		✓	✓	✓	✓	✓	✓					✓	✓			
	013	0.46	27	0.0				✓	✓	✓	✓		✓							
	020	0.26	39	1.6	**5**		✓		✓	✓	✓									
3053	008	2.20	23	0.1		✓			✓	✓	✓		✓							
	011	1.49	37	0.0			✓	✓	✓	✓	✓									
	015	0.22	39	0.7		✓	✓		✓	✓	✓		✓			✓				✓
3056	009	0.95	32	0.5		✓		✓	✓	✓	✓									
	013	0.25	28	41.5		✓	✓	✓	✓	✓	✓					✓	✓			
	014	0.33	21	14.8		✓	✓	✓	✓	✓	✓					✓	✓			✓
	015	0.73	22	0.3		✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓	✓
	022	0.26	29	1.6			✓		✓	✓	✓									✓

MSO

Animal	Unit	BF (kHz)	Thr (dB SPL)	SR (sp/s)	Holding time	Time to first MSO (hh:mm)	Lesion	Depth (µm)	Average spike amplitude (mV)	Binbeat at BF	Binbeat array	RFs			ITD	Phase locking	Dichotic pitch		cpHTC		frzHTC		rdmHTC		DCT
												I	C	D			NB	WB	loCe	leCo	loCe	leCo	loCe	leCo	
1589	016	0.48	50	7.9	01:40	13:29		5060.8	0.50	✓		✓	✓	✓	✓		✓	✓							
	018	0.75	34	6.4	01:00		**A**	5123.2	0.34	✓		✓	✓	✓	✓		✓		✓						
1623	002	0.70	59	3.5	01:00	03:20		4700.7	0.33	✓	✓	✓	✓	✓					✓						
	003	0.72	45	12.4	00:18		**C**	4769.8	0.27	✓		✓	✓	✓											
1633	001	0.13	55	7.5	04:04	05:07	**D**	5165.2	2.24	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓						
1636	022	0.32	48	6.9	00:47	20:21	**E**	5803.6	0.28	✓	✓	✓	✓	✓	✓				✓						
1654	006	0.56	26	3.6	00:56	11:56	**G**	4554.0	0.26	✓	✓	✓	✓	✓					✓						
1659	016	0.35	25	18.1	01:27	06:13	**H**	4760.7	0.67	✓	✓	✓	✓	✓					✓	✓					
1662	014	1.69	34	2.4	01:48	08:58	**I**	5619.6	0.54	✓	✓	✓	✓	✓			✓		✓	✓					
1667	001	0.46	34	7.9	01:24	02:40	**J**	5599.6	0.85	✓	✓	✓	✓	✓					✓	✓					
1675	006	0.59	33	19.4	00:31	08:41		4918.0	0.41	✓	✓	✓	✓	✓					✓						
	009	0.69	58	0.4	00:21		**K**	4378.0	0.31	✓	✓	✓	✓	✓											
3005	010	0.36	45	2.0	00:32	04:35		5860.1	0.48	✓	✓	✓	✓	✓					✓						
	011	0.17	52	23.8	01:00			5300.0	0.47	✓	✓	✓	✓	✓					✓						
	018	0.25	60	16.6	00:35		**L**	5434.9	0.60	✓	✓	✓	✓	✓											
3007	002	0.25	38	24.0	00:25	04:01		5930.0	0.29	✓	✓	✓		✓											
	014	0.74	24	0.1	01:32			5175.6	0.51		✓	✓	✓	✓	✓		✓		✓	✓					
	016	0.90	25	0.3	01:14		**blood track**	5214.1	0.42		✓	✓	✓	✓					✓	✓					
3010	010	0.14	47	3.3	00:39	08:51		5632.2	0.39	✓	✓			✓					✓	✓					
	012	0.14	46	3.6	01:13		**M**	5314.0	0.46	✓	✓	✓	✓	✓					✓	✓					
3021	005	0.96	28	0.2	00:51	03:45		5424.9	1.40	✓	✓	✓	✓	✓					✓	✓					
	008	0.39	25	0.8	03:38		**P**	5407.0	0.68	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓			
3022	005	0.96	18	6.2	05:59	03:42	**Q**	5791.0	2.10	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓			✓
3027	017	0.60	23	5.0	01:53	06:42		5400.0	0.31	✓	✓	✓	✓	✓		✓			✓	✓		✓			✓
	022	0.40	31	6.4	00:32			5153.2	0.32					✓							✓				
	023	0.36	28	1.0	00:21		**R**	4513.0	0.26	✓				✓							✓				
3038	014	0.41	28	14.3	03:13	05:55		5815.8	0.39	✓	✓	✓	✓	✓		✓			✓		✓	✓			✓
	018	0.31	24	11.7	02:07		**S**	5729.3	0.34	✓		✓	✓	✓		✓			✓		✓	✓			✓
3045	008	0.81	24	92.9	02:43	07:27		5358.0	0.97	✓	✓	✓	✓	✓		✓			✓		✓	✓			✓
	009	0.45	22	0.4	01:31			5407.0	1.36	✓	✓	✓	✓	✓					✓		✓	✓			✓
	011	0.60	27	2.6	00:31			5449.0	0.72		✓			✓					✓		✓				
	012	0.18	26	15.7	01:23			5538.0	2.07	✓		✓	✓	✓					✓		✓	✓			✓
	020	0.19	37	18.7	01:10		**T**	6043.2	0.55	✓		✓	✓	✓					✓		✓	✓			✓
3051	005	0.79	16	30.7	00:49	03:22		5384.0	1.41	✓		✓	✓	✓											✓
	010	0.39	36	0.7	00:22			5810.5	0.73												✓				
	011	1.14	14	1.4	02:25		**U**	5690.0	1.13	✓		✓	✓	✓					✓		✓	✓			✓
3052	006	0.61	20	4.4	00:22	05:48		5894.1	0.34	✓				✓											✓
	011	0.43	25	39.5	01:10		**V**	5328.0	0.29					✓					✓						✓
3053	012	0.24	31	9.8	00:51	07:02		4464.9	0.78	✓		✓	✓	✓					✓						✓
3055	009	0.62	13	17.0	00:32	04:31		5716.0	1.57	✓	✓	✓	✓	✓					✓						
	016	0.59	15	17.8	00:58			5096.6	0.42	✓	✓	✓	✓	✓					✓						✓
	020	0.26	28	5.7	01:05		**W**	5180.5	0.46	✓	✓	✓	✓	✓		✓			✓						
3056	027	1.39	17	0.7	00:39	16:10		6097.0	0.33	✓	✓	✓	✓	✓											
	028	0.90	25	0.2	00:54		**X**	5347.4			✓	✓	✓	✓					✓						
3057	003	0.85	31	0.0	00:42	02:47		5192.0	0.39		✓	✓	✓	✓					✓						
	012	0.56	28	1.0	00:42			4826.0	0.25	✓				✓					✓						✓
	013	0.63	26	4.6	00:23		**Y**	5127.3	0.20	✓				✓											✓

MSO, continued

Animal	Unit	BF (kHz)	Thr (dB SPL)	SR (sp/s)	Holding time	Time to first MSO (hh:mm)	Lesion	Depth (µm)	Average spike amplitude (mV)	Binbeat at BF	Binbeat array	RFs			ITD	Phase locking	Dichotic pitch		cpHTC		frzHTC		rdmHTC		DCT
												I	C	D			NB	WB	loCe	leCo	loCe	leCo	loCe	leCo	
3059	001	0.92	12	2.7	00:17	02:14		5390.2	0.33	✓				✓											
	007	2.00	32	1.6	00:35			5348.6	0.20			✓	✓	✓					✓						
	020	0.20	33	17.3	00:20		Z	5745.9		✓				✓											✓
3061	003	0.23	45	0.9	00:15	03:12		5488.9	0.45	✓				✓					✓						
	006	2.00	32	0.0	00:44			5036.8	0.45	✓		✓	✓	✓					✓	✓					
	008	0.74	14	11.3	02:37			5208.5	0.31	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓		✓	✓	✓
	009	0.25	30	9.3	01:03			5286.3	0.85	✓		✓	✓	✓				✓	✓				✓		
	012	0.17	40	49.6	01:23			5191.9	1.59	✓		✓	✓	✓		✓		✓	✓				✓		
	013	0.87	19	1.8	00:37			5177.0	0.24	✓				✓				✓							✓
	015	0.20	48	6.4	00:29		a	5261.0	0.26	✓				✓				✓							✓
3064	002	0.26	61	0.0	00:09	02:43		4998.1	0.27	✓				✓				✓							
	020	0.26	27	5.9	01:19			4906.0	0.35	✓		✓	✓	✓	✓			✓	✓	✓					
	021	1.46	27	0.0	02:00		b	5094.4	0.47	✓		✓	✓	✓		✓			✓	✓					✓

7.2 SOC COMPARISON TABLES

Table 1a **MNTB threshold distributions**

Species	Guinea pig	Cat	Gerbil
Range	19 to 55 dB-SPL	6 to 63 dB-SPL	
Mean±sd	36 ± 10 dB-SPL	21 dB-SPL	20 ± 5 dB-SPL (BF<6-kHz : 28 ± 8 dB-SPL)
Source	(this Thesis)	Smith *et al.*, 1998	Kopp-Scheinpflug *et al.*, 2003

Table 1b **MNTB $Q_{10\text{-}dB}$ distributions**

Species	Guinea pig	Cat	Gerbil	Rat (BFs>5-kHz)	Mouse (BFs>5-kHz)
Mean±sd	2.27 ± 1.45	9.20 ± 3.97	3.4 (BFs<6-kHz: 1.4)	7.5 ±0 .5	2.1 ± 0.2
Source	(this Thesis)	Tsuchitani, 1997	Kopp-Scheinpflug *et al.*, 2003	Kopp-Scheinpflug *et al.*, 2008	Kopp-Scheinpflug *et al.*, 2008

Table 1c **MNTB spontaneous rate distributions**

Species	Guinea pig	Cat	Gerbil	Rat	Mouse
Mean±sd	44.3 ± 24.5 spikes/s	27 spikes/s	30.0 ± 4.0 spikes/s	29.0 ± 2.5 spikes/s	32.0 ± 6.1 spikes/s
Source	(this Thesis)	Smith *et al.*, 1998	Kopp-Scheinpflug *et al.*, 2003	Kopp-Scheinpflug *et al.*, 2008	Kopp-Scheinpflug *et al.*, 2008

Table 2a **SPN threshold distributions**

Species	Guinea pig	Rat	Gerbil (offset-only)
Range	19 to 71 dB-SPL		20 to 70 dB-SPL
Mean±sd	47 ± 14 dB-SPL	25 ± 12 dB-SPL	
Source	(this study)	Kulesza *et al.*, 2003	Dehmel *et al.*, 2002

Table 2b **SPN $Q_{10\text{-}dB}$ distributions**

Species	Guinea pig	Gerbil	Rat
Mean±sd	2.27 ± 1.45	3.0 ± 2.9	6.77 ± 3.30
Source	(this study)	Behrend *et al.*, 2002	Kulesza *et al.*, 2003

Table 2c **SPN spontaneous rate distributions**

Species	Guinea pig	Gerbil	Rat
Mean±sd	0.4 ± 1.0 spikes/s	61% <10 spikes/s	'overwhelming majority' <6 spikes/s
Source	(this study)	Behrend *et al.*, 2002	Kulesza *et al.*, 2003

Table 3a **MSO $Q_{10\text{-}dB}$ distributions**

Species	Guinea pig	Chinchilla LL	Mouse
Mean±sd	1.18 ± 0.41	1.38	1.8
Source	(this Thesis)	Bremen & Joris, 2013	Fischl *et al.*, 2016

Table 3b **MSO first spike latency distributions**

Species	Guinea pig	Chinchilla LL	Gerbil
Mean±sd	5.9 ± 1.5-ms	6-ms	5.1 ± 0.2-ms
Source	(this Thesis)	Bremen & Joris, 2013	Spitzer & Semple, 1995

References

Abel, S.M. (1972) Duration discrimination of noise and tone bursts. *J Acoust Soc Am*, **51**, 1219–1223.

Adams, J.C. & Mugnaini, E. (1990) Immunocytochemical evidence for inhibitory and disinhibitory circuits in the superior olive. *Hear Res*, **49**, 281–298.

Agmon-Snir, H., Carr, C.E., & Rinzel, J. (1998) The role of dendrites in auditory coincidence detection. *Nature*, **393**, 268–272.

Arthur, R.M., Pfeiffer, R.R., & Suga, N. (1971) Properties of “Two-Tone Inhibition” in Primary Auditory Neurones. *J Physiol*, **212**, 593–609.

Babalian, A.L. (2008) Synaptic responses relevant for binaural interactions in the medial superior olive studied in the isolated whole brain preparation of the guinea pig. *FENS Abstr*, **4**, 1.

Babalian, A.L., Jacomme, A.-V., Doucet, J.R., Ryugo, D.K., & Rouiller, E.M. (2002) Commissural glycinergic inhibition of bushy and stellate cells in the anteroventral cochlear nucleus. *Neuroreport*, **13**, 555–558.

Babalian, A.L., Ryugo, D.K., Vischer, M.W., & Rouiller, E.M. (1999) Inhibitory synaptic interactions between cochlear nuclei: evidence from an in vitro whole brain study. *Neuroreport*, **10**, 1913–1917.

Bandyopadhyay, S. & Young, E.D. (2013) Nonlinear temporal receptive fields of neurons in the dorsal cochlear nucleus. *J Neurophysiol*, **110**, 2414–2425.

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

Batra, R., Kuwada, S., Fitzpatrick, D.C., Wang, L., Devore, S., Delgutte, B., & Colburn, H.S. (1997a) 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.

Batra, R., Kuwada, S., Fitzpatrick, D.C., Wang, L., Devore, S., Delgutte, B., & Colburn, H.S. (1997b) Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. coincidence detection. *J Neurophysiol*, **78**, 1222–1236.

Beckius, G.E., Batra, R., & Oliver, D.L. (1999) Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci*, **19**, 3146–3161.

Beerends, J.G. (1989) Pitches of simultaneous complex tones.

Beerends, J.G. & Houtsma, A.J.M. (1988) Pitch identification of simultaneous diotic and dichotic two-tone complexes. *J Acoust Soc Am*, **85**, 813–819.

Behrend, O., Brand, A., Kapfer, C., & Grothe, B. (2002) Auditory Response Properties in the Superior Paraolivary Nucleus of the Gerbil Auditory. *J Neurophysiol*, **87**, 2915–2928.

Békésy, G. V (1953) Description of Some Mechanical Properties of the Organ of Corti. *J Acoust Soc Am*, **25**, 770–785.

Benson, C.G. & Potashner, S.J. (1990) Retrograde transport of [3H]glycine from the cochlear nucleus to the superior olive in the guinea pig. *J Comp Neurol*, **296**, 415–426.

Bergsman, J.B., De Camilli, P., & McCormick, D.A. (2004) Multiple Large Inputs to Principal Cells in the Mouse Medial Nucleus of the Trapezoid Body. *J Neurophysiol*, 545–552.

Bernstein, J.G. & Oxenham, A.J. (2003) Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number? *J Acoust Soc Am*, **113**, 3323–3334.

Bilsen, F.A. (1972) Pitch of Dichotically Delayed Noise. In *Symposium on Hearing Theory 1972*. pp. 5–8.

Bilsen, F.A. (1976) Pronounced binaural pitch phenomenon. *J Acoust Soc Am*, **59**, 467–468.

Bilsen, F.A. (1977) Pitch of noise signals: Evidence for a “central spectrum.” *J Acoust Soc Am*, **61**, 150–161.

Bilsen, F.A. & Goldstein, J.L. (1974) Pitch of dichotically delayed noise and its possible spectral basis. *J Acoust Soc Am*, **55**, 292–296.

Bilsen, F.A. & Ritsma, R.J. (1969) Repetition Pitch and Its Implication for Hearing Theory. *Acustica*, **22**, 63–73.

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

Bledsoe, S.C., Snead, C.R., Helfert, R.H., Prasad, V., Wenthold, R.J., & Altschuler, R.A. (1990) Immunocytochemical and lesion studies support the hypothesis that the projection from the medial nucleus of the trapezoid body to the lateral superior olive is glycinergic. *Brain Res*, **517**, 189–194.

Bourk, T.R. (1976) Electrical responses of neural units in the anteroventral cochlear nucleus of the cat.

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

Brawer, J.R. & Morest, D.K. (1975) 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.

Bremen, P. & Joris, P.X. (2013) Axonal Recordings from Medial Superior Olive Neurons Obtained from the Lateral Lemniscus of the Chinchilla (Chinchilla laniger). *J Neurosci*, **33**, 17506–17518.

Brown, M.C. (1987) Morphology of labeled afferent fibers in the guinea pig cochlea. *J Comp Neurol*, **260**, 591–604.

Brown, M.C. (2011) Anatomy of Olivocochlear Neurons. In *Auditory and Vestibular Efferents*. pp. 17–37.

Brownell, W.E., Bader, C.R., Bertrand, D., & Ribaupierre, Y.D. (1985) Evoked Mechanical

Responses of Isolated Cochlear Outer Hair Cells. *Science (80-)*, **227**, 194–196.

Buunen, T.J.F. & Rhode, W.S. (1978) Response of fibers in the cat's auditory nerve to the cubic difference tone. *J Acoust Soc Am*, **64**, 772–781.

Caird, D., Göttl, K.H., & Klinke, R. (1980) Interaural attenuation in the cat, measured with single fibre data. *Hear Res*, **3**, 257–263.

Cant, N.B. (1981) The fine structure of two types of stellate cells in the anterior division of the anteroventral cochlear nucleus of the cat. *Neuroscience*, **6**, 2643–2655.

Cant, N.B. (1984) The fine structure of the lateral superior olivary nucleus of the cat. *J Comp Neurol*, **227**, 63–77.

Cant, N.B. & Benson, C.G. (2003) Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. *Brain Res Bull*, **60**, 457–474.

Cant, N.B. & Hyson, R.L. (1992) Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil. *Hear Res*, **58**, 26–34.

Cariani, P.A. & Delgutte, B. (1996a) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol*, **76**, 1698–1716.

Cariani, P.A. & Delgutte, B. (1996b) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J Neurophysiol*, **76**, 1717–1734.

Carlyon, R.P. & Shackleton, T.M. (1994) Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms? *J Acoust Soc Am*, **95**, 3541–3554.

Carr, C.E. & Konishi, M. (1988) Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci U S A*, **85**, 8311–8315.

Carr, C.E. & Konishi, M. (1990) A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci*, **10**, 3227–3246.

Colburn, H.S., Yan-an, H., & Culotta, C.P. (1990) Coincidence model of MSO responses. *Hear Res*, **49**, 335–346.

Cooper, N.P. & Guinan, J.J. (2006) Efferent-mediated control of basilar membrane motion. *J Physiol*, **576**, 49–54.

Cosentino, S., Marquardt, R., McAlpine, D., Culling, J.F., & Falk, T.H. (2014) A model that predicts the binaural advantage to speech intelligibility from the mixed target and interferer signals. *J Acoust Soc Am*, **135**, 796–807.

Counter, S.A. & Borg, E. (1993) Acoustic Middle Ear Muscle Reflex Protection against Magnetic Coil Impulse Noise. *Acta Otolaryngol*, **113**, 483–488.

Cox, R.M., DeChicchis, A.R., & Wark, D.J. (1981) Demonstration of binaural advantage in audiometric test rooms. *Ear Hear*, **2**, 194–201.

Cramer, E.M. & Huggins, W.H. (1958) Creation of pitch through binaural interaction. *J Acoust Soc Am*, **30**, 413–417.

Creelman, C.D. (1962) Human Discrimination of Auditory Duration. *J Acoust Soc Am*, **34**, 582.

Culling, J.F. (2000) Dichotic pitches as illusions of binaural unmasking. III. The existence region of the Fourcin pitch. *J Acoust Soc Am*, **107**, 2201–2208.

Culling, J.F., Marshall, D.H., & Summerfield, A.Q. (1998) Dichotic pitches as illusions of binaural unmasking. II. The Fourcin pitch and the dichotic repetition pitch. *J Acoust Soc Am*, **103**, 3527–3539.

Culling, J.F., Summerfield, A.Q., & Marshall, D.H. (1998) Dichotic pitches as illusions of binaural unmasking. I. Huggins' pitch and the "binaural edge pitch." *J Acoust Soc Am*, **103**, 3509–3526.

Culling, J.F. & Summerfield, Q. (1995) Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *J Acoust Soc Am*, **98**, 785.

Dallos, P. (1986) Neurobiology of cochlear inner and outer hair cells: intracellular recordings. *Hear Res*, **22**, 185–198.

Darrow, K.N., Maison, S.F., & Liberman, M.C. (2006) Cochlear efferent feedback balances interaural sensitivity. *Nat Neurosci*, **9**, 1474–1476.

Darrow, K.N., Maison, S.F., & Liberman, M.C. (2007) Selective removal of lateral olivocochlear efferents increases vulnerability to acute acoustic injury. *J Neurophysiol*, **97**, 1775–1785.

Dau, T., Püschel, D., & Kohlrausch, A. (1996) A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am*, **99**, 3615–3622.

Day, M.L. & Semple, M.N. (2011) Frequency-dependent interaural delays in the medial superior olive: implications for interaural cochlear delays. *J Neurophysiol*, **106**, 1985–1999.

de Boer, E. (1956a) Pitch of Inharmonic Signals. *Nature*, **178**, 535–536.

de Boer, E. (1956b) On the "Residue" in Hearing.

de Boer, E. (1976) On the "Residue" and Auditory Pitch Perception. In *Handbook of Sensory Physiology Volume V/3: Auditory System - Clinical and Special Topics*. pp. 479–584.

de Cheveigné, A. (1998) Cancellation model of pitch perception. *J Acoust Soc Am*, **103**, 1261–1271.

de Cheveigné, A. (2004) Pitch Perception Models. In *Pitch*. Springer, pp. 169–233.

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

Dehmel, S., Kopp-Scheinpflug, C., & Rübsamen, R. (2005) Interplay of excitation and inhibition in auditory brainstem processing at endbulbs of Held of the MNTB and AVCN. In *Plasticity of the Central Auditory System and Processing of Complex Acoustic Signals.* Springer Science and Business Media Inc., New York, pp. 15–36.

Dehmel, S., Kopp-Scheinpflug, C., Weick, M., Dörrscheidt, G.J., & Rübsamen, R. (2010) Transmission of phase-coupling accuracy from the auditory nerve to spherical bushy cells

in the Mongolian gerbil. *Hear Res*, **268**, 234–249.

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

Durlach, N.I. (1962) Note on the Creation of Pitch through Binaural Interaction. *J Acoust Soc Am*, **34**, 1096–1099.

Durlach, N.I. (1963) Equalization and Cancellation Theory of Binaural Masking-Level Differences. *J Acoust Soc Am*, **35**, 1206–1218.

Durlach, N.I. (1972) Binaural signal detection - Equalization and cancellation theory. In *Foundations of Modern Auditory Theory Vol. II*.

Durlach, N.I. & Colburn, H.S. (1978) Binaural Phenomena. In *Handbook of Perception Volume IV - Hearing*. pp. 365–466.

Evans, E.F. (1972) The Frequency Response and Other Properties of Single Fibres in the Guinea-Pig Cochlear Nerve. *J Physiol*, **226**, 263–287.

Evans, E.F. (1978) Place and Time Coding of Frequency in the Peripheral Auditory System: Some Physiological Pros and Cons. *Audiology*, **17**, 369–420.

Fekete, D.M., Rouiller, E.M., Liberman, M.C., & Ryugo, D.K. (1982) The central projections of intracellularly labeled auditory nerve fibers in cats. *J Comp Neurol*, **229**, 432–450.

Felix, R.A., Fridberger, A., Leijon, S., Berrebi, A.S., & Magnusson, A.K. (2011) Sound rhythms are encoded by postinhibitory rebound spiking in the superior paraolivary nucleus. *J Neurosci*, **31**, 12566–12578.

Felix, R.A., Magnusson, A.K., & Berrebi, A.S. (2014) The superior paraolivary nucleus shapes temporal response properties of neurons in the inferior colliculus. *Brain Struct Funct*,.

Felix II, R.A., Gourévitch, B., Gómez-Álvarez, M., Leijon, S.C.M., Saldaña, E., & Magnusson, A.K. (2017) Octopus Cells in the Posteroventral Cochlear Nucleus Provide the Main Excitatory Input to the Superior Paraolivary Nucleus. *Front Neural Circuits*, **11**.

Feng, L. & Wang, X. (2017) Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc Natl Acad Sci*,.

Fettiplace, R. & Hackney, C.M. (2006) The sensory and motor roles of auditory hair cells. *Nat Rev Neurosci*, **7**, 19–29.

Fischl, M.J., Burger, R.M., Schmidt-Pauly, M., Alexandrova, O., Sinclair, J.L., Grothe, B., Forsythe, I.D., & Kopp-Scheinpflug, C. (2016) Physiology and Anatomy of Neurons in the Medial Superior Olive (MSO) of the Mouse. *J Neurophysiol*, jn.00523.2016.

Fitzpatrick, D.C. & Kuwada, S. (2001) Tuning to Interaural Time Differences across Frequency. *J Neurosci*, **21**, 4844–4851.

Forsythe, I.D. & Barnes-Davies, M. (1993) The binaural auditory pathway: membrane currents limiting multiple action potential generation in the rat medial nucleus of the trapezoid body. *Proc R Soc B Biol Sci*, **251**, 143–150.

Franken, T.P., Roberts, M.T., Wei, L., Golding, N.L., & Joris, P.X. (2015) In vivo coincidence

detection in mammalian sound localization generates phase delays. *Nat Neurosci*, 1–11.

Frisina, R.D., Smith, R.L., & Chamberlain, S.C. (1990) Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hear Res*, **44**, 99–122.

Gibson, D.J. (1982) Interaural crosstalk in the cat. *Hear Res*, **7**, 325–333.

Glendenning, K.K., Baker, B.N., Hutson, K.A., & Masterton, R.B. (1992) Acoustic chiasm V: inhibition and excitation in the ipsilateral and contralateral projections of LSO. *J Comp Neurol*, **319**, 100–122.

Glendenning, K.K., Hutson, K.A., Nudo, R.J., & Masterton, R.B. (1985) Acoustic chiasm II: Anatomical basis of binaurality in lateral superior olive of cat. *J Comp Neurol*, **232**, 261–285.

Glendenning, K.K. & Masterton, R.B. (1983) Acoustic chiasm: efferent projections of the lateral superior olive. *J Neurosci*, **3**, 1521–1537.

Glueckert, R., Pfaller, K., Kinnefors, A., Rask-Andersen, H., & Schrott-Fischer, A. (2005) Ultrastructure of the normal human organ of Corti. New anatomical findings in surgical specimens. *Acta Otolaryngol*, **125**, 534–539.

Gockel, H.E., Carlyon, R.P., & Plack, C.J. (2004) Across-frequency interference effects in fundamental frequency discrimination: Questioning evidence for two pitch mechanisms. *J Acoust Soc Am*, **116**, 1092–1104.

Gockel, H.E., Carlyon, R.P., & Plack, C.J. (2009) Pitch discrimination interference between binaural and monaural or diotic pitches. *J Acoust Soc Am*, **126**, 281–290.

Gockel, H.E., Carlyon, R.P., & Plack, C.J. (2011) Combination of spectral and binaurally created harmonics in a common central pitch processor. *J Assoc Res Otolaryngol*, **12**, 253–260.

Gockel, H.E., Hafter, E.R., & Moore, B.C.J. (2009) Pitch discrimination interference: the role of ear of entry and of octave similarity. *J Acoust Soc Am*, **125**, 324–327.

Goldberg, J.M. & Brown, P.B. (1968) Functional Organization of the Dog Superior Olivary Complex: An Anatomical and Electrophysiological Study. *J Neurophysiol*, **31**, 639–656.

Goldberg, J.M. & Brown, P.B. (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol*, **32**, 613–636.

Goldstein, J.L. (1973) An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am*, **54**, 1496–1516.

Greene, N.T., Anbuhl, K.L., Williams, W., & Tollin, D.J. (2014) The acoustical cues to sound location in the guinea pig (Cavia porcellus). *Hear Res*, **316C**, 1–15.

Groff, J.A. & Liberman, M.C. (2003) Modulation of cochlear afferent response by the lateral olivocochlear system: activation via electrical stimulation of the inferior colliculus. *J Neurophysiol*, **90**, 3178–3200.

Grothe, B. (2000) The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Prog Neurobiol*, **61**, 581–610.

Grothe, B. & Neuweiler, G. (2000) The function of the medial superior olive in small mammals: temporal receptive fields in auditory analysis. *J Comp Physiol A*, **186**, 413–423.

Grothe, B. & Park, T.J. (1998) Sensitivity to interaural time differences in the medial superior olive of a small mammal, the Mexican free-tailed bat. *J Neurosci*, **18**, 6608–6622.

Grothe, B. & Park, T.J. (2000) Structure and Function of the Bat Superior Olivary Complex. *Microsc Res Tech*, **402**, 382–402.

Guinan, J.J. & Li, R.Y.S. (1990) Signal processing in brainstem auditory neurons which receive giant endings (calyces of Held) in the medial nucleus of the trapezoid body of the cat. *Hear Res*, **49**, 321–334.

Guinan, J.J., Norris, B.E., & Guinan, S.S. (1972a) Single Auditory Units in the Superior Olivary Complex: II: Locations of Unit Categories and Tonotopic Organisation. *Int J Neurosci*, 147–166.

Guinan, J.J., Norris, B.E., & Guinan, S.S. (1972b) Single Auditory Units in the Superior Olivary Complex: I: Responses to Sound and Classifications Based on Physiological Properties. *Int J Neurosci*,.

Guinan, J.J. & Peake, W.T. (1967) Middle-Ear Characteristics of Anesthetized Cats. *J Acoust Soc Am*, **41**, 1237–1261.

Guinan, J.J., Warr, W.B., & Norris, B.E. (1983) Differential Olivocochlear Projections From Lateral Versus Medial Zones of the Superior Olivary Complex. *J Comp Neurol*, **221**, 359–370.

Hackney, C.M., Osen, K.K., & Kolston, J. (1990) Anatomy of the cochlear nuclear complex of guinea pig. *Anat Embryol (Berl)*, **182**, 123–149.

Hall, J.L. (1965) Binaural Interaction in the Accessory Superior-Olivary Nucleus of the Cat. *J Acoust Soc Am*, **37**, 814–823.

Hancock, K.E. & Delgutte, B. (2000) Neural Correlates of the Huggins Dichotic Pitch. In *Association for Research in Otolaryngology Abstracts*.

Harrison, J.M. & Irving, R. (1965) The Anterior Ventral Cochlear Nucleus. *J Comp Neurol*, **124**, 15–41.

Hartmann, W.M. (1996) Pitch, periodicity, and auditory organization. *J Acoust Soc Am*, **100**, 3491–3502.

Hartmann, W.M. & Zhang, P.X. (2003) Binaural models and the strength of dichotic pitches. *J Acoust Soc Am*, **114**, 3317–3326.

Held, H. (1893) Die centrale Gehörleitung. *Arch für Anat und Physiol*, **17**, 201–248.

Helfert, R.H., Bonneau, J.M., Wenthold, R.J., & Altschuler, R.A. (1989) GABA and glycine immunoreactivity in the guinea pig superior olivary complex. *Brain Res*, **501**, 269–286.

Helfert, R.H. & Schwartz, I.R. (1986) Morphological evidence for the existence of multiple neuronal classes in the cat lateral superior olivary nucleus. *J Comp Neurol*, **244**, 533–549.

Helfert, R.H. & Schwartz, I.R. (1987) Morphological features of five neuronal classes in the

gerbil lateral superior olive. *Am J Anat*, **179**, 55–69.

Hermes, D.J. (1988) Measurement of pitch by subharmonic summation. *J Acoust Soc Am*, **83**, 257–264.

Hewitt, M.J., Meddis, R., & Shackleton, T.M. (1991) A computer model of a cochlear-nucleus stellate cell: Responses to amplitude-modulated and pure-tone stimuli. *J Acoust Soc Am*, **91**, 2096–2109.

Houtsma, A.J.M. (1981) Pitch of unequal-amplitude dichotic two-tone harmonic complexes. *J Acoust Soc Am*, **69**, 1778–1785.

Houtsma, A.J.M. & Goldstein, J.L. (1971) Perception of Musical Intervals: Evidence for the Central Origin of the Pitch of Complex Tones, MIT Research Laboratory of Electronics.

Houtsma, A.J.M. & Goldstein, J.L. (1972) The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition. *J Acoust Soc Am*, **51**, 520.

Ingham, N.J., Bleeck, S., & Winter, I.M. (2006) Contralateral inhibitory and excitatory frequency response maps in the mammalian cochlear nucleus. *Eur J Neurosci*, **24**, 2515–2529.

Jalabi, W., Kopp-Scheinpflug, C., Allen, P.D., Schiavon, E., DiGiacomo, R.R., Forsythe, I.D., & Maricich, S.M. (2013) Sound localization ability and glycinergic innervation of the superior olivary complex persist after genetic deletion of the medial nucleus of the trapezoid body. *J Neurosci*, **33**, 15044–15049.

Jeffress, L.A. (1948) A place theory of sound localization. *J Comp Physiol Psychol*, **41**, 35–39.

Jeffress, L.A. (1958) Medial Geniculate Body – A Disavowal. *J Acoust Soc Am*, **30**, 802–803.

Joris, P.X. (1996) Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J Neurophysiol*, **76**, 2137–2156.

Joris, P.X., Carney, L.H., Smith, P.H., & Yin, T.C. (1994a) Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the Tuning Curve Tail. *J Neurophysiol*, **71**, 1037–1051.

Joris, P.X., Carney, L.H., Smith, P.H., & Yin, T.C.T. (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol*, **71**, 1022–1036.

Joris, P.X., de Sande, B. V, Louage, D.H., & van der Heijden, M. (2006) Binaural and cochlear disparities. *Proc Natl Acad Sci U S A*, **103**, 12917–12922.

Joris, P.X., de Sande, B. V, & van der Heijden, M. (2005) Temporal Damping in Response to Broadband Noise. I. Inferior Colliculus. *J Neurophysiol*, **93**, 1857–1870.

Joris, P.X. & Smith, P.H. (2008) The volley theory and the spherical cell puzzle. *Neuroscience*, **154**, 65–76.

Joris, P.X. & van der Heijden, M. (2004) Temporal synchronization in the auditory periphery of macaque monkeys. *Soc Neurosci Abstr*, 650.16.

Joris, P.X. & Yin, T.C.T. (1995) Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J Neurophysiol*, **73**, 1043–1062.

Joris, P.X. & Yin, T.C.T. (1998) Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J Neurophysiol*, **79**, 253–269.

Kadner, A. & Berrebi, A.S. (2008) Encoding of temporal features of auditory stimuli in the medial nucleus of the trapezoid body and superior paraolivary nucleus of the rat. *Neuroscience*, **151**, 868–887.

Kadner, A., Kulesza, R.J., & Berrebi, A.S. (2006) Neurons in the Medial Nucleus of the Trapezoid Body and Superior Paraolivary Nucleus of the Rat May Play a Role in Sound Duration Coding. *J Neurophysiol*, 1499–1508.

Keine, C., Rübsamen, R., & Englitz, B. (2016) Inhibition in the auditory brainstem enhances signal representation and regulates gain in complex acoustic environments. *Elife*, **5**, 1–33.

Keys, J.W. (1947) Binaural Versus Monaural Hearing. *J Acoust Soc Am*, **19**, 629–631.

Kiang, N.Y.S., Pfeiffer, R.R., Warr, W.B., & Backus, A.S.N. (1965) Stimulus coding in the cochlear nucleus. *Ann Otol Rhinol Laryngol*, **74**, 463–485.

Kil, J., Hkageyama, G., Semple, M.N., & Kitzes, L.M. (1995) Development of ventral cochlear nucleus projections to the superior olivary complex in gerbil. *J Comp Neurol*, **353**, 317–340.

Kopp-Scheinpflug, C., Dehmel, S., Dörrscheidt, G.J., & Rübsamen, R. (2002) Interaction of excitation and inhibition in anteroventral cochlear nucleus neurons that receive large endbulb synaptic endings. *J Neurosci*, **22**, 11004–11018.

Kopp-Scheinpflug, C., Lippe, W.R., Dörrscheidt, G.J., & Rübsamen, R. (2003) The medial nucleus of the trapezoid body in the gerbil is more than a relay: Comparison of pre- and postsynaptic activity. *J Assoc Res Otolaryngol*, **4**, 1–23.

Kopp-Scheinpflug, C., Tolnai, S., Malmierca, M.S., & Rübsamen, R. (2008) The medial nucleus of the trapezoid body: comparative physiology. *Neuroscience*, **154**, 160–170.

Kopp-Scheinpflug, C., Tozer, A.J.B., Robinson, S.W., Tempel, B.L., Hennig, M.H., & Forsythe, I.D. (2011) The sound of silence: ionic mechanisms encoding sound termination. *Neuron*, **71**, 911–925.

Kuenzel, T., Borst, J.G.G., & van der Heijden, M. (2011) Factors Controlling the Input–Output Relationship of Spherical Bushy Cells in the Gerbil Cochlear Nucleus. *J Neurosci*, **31**, 4260–4273.

Kuhn, G.F. (1977) Model for the interaural time differences in the azimuthal plane. *J Acoust Soc Am*, **62**, 157–167.

Kulesza, R.J. & Berrebi, A.S. (2000) The Superior Paraolivary Nucleus of the Rat Is a GABAergic Nucleus. *J Assoc Res Otolaryngol*, **1**, 255–269.

Kulesza, R.J. & Grothe, B. (2015) Yes, there is a medial nucleus of the trapezoid body in humans. *Front Neuroanat*, **9**, 1–9.

Kulesza, R.J., Kadner, A., & Berrebi, A.S. (2007) Distinct roles for glycine and GABA in shaping the response properties of neurons in the superior paraolivary nucleus of the rat. *J Neurophysiol*, **97**, 1610–1620.

Kulesza, R.J., Spirou, G.A., & Berrebi, A.S. (2003) Physiological response properties of neurons in the superior paraolivary nucleus of the rat. *J Neurophysiol*, **89**, 2299–2312.

Kuwabara, N. & Zook, J.M. (1991) Classification of the principal cells of the medial nucleus of the trapezoid body. *J Comp Neurol*, **314**, 707–720.

Kuwabara, N. & Zook, J.M. (1999) Local collateral projections from the medial superior olive to the superior paraolivary nucleus in the gerbil. *Brain Res*, **846**, 59–71.

Kuwada, S. & Yin, T.C.T. (2012) The Superior Olivary Complex. *Transl Perspect Audit Neurosci Norm Asp Hear*, 161–184.

Langford, T.L. (1984) Responses elicited from medial superior olivary neurons by stimuli associated with binaural masking and unmasking. *Hear Res*, **15**, 39–50.

Langner, G. (1981) Neuronal mechanisms for pitch analysis in the time domain. *Exp Brain Res*, **44**, 450–454.

Langner, G. (1983) Monaural Phase Effects in Masking with Multicomponent Signals. In *HEARING - Physiological Bases and Psychophysics*. pp. 334–341.

Langner, G. (2015) *The Neural Code of Pitch and Harmony*.

Langner, G. & Schreiner, C.E. (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol*, **60**, 1799–1822.

Li, L. & Kelly, J.B. (1992) Binaural responses in rat inferior colliculus following kainic acid lesions of the superior olive: Interaural intensity difference functions. *Hear Res*, **61**, 73–85.

Liberman, M.C. (1978) Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am*, **63**, 442–455.

Liberman, M.C. (1991) Central Projections of Auditory-Nerve Fibers of Differing Spontaneous Rate. I. Anteroventral Cochlear Nucleus. *J Comp Neurol*, **313**, 240–258.

Liberman, M.C. (1993) Central projections of auditory nerve fibers of differing spontaneous rate, II. Posteroventral and dorsal cochlear nuclei. *J Comp Neurol*, **327**, 17–36.

Liberman, M.C., Dodds, L.W., & Pierce, S. (1990) Afferent and Efferent Innervation of the Cat Cochlea: Quantitative Analysis With Light and Electron Microscopy. *J Comp Neurol*, **301**, 443–460.

Licklider, J.C.R. (1951) A Duplex Thory of Pitch Perception. *Experientia*, **7**, 128–134.

Licklider, J.C.R. (1954) “Periodicity” pitch and “place” pitch. *J Acoust Soc Am*, **26**, 945.

Licklider, J.C.R. (1959) Three auditory theories. In *Psychology: A Study of a Science 1*. pp. 41–144.

Liu, L.-F., Palmer, A.R., & Wallace, M.N. (2006) Phase-locked responses to pure tones in the inferior colliculus. *J Neurophysiol*, **95**, 1926–1935.

Lord Rayleigh (1907) XII. On our Perception of Sound Direction. *Philos Mag Ser 6*, **13**, 214–232.

Lorente de Nó, R. (1947) Action Potential of the Motoneurons of the Hypoglossus Nucleus. *J Cell Comp Physiol*,.

Lorteije, J.A.M. & Borst, J.G.G. (2011) Contribution of the mouse calyx of Held synapse to tone adaptation. *Eur J Neurosci*, **33**, 251–258.

Lorteije, J.A.M., Rusu, S.I., Kushmerick, C., & Borst, J.G.G. (2009) Reliability and precision of the mouse calyx of Held synapse. *J Neurosci*, **29**, 13770–13784.

Louage, D.H.G., van der Heijden, M., & Joris, P.X. (2004) Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol*, **91**, 2051–2065.

Louage, D.H.G., van der Heijden, M., & Joris, P.X. (2005) Enhanced Temporal Response Properties of Anteroventral Cochlear Nucleus Neurons to Broadband Noise. *J Neurosci*, **25**, 1560–1570.

Luparello, T.J., Stein, M., & Park, C.D. (1964) A Stereotaxic Atlas of the Hypothalamus of the Guinea Pig. *J Comp Neurol*, **122**, 201–217.

Magnusson, A.K., Park, T.J., Pecka, M., Grothe, B., & Koch, U. (2008) Retrograde GABA Signaling Adjusts Sound Localization by Balancing Excitation and Inhibition in the Brainstem. *Neuron*, **59**, 125–137.

Maison, S.F. & Liberman, M.C. (2000) Predicting Vulnerability to Acoustic Injury with a Noninvasive Assay of Olivocochlear Reflex Strength. *J Neurosci*, **20**, 4701–4707.

Manis, P.B. & Marx, S.O. (1991) Outward currents in isolated ventral cochlear nucleus neurons. *J Neurosci*, **11**, 2865–2880.

Matho, K.S.H. (2013) Structure and dynamism in a central neural circuit in adulthood and postnatal development.

Mc Laughlin, M., van der Heijden, M., & Joris, P.X. (2008) How secure is in vivo synaptic transmission at the calyx of Held? *J Neurosci*, **28**, 10206–10219.

Mc Laughlin, M., Verschooten, E., & Joris, P.X. (2010) Oscillatory dipoles as a source of phase shifts in field potentials in the mammalian auditory brainstem. *J Neurosci*, **30**, 13472–13487.

McAlpine, D. (2004) Neural sensitivity to periodicity in the inferior colliculus: evidence for the role of cochlear distortions. *J Neurophysiol*, **92**, 1295–1311.

McAlpine, D., Jiang, D., & Palmer, A.R. (2001) A neural code for low-frequency sound localization in mammals. *Nat Neurosci*, **4**, 396–401.

Meddis, R. & Hewitt, M.J. (1991a) Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity. *J Acoust Soc Am*, **89**, 2883–2894.

Meddis, R. & Hewitt, M.J. (1991b) Virtual pitch and phase sensitivity of a computer model of the auditory peripery. I: Pitch identification. *J Acoust Soc Am*, **89**, 2866–2882.

Merrill, E.G. & Ainsworth, A. (1972) Glass-coated platinum-plated tungsten microelectrodes. *Med Biol Eng*, **10**, 662–672.

Mills, A.W. (1958) On the Minimum Audible Angle. *J Acoust Soc Am*, **30**, 237–246.

Moore, J.K. (1986) Cochlear Nuclei: Relationship to the Auditory Nerve. In *Neurobiology of Hearing: The Cochlea*. pp. 283–381.

Moore, M.J. & Caspary, D.M. (1983) Strychnine blocks binaural inhibition in lateral superior olivary neurons. *J Neurosci*, **3**, 237–242.

Morest, D.K. (1968) The growth of synaptic endings in the mammalian brain: A study of the calyces of the trapezoid body. *Z Anat Entwicklungsgesch*, **127**, 201–220.

Murugasu, E. & Russell, I.J. (1995) Salicylate Ototoxicity: The Effects on Basilar Membrane Displacement, Cochlear Microphonics, and Neural Responses in the Basal Turn of the Guinea Pig Cochlea. *Audit Neurosci*, **1**, 139–150.

Nakamoto, K.T., Shackleton, T.M., Magezi, D.A., & Palmer, A.R. (2014) A function for binaural integration in auditory grouping and segregation in the inferior colliculus. *J Neurophysiol*, 1819–1830.

Nicol, M.J. & Walmsley, B. (2002) Ultrastructural basis of synaptic transmission between endbulbs of Held and bushy cells in the rat cochlear nucleus. *J Physiol*, **539**, 713–723.

Oertel, D., Wright, S., Cao, X.J., Ferragamo, M., & Bal, R. (2011) The multiple functions of T stellate/multipolar/chopper cells in the ventral cochlear nucleus. *Hear Res*, **276**, 61–69.

Oertel, D. & Wu, S.H. (1990) Morphology and physiology of cells in slice preparations of the posteroventral cochlear nucleus of mice. *J Comp Neurol*, **295**, 136–154.

Oliver, D.L., Beckius, G.E., & Shneiderman, A. (1995) Axonal Projections From the Lateral and Medial Superior Olive To the Inferior Colliculus of the Cat - a Study Using Electron-Microscopic Autoradiography. *J Comp Neurol*, **360**, 17–32.

Palmer, A.R., Liu, L.-F., & Shackleton, T.M. (2007) Changes in interaural time sensitivity with interaural level differences in the inferior colliculus. *Hear Res*, **223**, 105–113.

Palmer, A.R. & Russell, I.J. (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res*, **24**, 1–15.

Palmer, A.R. & Shackleton, T.M. (2002) The Physiological Basis of the Binaural Masking Level Difference. *Acta Acust*, **88**, 312–319.

Palmer, A.R., Wallace, M.N., Arnott, R.H., & Shackleton, T.M. (2003) Morphology of physiologically characterised ventral cochlear nucleus stellate cells. *Exp Brain Res*, **153**, 418–426.

Palmer, A.R. & Winter, I.M. (1993) Coding of the Fundamental Frequency of Voiced Speech Sounds and Harmonic Complexes in the Cochlear Nerve and Ventral Cochlear Nucleus. In *The Mammalian Cochlear Nuclei*. pp. 373–384.

Paolini, A.G., FitzGerald, J. V, Burkitt, A.N., & Clark, G.M. (2001) Temporal processing from the auditory nerve to the medial nucleus of the trapezoid body in the rat. *Hear Res*, **159**, 101–116.

Park, T.J., Brand, A, Koch, U, Ikebuchi, M., & Grothe, B. (2008) Dynamic Changes in Level Influence Spatial Coding in the Lateral Superior Olive. *Hear Res*, **238**, 58–67.

Park, T.J., Klug, A., Holinstat, M., & Grothe, B. (2004) Interaural level difference processing in the lateral superior olive and the inferior colliculus. *J Neurophysiol*, **92**, 289–301.

Patterson, R.D. (1969) Noise Masking of a Change in Residue Pitch. *J Acoust Soc Am*, **45**, 1520–1524.

Patterson, R.D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., & Allerhand, M. (1992) Complex sounds and auditory images. In *Auditory Physiology and Perception, Proc. 9th International Symposium on Hearing*. pp. 429–446.

Patterson, R.D. & Wightman, F.L. (1976) Residue pitch as a function of component spacing. *J Acoust Soc Am*, **59**, 1450–1459.

Pecka, M., Brand, A., Behrend, O., & Grothe, B. (2008) Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *J Neurosci*, **28**, 6914–6925.

Pecka, M. & Siveke, I. (2010) Enhancement of ITD coding within the initial stages of the auditory pathway. *J Neurophysiol*, **103**, 38–46.

Pfeiffer, R.R. (1966a) Anteroventral Cochlear Nucleus: Wave Forms of Extracellularly Recorded Spike Potentials. *Science (80-)*, **154**, 667–668.

Pfeiffer, R.R. (1966b) Classification of response patterns of spike discharges for units in the cochlear nucleus: Tone-burst stimulation. *Exp Brain Res*, **1**, 220–235.

Pickles, J.O., Comis, S.D., & Osborne, M.P. (1984) Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hear Res*, **15**, 103–112.

Plauška, A., Borst, J.G.G., & van der Heijden, M. (2016) Predicting binaural responses from monaural responses in the gerbil medial superior olive. *J Neurophysiol*, jn.01146.2015.

Plauška, A., van der Heijden, M., & Borst, J.G.G. (2017) A test of the stereausis hypothesis for sound localization in mammals. *J Neurosci*, 0233–17.

Pollak, G.D., Burger, R.M., & Klug, A. (2003) Dissecting the circuitry of the auditory system. *Trends Neurosci*, **26**, 33–39.

Pollak, G.D., Burger, R.M., Park, T.J., Klug, A., & Bauer, E.E. (2002) Roles of inhibition for transforming binaural properties in the brainstem auditory system. *Hear Res*, **168**, 60–78.

Pralong, D. & Carlile, S. (1994) Measuring the human head-related transfer functions: A novel method for the construction and calibration of a miniature “in-ear” recording system. *J Acoust Soc Am*, **95**, 3435–3444.

Pressnitzer, D. & Patterson, R.D. (2001) *Distortion Products and the Perceived Pitch of Harmonic Complex Tones*, Physiological and Psychophysical Bases of Auditory Function.

Quiroga, R.Q., Nadasdy, Z., & Ben-Shaul, Y. (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput*, **16**, 1661–1687.

Raatgever, J. & Bilsen, F.A. (1986) A central spectrum theory of binaural processing. Evidence from dichotic pitch. *J Acoust Soc Am*, **80**, 429–441.

Ramón y Cajal, S. (1909) *Histologie Du Système Nerveux de l’Homme et Des Vertébrés (L. Azoulay, Transi.), 2 Volumes*.

Recio-Spinoso, A. (2012) Enhancement and Distortion in the Temporal Representation of

Sounds in the Ventral Cochlear Nucleus of Chinchillas and Cats. *PLoS One*, **7**.

Rhode, W.S. (1995) Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. *J Acoust Soc Am*, **97**, 2414–2429.

Rhode, W.S., Oertel, D., & Smith, P.H. (1983) Physiological Response Properties of Cells Labeled Intracellularly With Horseradish Peroxidase in Cat Ventral Cochlear Nucleus. *J Comp Neurol*, **213**, 448–463.

Rice, J.J., May, B.J., Spirou, G.A., & Young, E.D. (1992) Pinna-based spectral cues for sound localization in cat. *Hear Res*, **58**, 132–152.

Ritsma, R.J. (1967) Frequencies dominant in the perception of the pitch of complex sounds. *J Acoust Soc Am*, **42**, 191–198.

Ritsma, R.J. & Bilsen, F.A. (1970) Spectral Regions Dominant in the Perception of Repetition Pitch. *Acustica*, **23**, 334–339.

Ritsma, R.J. & Engel, F.L. (1964) Pitch of Frequency-Modulated Signals. *J Acoust Soc Am*, **36**, 1637–1644.

Roberts, M.T., Seeman, S.C., & Golding, N.L. (2014) The relative contributions of MNTB and LNTB neurons to inhibition in the medial superior olive assessed through single and paired recordings. *Front Neural Circuits*, **8**, 49.

Robertson, D. & Irvine, D.R.F. (1989) Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J Comp Neurol*, **282**, 456–471.

Robertson, D. & Winter, I.M. (1988) Cochlear nucleus inputs to olivocochlear neurones revealed by combined anterograde and retrograde labelling in the guinea pig. *Brain Res*, **462**, 47–55.

Robles, L. & Ruggero, M.A. (2001) Mechanics of the Mammalian Cochlea. *Physiol Rev*, **81**, 1305–1352.

Rothman, J.S. & Young, E.D. (1996) Enhancement of neural synchronization in computational models of ventral cochlear nucleus bushy cells. *Audit Neurosci*, **2**, 47–62.

Rothman, J.S., Young, E.D., & Manis, P.B. (1993) Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J Neurophysiol*, **70**, 2562–2583.

Rouiller, E.M., Cronin-Schreiber, R., Fekete, D.M., & Ryugo, D.K. (1986) The Central Projections of Intracellularly Labeled Auditory Nerve Fibers in Cats: An Analysis of Terminal Morphology. *J Comp Neurol*, **249**, 261–278.

Rouiller, E.M. & Ryugo, D.K. (1984) Intracellular marking of physiologically characterized cells in the ventral cochlear nucleus of the cat. *J Comp Neurol*, **225**, 167–186.

Russell, I.J. & Sellick, P.M. (1978) Intracellular studies of hair cells in the mammalian cochlea. *J Physiol*, **284**, 261–290.

Rybalko, N., Popelář, J., Šuta, D., Large, C.H., & Syka, J. (2014) Impact of Kv3 Channel Modulator AUT3 on Auditory Temporal Resolution in Rats. In *Association for Research in Otolaryngology Abstracts*.

Ryugo, D.K. & Sento, S. (1991) Synaptic Connections of the Auditory Nerve in Cats: Relationship Between Endbulbs of Held and Spherical Bushy Cells. *J Comp Neurol*, **303:35**, 35–48.

Ryugo, D.K., Wu, M.M., & Pongstaporn, T. (1996) Activity-Related Features of Synapse Morphology: A Study of Endbulbs of Held. *J Comp Neurol*, **158**, 141–158.

Saint Marie, R.L. & Baker, R.A. (1990) Neurotransmitter-specific uptake and retrograde transport of [3H]glycine from the inferior colliculus by ipsilateral projections of the superior olivary complex and nuclei of the lateral lemniscus. *Brain Res*, **524**, 244–253.

Saldaña, E., Aparicio, M.-A., Fuentes-Santamaría, V., & Berrebi, A.S. (2009) Connections of the superior paraolivary nucleus of the rat: projections to the inferior colliculus. *Neuroscience*, **163**, 372–387.

Sanes, D.H. & Wooten, G.F. (1987) Development of glycine receptor distribution in the lateral superior olive of the gerbil. *J Neurosci*, **7**, 3803–3811.

Sayles, M., Smith, P., & Joris, P.X. (2016) Inter-aural Time Sensitivity of Superior-olivary-complex Neurons is Shaped by Systematic Cochlear Disparities. In *Association for Research in Otolaryngology Abstracts*. p. 273.

Sayles, M. & Winter, I.M. (2007) The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Res*, **1171**, 52–66.

Sayles, M. & Winter, I.M. (2008a) Reverberation Challenges the Temporal Representation of the Pitch of Complex Sounds. *Neuron*, **58**, 789–801.

Sayles, M. & Winter, I.M. (2008b) Ambiguous pitch and the temporal representation of inharmonic iterated rippled noise in the ventral cochlear nucleus. *J Neurosci*, **28**, 11925–11938.

Schofield, B.R. (1991) Superior paraolivary nucleus in the pigmented guinea pig: separate classes of neurons project to the inferior colliculus and the cochlear nucleus. *J Comp Neurol*, **312**, 68–76.

Schofield, B.R. (1995) Projections from the cochlear nucleus to the superior paraolivary nucleus in guinea pigs. *J Comp Neurol*, **360**, 135–149.

Schofield, B.R. & Cant, N.B. (1991) Organization of the superior olivary complex in the guinea pig. I. Cytoarchitecture, cytochrome oxidase histochemistry, and dendritic morphology. *J Comp Neurol*, **314**, 645–670.

Schofield, B.R. & Cant, N.B. (1992) Organization of the superior olivary complex in the guinea pig: II. Patterns of projection from the periolivary nuclei to the inferior colliculus. *J Comp Neurol*, **317**, 438–455.

Schofield, B.R., Mellott, J.G., & Motts, S.D. (2014) Subcollicular projections to the auditory thalamus and collateral projections to the inferior colliculus. *Front Neuroanat*, **8**, 1–16.

Schouten, J.F. (1940) The residue and the mechanism of hearing. *Proc K Ned Akad von Wet*, **43**, 991–999.

Schouten, J.F., Ritsma, R.J., & Cardozo, B.L. (1962) Pitch of the Residue. *J Acoust Soc Am*, **34**, 1418–1424.

Schreiner, C.E. & Langner, G. (1988) Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol*, **60**, 1823–1840.

Schroeder, M.R. (1977) New viewpoints in binaural interactions. *Psychophys Physiol Hear*, 455–467.

Scott, L.L., Hage, T.A., & Golding, N.L. (2007) 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.

Scott, L.L., Mathews, P.J., & Golding, N.L. (2005) Posthearing developmental refinement of temporal processing in principal neurons of the medial superior olive. *J Neurosci*, **25**, 7887–7895.

Seidl, A.H. & Grothe, B. (2005) Development of Sound Localization Mechanisms in the Mongolian Gerbil Is Shaped by Early Acoustic Experience. *J Neurophysiol*, **94**, 1028–1036.

Shackleton, T.M. & Carlyon, R.P. (1994) The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J Acoust Soc Am*, **95**, 3529–3540.

Shackleton, T.M., Liu, L.-F., & Palmer, A.R. (2009) Responses to diotic, dichotic, and alternating phase harmonic stimuli in the inferior colliculus of guinea pigs. *J Assoc Res Otolaryngol*, **10**, 76–90.

Shailer, M.J. & Moore, B.C.J. (1987) Gap detection and the auditory filter: Phase effects using sinusoidal stimuli. *J Acoust Soc Am*, **81**, 1110.

Shamma, S.A., Shen, N., & Gopalaswamy, P. (1989) Stereausis: binaural processing without neural delays. *J Acoust Soc Am*, **86**, 989–1006.

Shamma, S. & Klein, D. (2000) The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *J Acoust Soc Am*, **107**, 2631–2644.

Shofner, W.P. (1991) Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. *J Acoust Soc Am*, **99**, 2450–2466.

Shore, S.E. (2005) Multisensory integration in the dorsal cochlear nucleus: Unit responses to acoustic and trigeminal ganglion stimulation. *Eur J Neurosci*, **21**, 3334–3348.

Shore, S.E., Godfrey, D.A., Helfert, R.H., Altschuler, R.A., & Bledsoe, S.C. (1992) Connections between the cochlear nuclei in guinea pig. *Hear Res*, **62**, 16–26.

Shore, S.E., Sumner, C.J., Bledsoe, S.C., & Lu, J. (2003) Effects of contralateral sound stimulation on unit activity of ventral cochlear nucleus neurons. *Exp Brain Res*, **153**, 427–435.

Sinex, D.G. (2008) Responses of cochlear nucleus neurons to harmonic and mistuned complex tones. *Hear Res*, **238**, 39–48.

Smith, P.H. (1995) Structural and functional differences distinguish principal from nonprincipal cells in the guinea pig MSO slice. *J Neurophysiol*, **73**, 1653–1667.

Smith, P.H., Joris, P.X., Carney, L.H., & Yin, T.C.T. (1991) Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *J Comp*

Neurol, **304**, 387–407.

Smith, P.H., Joris, P.X., & Yin, T.C.T. (1993) Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J Comp Neurol*, **331**, 245–260.

Smith, P.H., Joris, P.X., & Yin, T.C.T. (1998) Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. *J Neurophysiol*, **79**, 3127–3142.

Smith, P.H. & Rhode, W.S. (1987) Characterization of HRP-labeled globular bushy cells in the cat anteroventral cochlear nucleus. *J Comp Neurol*, **266**, 360–375.

Smith, P.H. & Rhode, W.S. (1989) Structural and functional properties distinguish two types of multipolar cells in the ventral cochlear nucleus. *J Comp Neurol*, **282**, 595–616.

Smoorenburg, G.F. (1970) Pitch perception of two-frequency stimuli. *J Acoust Soc Am*, **48**, 924–942.

Sommer, I., Lingenhöhl, K., & Friauf, E. (1993) Principal cells of the rat medial nucleus of the trapezoid body: an intracellular in vivo study of their physiology and morphology. *Exp Brain Res*, **95**, 223–239.

Spangler, K.M., Warr, W.B., & Henkel, C.K. (1985) The projections of principal cells of the medial nucleus of the trapezoid body in the cat. *J Comp Neurol*, **238**, 249–262.

Spirou, G.A. & Berrebi, A.S. (1997) Glycine immunoreactivity in the lateral nucleus of the trapezoid body of the cat. *J Comp Neurol*, **383**, 473–488.

Spirou, G.A., Rager, J., & Manis, P.B. (2005) Convergence of auditory-nerve fiber projections onto globular bushy cells. *Neuroscience*, **136**, 843–863.

Spitzer, M.W. & Semple, M.N. (1995) Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *J Neurophysiol*, **73**, 1668–1690.

Spoendlin, H. & Schrott, A. (1989) Analysis of the human auditory nerve. *Hear Res*, **43**, 25–38.

Stange, A., Myoga, M.H., Lingner, A., Ford, M.C., Alexandrova, O., Felmy, F., Pecka, M., Siveke, I., & Grothe, B. (2013) Adaptation in sound localization: from GABA(B) receptor-mediated synaptic modulation to perception. *Nat Neurosci*, **16**, 1840–1847.

Terhardt, E. (1978) Psychoacoustic evaluation of musical sounds. *Percept Psychophys*, **23**, 483–492.

Terhardt, E. (1979) Calculating virtual pitch. *Hear Res*, **1**, 155–182.

Thompson, A.M. & Thompson, G.C. (1991) Projections from the posteroventral cochlear nucleus to the superior olivary complex in guinea pig: light and EM observations with the PHA-L method. *J Comp Neurol*, **311**, 495–508.

Tollin, D.J. (2003) The Lateral Superior Olive: A Functional Role in Sound Source Localization. *Neurosci*, **9**, 127–143.

Tollin, D.J., Koka, K., & Tsai, J.J. (2008) Interaural Level Difference Discrimination Thresholds for Single Neurons in the Lateral Superior Olive. *J Neurosci*, **28**, 4848–4860.

Tollin, D.J. & Yin, T.C.T. (2005) Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *J Neurosci*, **25**, 10648–10657.

Tsai, J.J., Koka, K., & Tollin, D.J. (2010) Varying overall sound intensity to the two ears impacts interaural level difference discrimination thresholds by single neurons in the lateral superior olive. *J Neurophysiol*, 875–886.

Tsuchitani, C. (1977) Functional organization of lateral cell groups of cat superior olivary complex. *J Neurophysiol*, **40**, 296–318.

Tsuchitani, C. (1988a) The inhibition of cat lateral superior olive unit excitatory responses to binaural tone bursts. I. The transient chopper response. *J Neurophysiol*, **59**, 184–211.

Tsuchitani, C. (1988b) The inhibition of cat lateral superior olive unit excitatory responses to binaural tone bursts. II. The sustained discharges. *J Neurophysiol*, **59**, 184–211.

Tsuchitani, C. (1997) Input from the medial nucleus of the trapezoid body to an interaural level detector. *Hear Res*, **105**, 211–224.

Tsuchitani, C. & Boudreau, J.C. (1969) Stimulus level of dichotically presented tones and cat superior olive S-segment cell discharge. *J Acoust Soc Am*, **46**, 979–988.

Tsuji, J. & Liberman, M.C. (1997) Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections. *J Comp Neurol*, **381**, 188–202.

Turecek, R. & Trussell, L.O. (2001) Presynaptic glycine receptors enhance transmitter release at a mammalian central synapse. *Nature*, **411**, 587–590.

Typlt, M., Haustein, M.D., Dietz, B., Steinert, J.R., Witte, M., Englitz, B., Milenkovic, I., Kopp-Scheinpflug, C., Forsythe, I.D., & Rübsamen, R. (2010) Presynaptic and postsynaptic origin of multicomponent extracellular waveforms at the endbulb of Held-spherical bushy cell synapse. *Eur J Neurosci*, **31**, 1574–1581.

van der Heijden, M., Lorteije, J.A.M., Plauška, A., Roberts, M.T., Golding, N.L., & Borst, J.G.G. (2013) Directional Hearing by Linear Summation of Binaural Inputs at the Medial Superior Olive. *Neuron*, **78**, 936–948.

van der Heijden, M., Louage, D.H.G., & Joris, P.X. (2011) Responses of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise: implications for the sensitivity to interaural delays. *J Assoc Res Otolaryngol*, **12**, 485–502.

Vonderschen, K. & Wagner, H. (2014) Detecting interaural time differences and remodeling their representation. *Trends Neurosci*, **37**, 289–300.

Wakeford, O.S. & Robinson, D.E. (1974) Lateralization of tonal stimuli by the cat. *J Acoust Soc Am*, **55**, 649–652.

Wallace, M.N., Rutkowski, R.G., Shackleton, T.M., & Palmer, A.R. (2000) Phase-locked responses to pure tones in guinea pig auditory cortex. *Neuroreport*, **11**, 3989–3993.

Wallace, M.N., Shackleton, T.M., & Palmer, A.R. (2002) Phase-locked responses to pure tones in the primary auditory cortex. *Hear Res*, **172**, 160–171.

Wang, Y.-X., Wenthold, R.J., Ottersen, O.P., & Petralia, R.S. (1998) Endbulb synapses in the anteroventral cochlear nucleus express a specific subset of AMPA-type glutamate receptor subunits. *J Neurosci*, **18**, 1148–1160.

Wenthold, R.J., Huie, D., Altschuler, R.A., & Reeks, K.A. (1987) Glycine immunoreactivity localized in the cochlear nucleus and superior olivary complex. *Neuroscience*, **22**, 897–912.

Wiegrebe, L. & Meddis, R. (2004) The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. *J Acoust Soc Am*, **115**, 1207–1218.

Wiersinga-Post, E. & Duifhuis, H. (1994) Probability distributions of the pitch of harmonic and inharmonic three tone complexes: Symmetric mistuning. In *Advances in Hearing Research: Proceedings of the 10th Internation Symposium on Hearing*. World Scientific, pp. 488–497.

Wightman, F.L. (1973) Pitch and stimulus fine structure. *J Acoust Soc Am*, **54**, 397–406.

Winter, I.M. & Palmer, A.R. (1990) Responses of single units in the anteroventral cochlear nucleus of the guinea pig. *Hear Res*, **44**, 161–178.

Winter, I.M., Robertson, D., & Cole, K.S. (1989) Descending projections from auditory brainstem nuclei to the cochlea and cochlear nucleus of the guinea pig. *J Comp Neurol*, **280**, 143–157.

Winter, I.M., Wiegrebe, L., & Patterson, R.D. (2001) Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus. *J Neurophysiol*, **85**, 1206–1219.

Wright, A., Davis, A., Bredberg, G., Ülehlová, L., & Spencer, H. (1987) Hair Cell Distributions in the Normal Human Cochlea: A Report of a European Working Group. *Acta Otolaryngol*, **436**, 15–24.

Wu, S.H. & Kelly, J.B. (1992) Synaptic Pharmacology of the Superior Olivary Complex Studied in Mouse Brain Slice. *J Neurosci*, **12**, 3084–3097.

Wu, S.H. & Kelly, J.B. (1994) Physiological evidence for ipsilateral inhibition in the lateral superior olive: synaptic responses in mouse brain slice. *Hear Res*, **73**, 57–64.

Yin, T.C.T. & Chan, J.C.K. (1990) Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol*, **64**, 465–488.

Yin, T.C.T., Chan, J.C.K., & Carney, L.H. (1987) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for Cross-Correlation. *J Neurophysiol*, **58**, 562–583.

Yin, T.C.T. & Kuwada, S. (1983) Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J Neurophysiol*, **50**, 1020–1042.

Yost, W.A. (1996) Pitch of iterated rippled noise. *J Acoust Soc Am*, **100**, 511–518.

Yost, W.A. & Dye, R.H. (1988) Discrimination of interaural differences of level as a function of frequency. *J Acoust Soc Am*, **83**, 1846–1851.

Yost, W.A., Patterson, R.D., & Sheft, S. (1996) A time domain description for the pitch strength of iterated rippled noise. *J Acoust Soc Am*, **99**, 1066–1078.

Yost, W.A., Patterson, R.D., & Sheft, S. (1998) The role of the envelope in processing iterated rippled noise. *J Acoust Soc Am*, **104**, 2349–2361.