Behavioural and neural correlates of binaural hearing

Joseph A. Sollini

Acknowledgements

I would like to thank Chris for having the patience to help me through my PhD, it was not an easy journey. Also to everyone at the IHR who helped me. Trevor and Mark, you both helped me greatly with your vast knowledge of the theory and technical expertise. To Rachel and Zoe for always doing your utmost to help out and generally with a smile. To Ben and Sue for making the office not just a place of work. To Scholes and Nick for always being in the pub, I can pay you no higher compliment. Basically to everyone who helped proof read (Kate, Joel and Jess particularly) and gave me advice. Obviously thanks to my Dad, Marilyn, Mum and Paul I love you all and without your emotional and financial support this would never have been possible. One day I only hope I can repay you (not financially, the moneys gone!). I had always hoped to end this thesis by writing something profound and lasting. I'm just going to settle for "it's finally over".

Abstract

The work in this thesis involves two separate projects. The first project involves the behavioural measurement of auditory thresholds in the ferret (*Mustela Putorius*). A new behavioural paradigm using a sound localisation task was developed which produces reliable psychophysical detection thresholds in animals. Initial attempts to use the task failed and after further investigation improvements were made. These changes produced a task that successfully produced reliably low thresholds. Different methods of testing, and the number of experimental trials required, where then explored systemically. The refined data collection method was then used to investigate frequency resolution in the ferret. These data demonstrated that the method was suitable for measuring perceptual frequency selectivity. It revealed that the auditory filters of ferrets are broader than several other species. In some cases this was also broader than neural estimates would suggest.

The second project involved the measurement of neural data in the Guinea Pig (*Cavia porecellus*). More specifically the project aimed to test the ability of the primary auditory cortex (AI) to integrate high frequency spatial cues. Two experiments were required to elucidate these data. The first experiment demonstrated a relationship between frequency and space, though these data proved noisy. A second experiment was conducted, focussing on improving the quality of the data this allowed for a more quantitative approach to be applied. The results highlighted that though AI neurons are responsive over a broad frequency range, inhibitory binaural interactions integrate spatial information over a smaller range. Binaural interactions were only strong when sounds in either ear were closely matched in frequency. In contrast, excitatory binaural interactions did not generally depend on the interaural frequency difference. These findings place important constraints on the across frequency integration of binaural level cues.

Contents

2 ELECTROPHYSIOLOGY 100

Abbreviations

ABL	**=**	**Average binaural level**
AC	**=**	**Auditory cortex**
AFC	**=**	**Alternative forced choice**
AI	**=**	**Primary auditory cortex**
AN	**=**	**Auditory nerve**
AVCN	**=**	**Anterior ventral cochlear nucleus**
BF	**=**	**Best frequency**
BM	**=**	**Basilar membrane**
BMLD	**=**	**Binaural masking level difference**
CB	**=**	**Critical bandwidth**
CDF	**=**	**Cumulative distribution function**
cFRA	**=**	**Contralateral frequency response area**
CN	**=**	**Cochlear nucleus**
CNIC	**=**	**Central cortex of the inferior colliculus**
COM	**=**	**Centre of mass**
CR	**=**	**Critical ratio**
CTF	**=**	**Contralateral tone frequency**
dB	**=**	**Decibel**
DCIC	**=**	**Dorsal cortex of the inferior colliculus**
DCN	**=**	**Dorsal cochlear nucleus**
DLPO	**=**	**Dorsolateral peri-olivary nucleus**
DMPO	**=**	**Dorsomedial peri-olivary nucleus**
DNLL	**=**	**Dorsal nucleus of the lateral lemniscus**
DPO	**=**	**Dorsal peri-olivary nucleus**
DTF	**=**	**Directional transfer function**
ECIC	**=**	**External cortex of the inferior colliculus**
ERB	**=**	**Equivalent rectangular bandwidth**
FRA	**=**	**Frequency response area**

fBF	**=**	**Facilitatory best frequency**
GOM	**=**	**Growth of masking**
HRTF	**=**	**Head related transfer function**
Hz	**=**	**Hertz**
IC	**=**	**Inferior colliculus**
iBF	**=**	**Inhibitory best frequency**
iCF	**=**	**Inhibitory characteristic frequency**
iFRA	**=**	**Ipsilateral frequency response area**
IFD	**=**	**Ipsilateral frequency difference**
IHC	**=**	**Inner hair cell**
ILD	**=**	**Interaural level difference**
ITD	**=**	**Interaural time difference**
ITF	**=**	**Ipsilateral tone frequency**
JND	**=**	**Just noticeable difference**
kHz	**=**	**Kilohertz**
LSO	**=**	**Lateral superior olive**
LTB	**=**	**Lateral nucleus of the trapezoid body**
MAA	**=**	**Minimum audible angle**
MAF	**=**	**Minimum audible field**
MANOVA	**=**	**Multivariate analysis of variance**
MAP	**=**	**Minimum audible pressure**
MGB	**=**	**Medial geniculate body**
MGB	**=**	**Medial geniculate body**
MLD	**=**	**Masking level difference**
MNTB	**=**	**Medial nuclei of the trapezoid body**
MPO	**=**	**Medial superior olivary complex**
MS	**=**	**Milliseconds**
MSO	**=**	**Medial superior olive**
OHC	**=**	**Outer hair cell**
PSTH	**=**	**Peri-stimulus time histogram**
PTC	**=**	**Psychophysical tuning curve**

PVCN	**=**	**Posteroventral cochlear nucleus**
SDT	**=**	**Signal detection theory**
SL	**=**	**Sensation level**
SNR	**=**	**Signal to noise ratio**
SOC	**=**	**Superior olivary complex**
SPL	**=**	**Sound pressure level**
VAS	**=**	**Virtual auditory space**
VCN	**=**	**Ventral cochlear nucleus**
VMPO	**=**	**Ventromedial peri-olivary nucleus**

1 Animal psychophysics

1.1 Literature review

There were two main objectives for the research presented in this chapter. The first objective was to find a behavioural paradigm which animals find easy to learn and yields low sound detection thresholds. The traditional approach to measuring sound detection thresholds has been, in essence, to ask the listener the question “Did you hear a sound?”. With the listener choosing to respond either “yes” or “no”, depending on their perception. Here an approach was taken where the listener was asked a different question i.e. “Where did the sound come from?” while varying the sound level. In the subsequent two sections I will develop and validate a novel method for measuring detection thresholds in ferrets, in which their task is to approach the direction of target sounds if they can hear them. In section 1.2 (three location discrimination) I will investigate ferrets’ ability to discriminate narrow band signals from three different locations as a function of signal to noise ratio (SNR). Performance on this task was poor, and did not appear to reflect the ability to detect sounds. In section 1.3 (two location discrimination) the task is reduced to two locations. The dependence on SNR in this task appears to reflect detection ability, yielding thresholds that are lower and more reliable than a previously used method in ferrets.

A sound localisation approach was chosen because it was hoped this would be a more intuitive task and hence require less training than a traditional sound detection task. A number of factors would influence the effectiveness of this task to collect detection thresholds. These included: the ability of the animals to use the sound localisation cues available to them, the relationship between sound level and localisation ability and any potential factors which could reduce the difficulty of the task. A background to these issues are introduced here by

discussing the cues available, how localisation ability is affected at low sound levels and factors which could be used to improve sound localisation ability.

The second objective of this chapter was to apply the developed paradigm to measure auditory filter widths in the ferret. The concept of the auditory filter width comes from extensive research in human behavioural psychophysics with accompanying research in animal electrophysiology. There has been very little research that measures animal behavioural auditory filter widths to accompany the animal electrophysiological data. In addition these data that do provide data on animal auditory filter widths often only estimate the filter widths as opposed to directly measuring them. In section 1.4 (auditory filter functions), this new method is applied to measurement of perceptual frequency resolution in ferrets.

1.1.1 Sound localisation ability

To assess the appropriateness of using a sound localisation task for measuring signal detection ability it is first necessary to discuss how sounds are localised and what factors may affect performance on such a task. To understand how sounds are converted into spatial objects it is first necessary to understand the localisation cues that the auditory system receives.

1.1.1.1 Sound localisation cues

The input signals to the ears are filtered differently depending on the position of a sound in space relative to the head. Fig 1.1.1 demonstrates the Head Related Transfer Function (HRTF) of a human. This is the frequency transfer function close to the ear drum, as the horizontal position of a sound source is varied in space. The grey colour scale indicates signal level at the ear drum relative to when the head was not present. Gain is greatest within the frequency range 4-6 kHz and at ~45˚. In addition to the gain introduced by the HRTF we can also

observe that at some locations and frequencies the HRTF attenuates the signal. This is particularly true for contralateral space (negative angles), where the relative level is much lower than in ipsilateral space. Detection thresholds measured in the free field are lowest around ±45˚ corresponding to the regions where the HRTF produces the most gain (Saberi *et al.*, 1991; Sabin *et al.*, 2005).

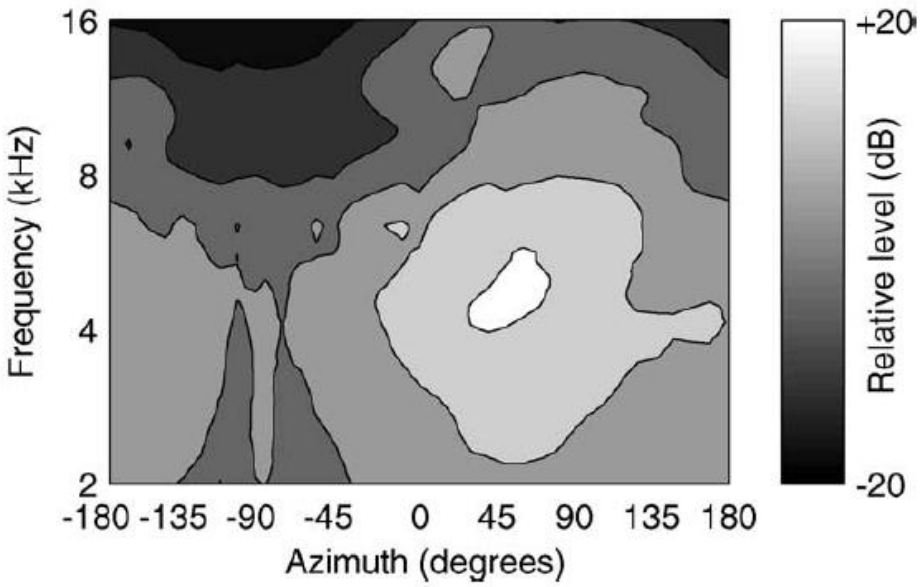

Fig 1.1.1. The head related transfer function for the right ear of a human listener. Negative azimuths indicate the left side of space (relative to the listener), positive values the right. Taken from Sabin et. al 2004.

The two ears, and their respective HRTFs, create a number of cues that are useful for sound localisation including: interaural time differences (ITDs), interaural level differences (ILDs) and spectral cues. The two dominant cues for azimuthal sound localization, ILDs and ITDs, have been known for over one hundred years (Rayleigh, 1907).

1.1.1.1.1 Interaural time differences (ITDs)

ITDs are produced by the sound travelling to each ear taking different paths and hence travelling different distances (Fig 1.1.2). When a sound is directly in front of a listener the signal should reach the two ears at the same time, as the path

length to the two ears is approximately equal. However when a signal is presented from a more lateral location the path lengths to the two ears will differ, hence creating a difference in arrival time to the two ears.

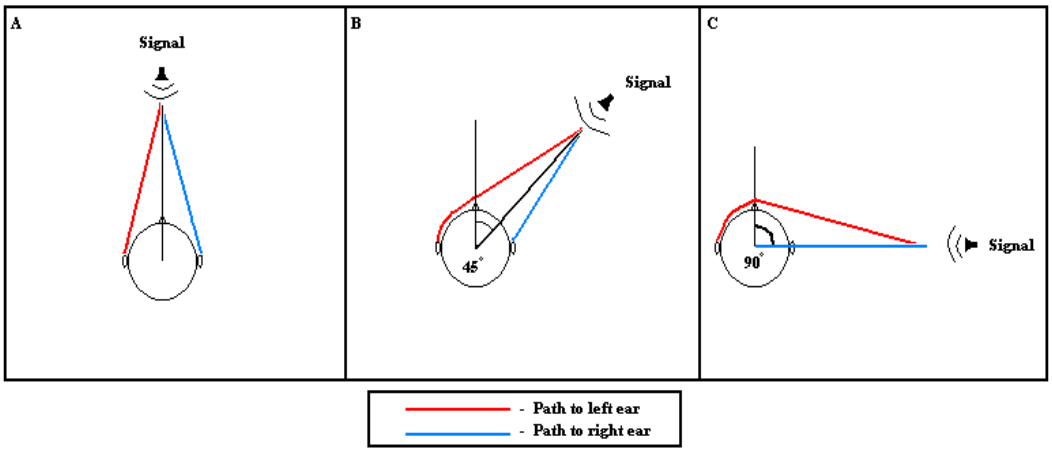

Fig 1.1.2 Illustrative diagram demonstrating differences in the paths to each ear from a single sound source. Sound originating from A) straight ahead (0˚), B) 45˚ C) 90˚.

Feddersen et al. (1957) demonstrated the dependence of ITD cues on azimuth (Fig 1.1.3). ITD increases, almost linearly, with azimuthal angle until close to 90˚. The ITD/azimuth function is symmetrical around 90˚ which means that for sound sources equidistant either side of 90˚ the ITD cue will be equal (von Hornbostel and Wertheimer, 1920). This has been called the "cone of confusion". The size of the ITD cue is also influenced by the frequency being presented, generally as the frequency is increased the size of the ITD cue is reduced (Kuhn, 1977).

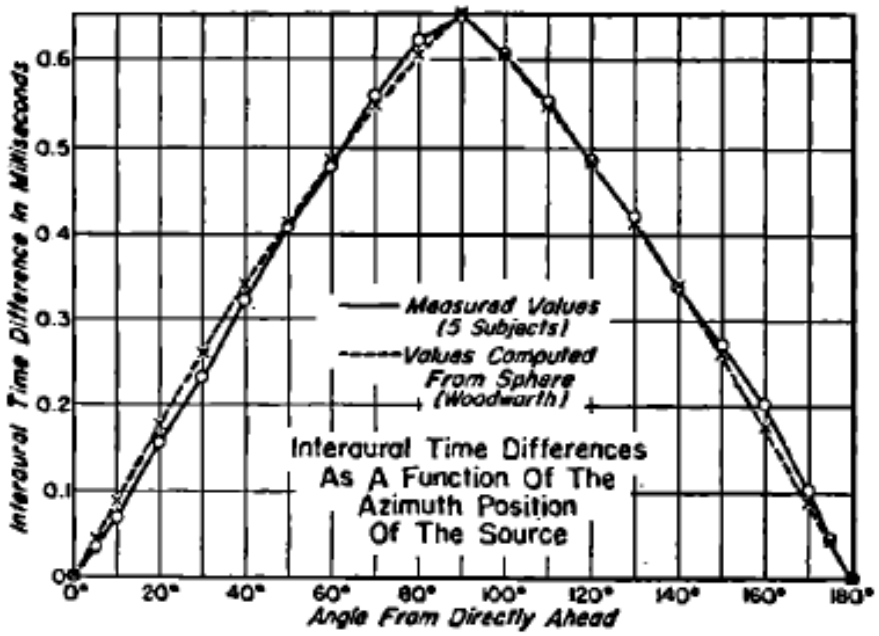

Fig 1.1.3. The average ITD of tones presented at a selection of azimuths, and averaged across frequency, to 5 human listeners. Taken from Feddersen et. al

1.1.1.1.2 Interaural level differences (ILDs)

ILDs are created by the signal to the ear contralateral to the sound source being attenuated by the head and ear, as in the example HRTF (Fig 1.1.1). Feddersen et al. (1957) studied the effect of varying the frequency of a pure tone signal on the size of the ILD cue, in humans (Fig 1.1.4). They found the size of the ILD is dependent both on frequency and location (Feddersen *et al.*, 1957). Relatively high frequencies (e.g. 3-6kHz) created large changes in ILD when the angle was moved from 0 to 90˚. At lower frequencies (200-500Hz) the size of the ILD cue was relatively unaffected by changes in azimuth. Put another way the usefulness of the ILD cue was restricted to high frequencies. At low frequencies ILD gave much less information about the position of a signal. This is in contrast to ITDs which, as mentioned previously, are larger at lower frequencies. They also noted that close to 90˚ ILD cues vary little with angle, suggesting at more lateral angles localisation performance should be poorer.

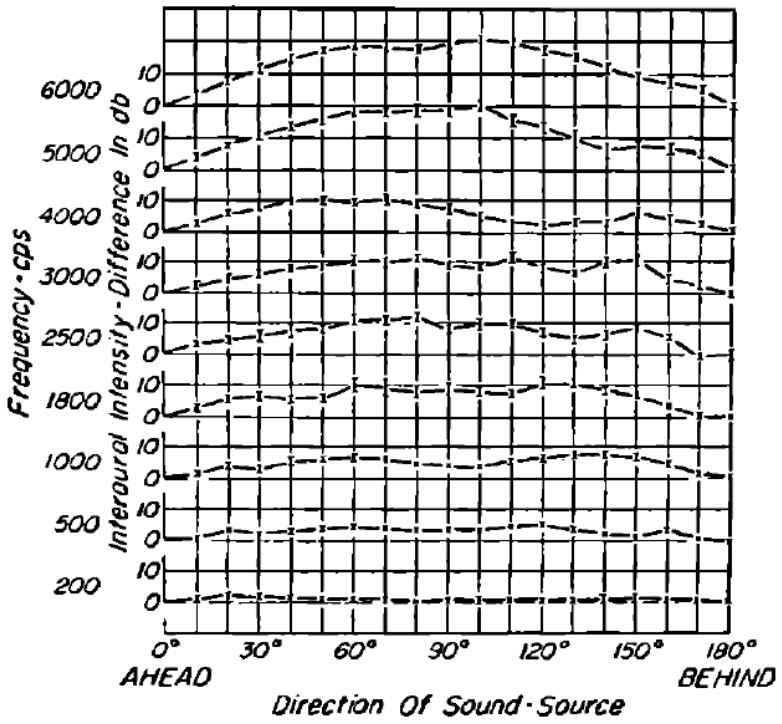

Fig 1.1.4 The average ILD of tones presented at a selection of frequencies and azimuths to 5 human listeners. Taken from Feddersen et. al 1957.

1.1.1.1.3 Spectral cues

The HRTF demonstrates that sounds are differentially amplified and attenuated by the head depending on frequency and source location. If the listener perceives the changes of sound level within frequency bands then these differences can be used as cues for sound localisation. An example of these differences can be seen in Fig 1.1.5, this demonstrates the spectra of broadband noise presented from 3 different elevations. As can be seen the entire spectral profile changes for each elevation. Some changes are systematic, e.g. in the 5-7 kHz and 10-11kHz regions increases in amplitudes directly relate to increases in elevation. These localisation cues are called spectral cues. Spectral cues are primarily used for localization in elevation but they are also very useful in resolving front/back errors in azimuthal localisation (Middlebrooks and Green, 1991).

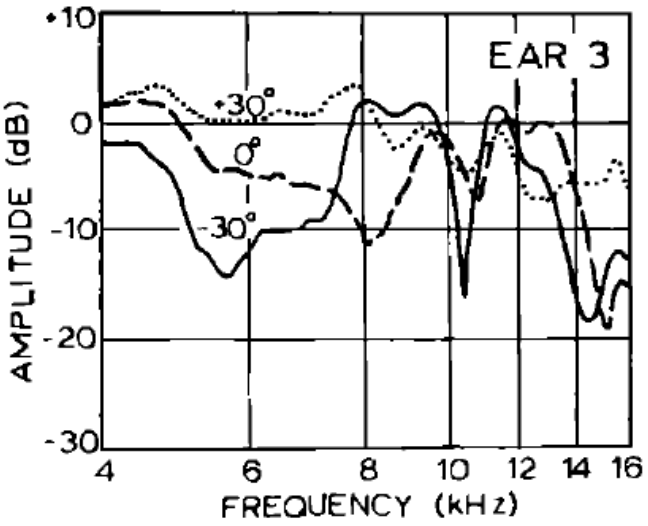

Fig 1.1.5. The spectrum of a flat broadband noise presented from three different elevations recorded from the position of the ear drum in an artificial

1.1.1.2 The duplex theory

The "duplex" theory states that humans use ITD cues to localise low frequency sounds, where they are largest, and ILD cues to localise high frequency sound, where they are largest (Rayleigh, 1907; Feddersen, 1957; Kuhn, 1977). However, the usefulness of the cue does not necessarily determine whether that cue is used, investigation of psychophysics is necessary to determine how these cues are used.

Studies over headphones have found the ability to detect ITDs declines above 1kHz and ITDs cues are, for pure tone stimuli, unusable above this (Durlach and Colburn, 1978). The ability to detect ILDs does not vary much between 200Hz and 10kHz, where ILDs <1dB could be detected at all tested frequencies (Durlach and Colburn, 1978). This suggests that at low frequencies humans could use both ILD and ITDs. In practice humans favour ITD cues at low frequencies as, in terms of angular space, ITD discrimination performance equates to better localisation ability (Moore, 2003). As ITDs are unusable at high frequencies ILDs become the dominant cue. Direct study has observed the relative dominance of ITD at low frequencies and ILD at high frequencies (Wightmann and Kistler, 1992; Macpherson and Middlebrooks, 2002).

Broadly speaking the duplex theory remains true but there are a number of caveats. While it is true that ITDs in the fine structure become unusable at high frequencies, ITDs in the envelope of high frequency signals can be used to lateralize sounds (Yost, 1971; Henning, 1974). In addition spectral cues can be used to resolve front/back confusions (Musicant and Butler, 1984; Middlebrooks and Green, 1991). While these additional cues are useful they do not play a dominant role in localising sounds (Macpherson and Middlebrooks, 2002).

1.1.1.3 Other factors affecting localisation

Frequency content is just one factor that influences the way in which a sound is localised. A number of factors can influence localisation ability including; the reference angle (Mills, 1958; Mills, 1960; Yost, 1975; Yost and Dye Jr, 1988), the signal bandwidth (Trahoitis and Stern, 1989; Terhune, 1974; Butler, 1986; King and Oldfield, 1997; Eberle *et al.*, 2000; Trahoitis *et al.*, 2001; Yost *et al.*, 2007), the signal detectability (Altshuler and Comalli, 1975; Comalli and Altshuler, 1976; Good and Gilkey, 1996; Lorenzi *et al.*, 1999; Sabin *et al.*, 2005), the number of response locations (Hartmann *et al.*, 1988), the stimulus duration (Tobias and Schubert, 1959; Tobias and Zerlin, 1959; Abel and Khunov, 1983), the stimulus rise time (Abel and Khunov, 1983), the number of sounds sources (Blauert, 1997) and the amount of reverberation in an enclosed space (Giguere and Abel, 1993). While all of these factors influence sound localisation not all of these factors are directly influential on the studies conducted here. For this reason a selection of these factors are focussed on here.

1.1.1.3.1 Reference angle

A number of studies have demonstrated that spatial acuity is greatest close to midline (0˚) and decreases with increasing angle from midline (Mills, 1958; Mills, 1960; Yost, 1975; Yost and Dye Jr, 1988). Potentially, this could be solely due to

the cues available in the free field, at more lateral angles the change in ILD and ITD with angle becomes smaller (Feddersen *et al.*, 1957). Though the fact the majority of these studies were conducted over headphones, where the cue available is directly controlled and hence available, demonstrates it is localisation ability which is poorer at lateral angles and not due to the availability of cues.

1.1.1.3.2 Signal bandwidth

It has been known for over 100 years that localisation improves as the signal bandwidth is increased from very narrowband stimuli, such as tones, to broadband stimuli, such as noise (Pierce, 1901). There are a number of listening situations where increasing bandwidth affects localisation performance. Firstly increasing the bandwidth of a high-pass sound to a broadband sound improves localisation performance (Wightman and Kistler, 1992). Increasing bandwidth changes the perceived laterality of low frequency sounds (Trahoitis and Stern, 1989), but this still results in better lateralisation performance (Trahoitis and Stern, 1989; Trahoitis *et al.*, 2001; Yost *et al.*, 2007). Localisation ability also improves with increasing bandwidth at high frequencies (Terhune, 1974; Butler, 1986; King and Oldfield, 1997; Eberle *et al.*, 2000).

A number of reasons could be posited for why increasing bandwidth improves performance. In the case of the first example it is due to the fact that the ITD information (low frequency) provides better estimates of the source location (Wightmann and Kistler, 1992). In free-field localisation studies increasing the bandwidth of a signal from a pure tone to broadband noise will introduce spectral cues which in turn will help resolve front/back confusions (Butler, 1986; Middlebrooks and Green, 1991). However, headphone studies have also demonstrated improvements in ILD and ITD resolution with increasing bandwidth (Trahoitis and Stern, 1989; Trahoitis *et al.*, 2001; Hartmann and Constan, 2002; Yost *et al.*, 2007). Given that spectral cues are specific to the

HRTF and the headphone studies discussed do not artificially contain spectral cues this suggests spectral cues are not solely responsible for these improvements. Butler (1986) removed improvements due to front/back resolutions from localisation data and found increasing signal bandwidth lead to modest improvements in free field localisation performance.

1.1.1.3.3 Signal detectability

It is now well known that the ability to identify a signal's position declines as the signal level approaches detection threshold. This has been shown over headphones (Altshuler and Comalli, 1975; Comalli and Altshuler, 1976) and open field studies, both in silence (Sabin *et al.*, 2005) and in the presence of noise (Good and Gilkey, 1996; Lorenzi *et al.*, 1999). It has also been shown for different types of stimuli including tones (Jacobsen, 1976), narrow band noise (Abel and Hay, 1996) broadband noise (Sabin *et al.*, 2005) and filtered pulse trains (Good and Gilkey, 1996; Lorenzi *et al.*, 1999). The relevant findings to these studies can be summarised as follows. Azimuthal localisation is less susceptible to low signal levels than vertical localisation (Su and Recanzone, 2001). Localisation ability in the azimuthal plane asymptotes at around 20dB SL for localisation in silence (Sabin *et al.*, 2005) and between 0-12 dB SNR in the presence of noise (Good and Gilkey, 1996; Lorenzi *et al.*, 1999) and declines at lower signal levels. While performance at lower signal levels declines for signals presented at the most lateral signal locations a large majority of errors are restricted to the same hemisphere (Good and Gilkey, 1996; Lorenzi *et al.*, 1999; Sabin *et al.*, 2005). Finally, localisation ability is differentially affected by noise dependant on the frequency of the sound being presented (Good and Gilkey, 1996; Lorenzi *et al.*, 1999; Sabin *et al.*, 2005). Whereas, in localisation in quiet, low frequency ITD cues govern responses to broadband stimuli (Wightman and Kistler, 1992), in noise, high frequency ILD and spectral cues are more influential in particular listening conditions i.e. when the masker is not near to midline (Lorenzi *et al.*, 1999).

1.1.1.3.4 Number of response locations

Early study of location identification were restricted to either a single plane (Stevens and Newman, 1936; Sandel *et al.*, 1955; Terhune, 1974; Perrott *et al.*, 1987) or used a highly reduced subset of possible speaker locations (Blauert, 1969; Shelton and Searle, 1978; Butler, 1986). Recent study has demonstrated that reducing the number of speaker locations improves localisation performance (Hartmann *et al.*, 1998). This is for the simple reason that reducing the number of speakers over a given span (angular distance between the first and last speaker) reduces the overlap between the underlying distribution of space and response locations. To ensure precise measurement of the underlying distribution the speaker separation angle must be much smaller than the underlying distribution, Hartmann's data implies that, for humans speaker separations of around 4˚ should suffice. Obviously presenting sounds from speakers at 4˚ separations through the entirety of a sphere surrounding a listener is incredibly impractical (this would require over 8000 speakers). Fortunately the restriction is on the number of response locations as opposed to the location of presentation. Therefore this problem can be easily sidestepped by using techniques that allow listeners to freely indicate the sound source, such as hand (Djelani *et al.*, 2000), head (Thurlow and Runge, 1967; Makous and Middlebrooks, 1990; Bronkhorst, 1995; Lewald *et al.*, 2000) or eye pointing (Frens and Opstal, 1995; Yao and Peck, 1997) or using a laser spot (Seeber, 2002) if the uncertainty of this method is substantially less than that due to the sensory phenomena under scrutiny (Makous and Middlebrooks, 1990).

1.1.1.4 Sound localisation ability of animals

Location discrimination has been tested in many non-human species (Renaud and Popper, 1975; Heffner, 1978; Brown *et al.*, 1980; Heffner and Heffner, 1980; Kelly, 1980; Heffner and Heffner, 1982; Heffner and Heffner, 1984; Heffner and

Heffner, 1987; 1988b; a; c). Of all species of mammals tested, humans remain the "gold standard" for sound localisation performance (Heffner and Masterton, 1990). For a number of years it has been believed that the owl holds the title for the greatest sound localisation acuity of all tested terrestrial animals (Knudsen *et al.*, 1979), although recent study has suggested that the cat may, in fact, have better localisation acuity (Tollin *et al.*, 2005). This currently remains the source of some debate (Heffner *et al.*, 2005; Tollin *et al.*, 2005; Moore *et al.*, 2008).

In terms of location discrimination humans yield considerably lower thresholds than the majority of mammals. Thresholds on MAA tasks range from 0.9 to greater than 60˚, starting with dolphins through to the pocket gopher (Heffner and Masterton, 1990). For non-human mammals, the mean threshold is 15˚ with a median of 12˚, with only the elephant and dolphin demonstrating comparative performance to humans (for those included in Heffner's review). MAAs have been measured in the ferret using single click stimuli (Kavanagh and Kelly, 1987). MAAs at midline were ~18˚. It is difficult to compare this value with those other species mentioned, as the majority of these studies have used click trains and long duration noise. These two factors improve localisation performance (Kavanagh and Kelly, 1987). As with humans, MAAs tested at more lateral locations (reference location of ±60˚) were larger in the ferret, ~30˚ (Kavanagh and Kelly, 1987).

Houben and Gourevitch measured ITD and ILD discrimination thresholds over headphones in the macaque. These authors found that the smallest detectable ILD difference was 3.5 dB at 500 Hz. This is noticeably higher than those recorded in humans. Surprisingly, the authors also found that ILD discrimination thresholds increased with frequency, as opposed to in humans, where they remain relatively stable for a wide range of frequencies (Fig 1.1.6). Also, in contrast to humans, the macaque can make ITD discriminations up to 2000 Hz and possibly higher. The smallest measured ITD threshold was 42 µs at

1.5 kHz, again larger than in humans. The ITD performance itself was a good fit for free-field discrimination thresholds, suggesting that the macaque uses ITD cues at low frequencies and ignores the ILD cues available (Houben and Gourevitch, 1979).

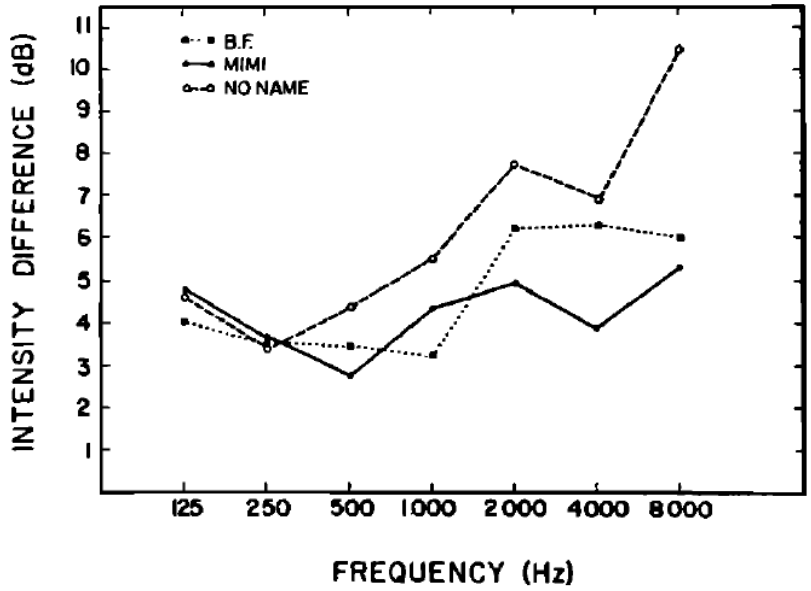

Fig 1.1.6. Interaural intensity difference discrimination thresholds in the macaque at different frequencies. Symbols are for each animal. Taken from Houben and Gourevitch, 1979.

ITD and ILD discrimination thresholds have also been measured via headphones in the cat, though over a more restricted frequency range (Wakeford and Robinson, 1974). The authors demonstrated that, unlike with the monkey, IID thresholds actually decreased with frequency over the range tested (Fig 1.1.7). The IID thresholds were small (between 0.5 and 1.8dB) and comparable to those gained in humans (Durlach and Colburn, 1978). ITD thresholds at 0.5 and 1kHz were ~20µs, noticeably smaller than in the macaque and comparable to those found in humans. Above 1 kHz the ITD thresholds increased, potentially demonstrating the cut-off of useable ITD information. ITD thresholds have also been studied in the rabbit, yielding an average threshold of 50-60 µs, larger than that of the cat (Ebert Jr *et al.*, 2008). In general there is ample evidence available to suggest that mammals are capable of using both ITD and ILD cues

and, from the studies mentioned here, the use of the cues in isolation suggests that there is some evidence to suggest the duplex theory holds.

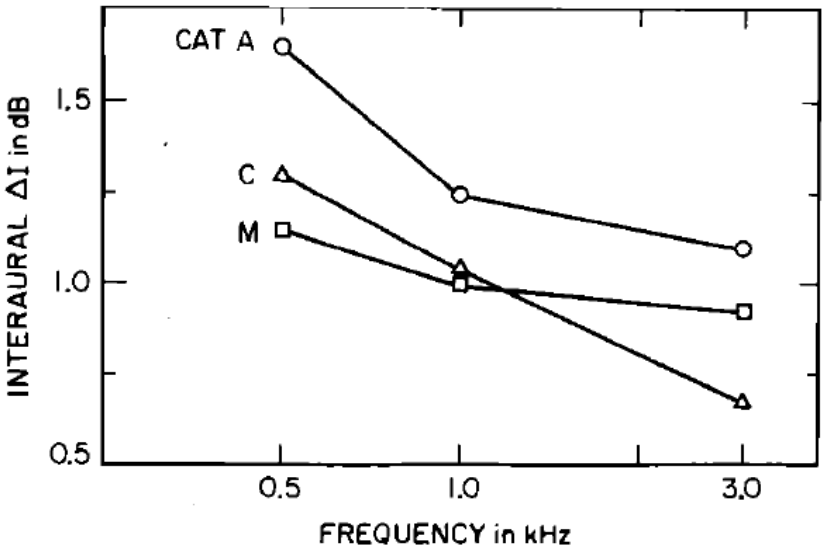

Fig 1.1.7. Interaural intensity difference discrimination thresholds in the cat at different frequencies. Symbols are for each animal. Taken from Wakeford and Robinson, 1974.

The effect of signal bandwidth has also been demonstrated in the macaque (Brown *et al.*, 1980). Brown et al. (1980) measured MAAs in macaque (*macaca*) in the free field. Animals were tested using band-pass noise with a range of centre frequencies and a range of bandwidths. The results can be summarised as follows: at narrow bandwidths thresholds were lower at low frequencies, but at broader bandwidths (8 kHz wide) thresholds were lower and increases in the signal bandwidth led to lower discrimination thresholds at the majority of centre frequencies tested.

Increasing bandwidth has also been shown to improve performance in ferrets performing a location identification task (Kacelnik *et al.*, 2006). Ferrets were tested using broadband and 1/6 octave wide noise-bands centred at 1 and 15 kHz on a 12 speaker localisation task. Performance in the broadband condition was markedly better on broadband signals than on narrowband signals. Study of location identification in the monkey (*Macaca*) has also demonstrated that,

when moving from broadband noise to pure tone signals accuracy of orientation behaviour decreased for narrower bandwidth stimuli (Recanzone *et al.*, 2000). Huang and May (1996) measured sound orientation behaviour in the cat. They found that reductions in bandwidth from a broadband noise to a narrowband noise had no effect on head orientation performance unless the frequency information at higher frequencies coincided with a spectral notch in the individual cat HRTF (Huang and May, 1996; May and Huang, 1996).

1.1.1.5 Localisation summary

For each cue there are limitations on their size, for instance at low frequencies ITDs are larger and at high frequencies ILDs are larger. There are also limitations on their availability, for example narrow band sounds like pure tones contain no spectral cues. In addition for each cue there is a corresponding ability to use this cue, below 1kHz ITDs in the fine structure produce the best performance of all localisation cues but despite ITDs being present above 1.5kHz we are unable to use them. Generally in each listening situation the auditory system gives the most weight to the most useful cue. Whether this be to favour ITD information instead of ILD information when both are available (Wightmann and Kistler, 1992; Macpherson and Middlebrooks, 2002), to favour ILD information over spectral cues and envelope ITDs at high frequencies (Eberle *et al.*, 2000; Macpherson and Middlebrooks, 2002) or to use ILD and spectral cues at low frequencies under particular listening conditions (Lorenzi *et al.*, 1999).

Localisation studies have also been conducted in other mammals. Decreasing signal bandwidth, generally, decreases sound localisation performance in mammals. The notable exception is the cat, which appears to heavily rely on spectral notches provided by the HRTF. Provided the frequency information within this notch region is present, localisation performance remains relatively stable. It is not entirely clear whether, like humans, non-human mammals have

increased location sensitivity at 0˚. In the ferret, MAAs have been shown to increase as the reference location is moved from 0˚ to more lateral locations, whereas very little difference was found in the monkey (Kavanagh and Kelly, 1987; Recanzone *et al.*, 1998).

1.1.2 Sound detection ability

The ability to hear a sound, signal detection ability, is influenced by a number of factors. These can be due to properties of the signal. For example: the frequency and the level (Wienn, 1899; Abraham, 1907; Fletcher and Wegel; 1922), the duration (Hughes, 1946), the rise time (Walton and Wilson, 1974) and the spatial position of a signal (Sabin *et al.*, 2005). It can also be due to properties of interfering sounds. For example the level of masking sounds relative to the signal (Mayer, 1876), the frequency content of the masker (Mayer, 1876; Wegel and Lane, 1924; Fletcher, 1940), the frequency content of signal relative to the masker (Green, 1960), the onset of signal relative to the masker (Bronstein and Churilova, 1936; Samoilova, 1956; Gustafsson and Arlinger, 1994) the relative spatial location of signal and masker (Jeffress *et al*., 1956; Saberi, 1994) and the degree of modulation in the signal and masker (Carhart *et al*., 1966; Hall *et al*., Gustafsson and Arlinger, 1994). Furthermore these effects interact with each other and are often related. For example the relationship between signal intensity, duration and rise time (Hughes, 1946; Pedersen and Elberling, 1972). There is neither the time nor necessity to discuss all of these factors here but a number are important to studies in this thesis.

1.1.2.1 The auditory filter

Fletcher (1940) used band-pass noise maskers of increasing bandwidth, but constant spectral density, to test hearing thresholds to a fixed frequency tonal signal. The masker was centred on the signal frequency. Fletcher found that, as

the masker bandwidth was increased, the signal threshold was progressively increased before flattening off, where further increases in the masker bandwidth did not significantly alter detection thresholds. As the spectral density remained constant throughout, each increment in masker bandwidth added additional power to the delivered masker. Once this additional power falls at distant frequency regions no further increase in threshold is measured. The result is interpreted as demonstrating a band-pass filter centred on the signal frequency. When the masker bandwidth is wide enough, additional power added falls outside of this filter and hence does not additionally mask the signal. Zwicker (1957) termed the bandwidth at which additional increases in masker bandwidth no longer increase the signal threshold the "critical bandwidth".

Fletcher suggested that this result implied the peripheral auditory system (specifically the BM) behaves as if it contains a bank of band-pass filters. The filters measured by psychophysical experiments are now termed "auditory filters". It is not clear precisely how or where the bandwidths demonstrated in psychophysical experiments are derived or represented in the auditory system. It does, however, seem likely that the frequency analysis inherited from the BM is involved (Moore, 2003).

1.1.2.2 Describing auditory filters

For the sake of convenience Fletcher assumed that the auditory filter could be approximated with a flat top and vertical sides. Thus the filter approximates equal weight to all frequencies within its pass-band (the critical band, or CB, not to be confused with the critical bandwidth). Fletcher pointed out that it was possible to estimate the auditory filter width by calculating the critical band. This relies on two hypotheses, firstly that only a narrow-band of frequencies surrounding the signal frequency contribute to masking, as discussed above. Secondly that the following relationship holds:

$$\frac{P}{WN_o} = k$$

Where *P* is the power of the signal and WN_o is the power of the noise within the CB. *W* is the width of the critical band (as opposed to the critical bandwidth) in Hz and N_o is the power density of the noise itself, as it is white noise it has a flat spectrum so equal power at all frequencies. *k* is presumed to be a constant and is regarded as the efficiency of the detection process. This can be reworked:

$$W = \frac{P}{kN_o}$$

As *P* and N_o are known, and by estimating *k*, it is possible to estimate the CB. Fletcher estimated that *k* was equal to 1, therefore:

$$W = \frac{P}{N_o}$$

P/N_o is referred to as the critical ratio and is often used to estimate auditory filter widths. More recent studies have demonstrated that *k* is typically around 0.4 (Scharf, 1970), furthermore it has been demonstrated that *k* can vary with centre frequency so the critical ratio does not give correct indication of how the CB will vary with centre frequency.

The CB estimates the bandwidth of an auditory filter. Other approaches have attempted to estimate the shape of the auditory filter. The psychophysical tuning curve (Chistovich, 1957) attempts to directly estimate psychophysical tuning shape whereas the notch-noise method (Patterson, 1976) attempts to estimate it indirectly.

1.1.2.3 Auditory filters in animals

Behaviourally measuring auditory filter widths in animals can be time consuming (Yost and Shofner, 2009). This is because each animal has to be trained to the task, animal performance is often variable (requiring extra repeats) and because a number of conditions need to be measured at a number of frequencies. This has meant very few animal studies have directly measured auditory filter widths (Fay, 1988). As described previously, the critical ratio is an estimate of the critical bandwidth. Often this has been used to estimate the frequency selectivity of animals. Fig 1.1.8 demonstrates the resulting measures of critical ratios found in the cat, macaque and chinchilla.

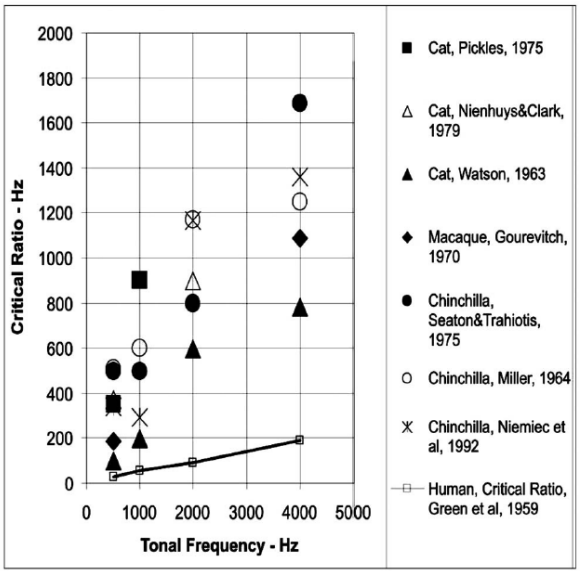

Fig 1.1.8. Critical measure estimates based on behavioural data for the cat, chinchilla, macaque and human. Symbols indicate animal and study. Taken from Yost and Schofner, 2009.

These critical ratio data suggest that frequency selectivity in animals is poor relative to humans, with the species demonstrated here yielding CRs several orders of magnitude larger. However, direct measurement of CBs result in markedly lower values, suggesting that CRs yield poor estimates of CBs in animals, though still generally higher than in humans (May *et al.*, 2006; Yost and Shofner, 2009).

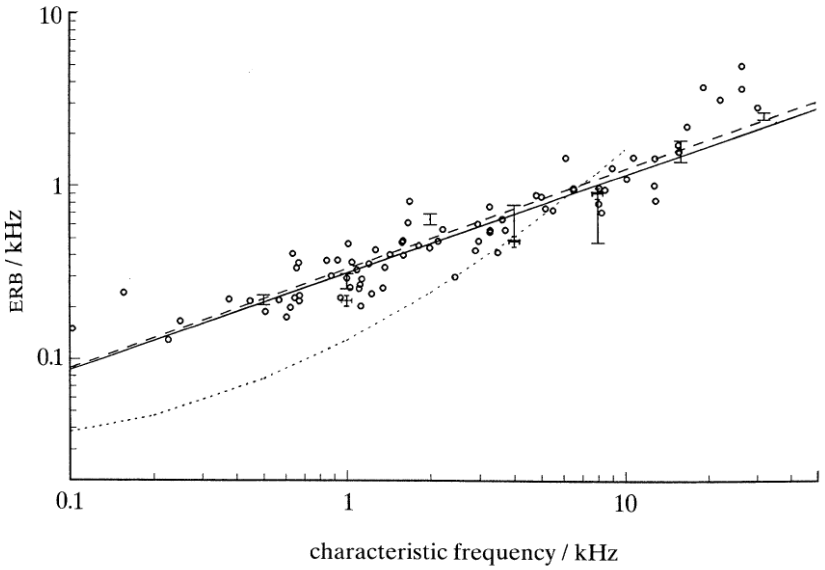

Fig 1.1.9. Comparison of ERBs measured physiologically in the auditory nerve and measured behaviourally in the guinea pig. Bracket symbols: behavioural ERBs derived from comb-filtered noise masking. Circles: behavioural ERBs derived from band-stop noise masking. Continuous line: regression through the mean EBS (comb filtered). Dashed line: regression through the physiological data points. The dotted line represents human ERBs. Taken from Evans, 1992.

A few studies have demonstrated a close relationship between physiological and behavioural measures of frequency selectivity in animals (Evans, 1972; Pickles, 1975; Ehret and Schreiner, 1997). Fig 1.1.9 demonstrates one such example comparing physiological responses of auditory nerve fibers against behavioural measures of frequency selectivity (measured in ERBs). As can be seen, there is a close correspondence between the behavioural and physiological results. At high levels of the auditory system, such as auditory cortex, this relationship is not so obvious, though a close correspondence has been found in inferior colliculus cells in the cat (Ehret and Schreiner, 1997).

1.1.2.4 Signal detection summary

A number of factors determine the detectability of a signal. Those that are important here are that the detectability of signal is influenced by interfering noise but only over a particular frequency range (the critical bandwidth). This is

because of the filtering effect of the auditory filter. Studies in animals suggest that critical bandwidths might be larger in animals, though this could be due to methodological differences. There is as yet no consensus on whether behavioural frequency selectivity can be measured directly from neurons.

1.2 Three-location discrimination

1.2.1 Introduction

The aim of perceptual psychophysical testing is to measure the limits of perception. This can often be difficult in animal models where task design can have a profound effect on the thresholds measured and the length of time required to train the animals (Birch, 1968; Ehret, 1974; Heffner and Masterson, 1980; Zheng et al., 1999; Prosen et al., 2003; Klink et al. 2006 and Radziwon et al., 2009). In addition tasks are rarely generalisable and so to measure two different perceptual phenomena, for instance sound detection and localisation ability, requires training for and measurement on two different tasks (Kavanagh and Kelly, 1988). The amount of training needed for sound localisation tasks, in animal models, is often less than that required for sound detection tasks, suggesting that it could be a more intuitive or ecological task for animals to perform. The purpose of this study was to investigate the potential of using a simplified sound localisation task to collect signal detection thresholds.

While sound detection and localisation are considered different abilities the two are related, it is well known that reducing the detectability of a sound reduces the ability to localise it (Sabin, Good, Lorenzi, Jacobsen, Abel). For humans localising sounds in background noise performance begins to become affected between 0 and 6 dB SNR (Good and Gilkey, 1996; Lorenzi et al., 1999). Whereas detection thresholds in noise are much lower than this (Lorenzi et al., 1999). This would suggest using a sound localisation task to measure detection thresholds would result in elevated detection thresholds. These studies, however, were conducted using the full 360˚ of the azimuthal plane. Reducing the number of response locations and increasing the angle between sound sources improves localisation ability (Hartmann, 2002). It was hoped that by reducing the number of sound sources and hence demands on localisation

ability the degradation of localisation performance would occur at a similar SNR to detection threshold.

An animal model, the ferret (*Mustela putorius*), was used to test this idea. First animals were trained on a localisation task and then a reduced localisation-in-noise task employed to test if detection-in-noise thresholds could be measured. If successful this could result in a task that when compared to traditional sound detection tasks would require less training time, produce similar thresholds and be generalisable enough to collect data for two different perceptual phenomena.

1.2.2 Methods

1.2.2.1 Subjects

Subjects were 3 male and 1 female pigmented ferrets (*Mustela Putorius*) which were either bred in-house (2/4) or brought in from an accredited breeding farm (2/4). Those ferrets obtained externally were kept in quarantine for 6 weeks to ensure they were free of infection and ear mites before being moved to the main housing rooms. Male ferrets were housed one to a cage. Females were either housed separately or in a group depending on the nature of the ferret but were separated while on water regulation.

The housing environment complied with the Code of Practice for the Housing and Care of Animals in Designated Breeding and Supplying Establishments issued under section 21 of the Animals (Scientific Procedures) Act 1986. Cages were 72×85×50cm with a total floor area of 6120cm². Seasonal changes to the animal environment were applied dependent on the season. Seasons were defined so that November through to the end of February was winter, March was spring, April through to September was summer and October was Autumn. Temperature was varied seasonally at 17, 18 and 19˚c ± 2˚c in winter,

spring/autumn and summer. Light was also varied seasonally for winter (8hrs light and 16 hrs dark) Autumn/Spring (12 hrs light and 12 hrs dark) and summer (16 hrs light and 8 hrs dark). Cages were cleaned once a week. Environmental enrichment was provided in the form of tubes, hammocks, nesting materials and balls. In addition animals were allowed out of their cages on a daily basis to allow them to exercise, interact and to stimulate investigative and play behaviour.

Animal water was regulated when they participated in behavioural tasks. Water was withdrawn from their cage and delivered as reward for the required behaviour during behavioural sessions. When on regulation animals were given wet alternative to their normal food diet composed of their protein pellets ground up with a food supplement (Cimicat, Petlife International Ltd., UK). Water regulation periods could be as long as 14 days with the animals completing at least two behavioural sessions each day but on occasion this could be extended to three. After each regulation period animals were provided water *ad libum* for a minimum of three days. After three regulation periods and two rest periods of this sort the animals were given an extended rest period anywhere between 5 and 14 days. Rest periods and extended rest periods were intended to help the animals regain any weight lost while the animal's water was regulated. Weight was monitored and health of the animal was checked on a daily basis.

1.2.2.2 Apparatus

Animals were tested in one of two identical circular arenas measuring 150 cm in diameter (Fig 1.2.1).

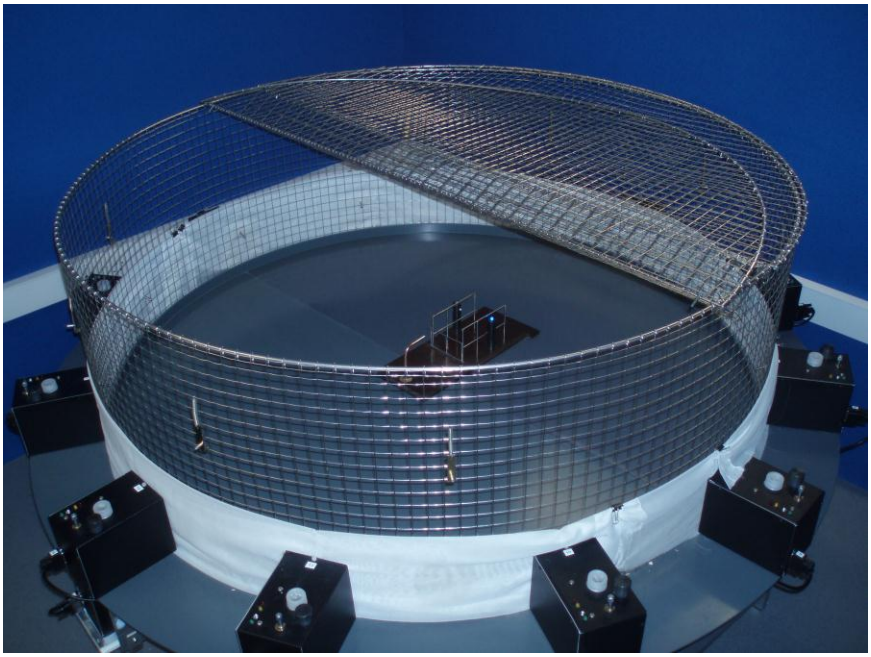

Fig 1.2.1. A behavioural arena used for animal psychophysics. Located in the centre of the arena is a platform (Fig 1.2.2), see text for details.

Both arenas were surrounded laterally around the circumference by a mesh fence and covered by a mesh lid (Fig 1.2.1). Each arena was housed inside a sound-attenuated room. The base of each arena was made of uncovered hard PVC. In the centre of the arena floor was a raised platform holding a fence, two posts and a spout (Fig 1.2.3). The centre platform was designed to encourage the same head and body position for each trial. A nose cone was fitted if an animal was found to have too varied or irregular head position during signal delivery.

The enclosed arena was surrounded by 12 modules. Each module had a speaker for sound delivery and a water spout for water delivery. Each spout was connected to the modules own reservoir which was filled with distilled water. There modules separated by 30˚ along the azimuthal plane covering 360˚. The speakers on the modules were Visaton FX 10 coaxial speakers with a maximum power of 70 W and a frequency response range between 70-22000Hz. The spouts detected a lick by using an infrared beam which would be broken when the end of the spout was covered (i.e. when the spout was licked).

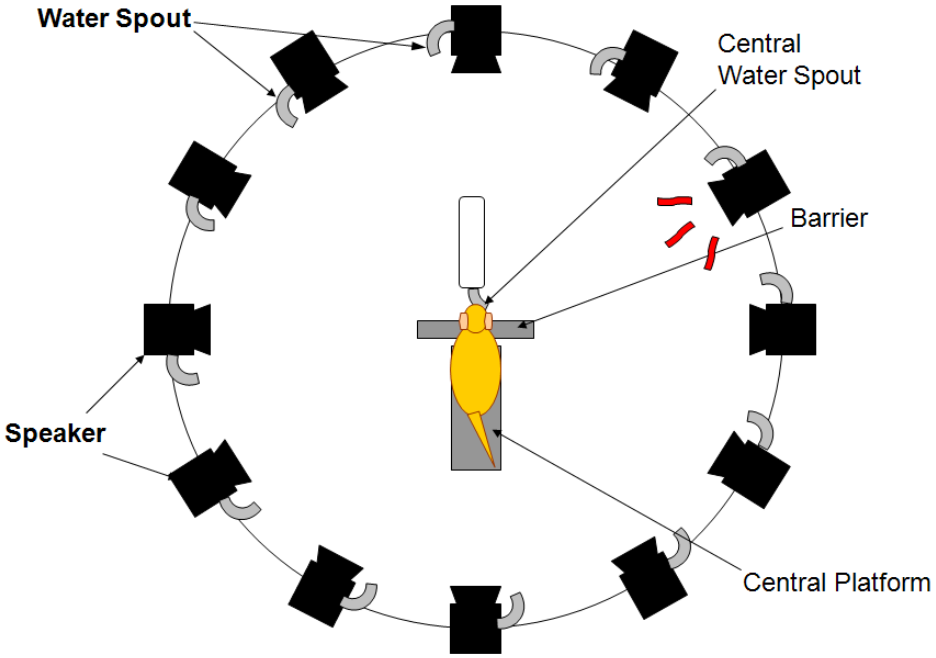

Fig 1.2.2. Arena used to train and test ferrets for localisation ability on the horizontal plane.

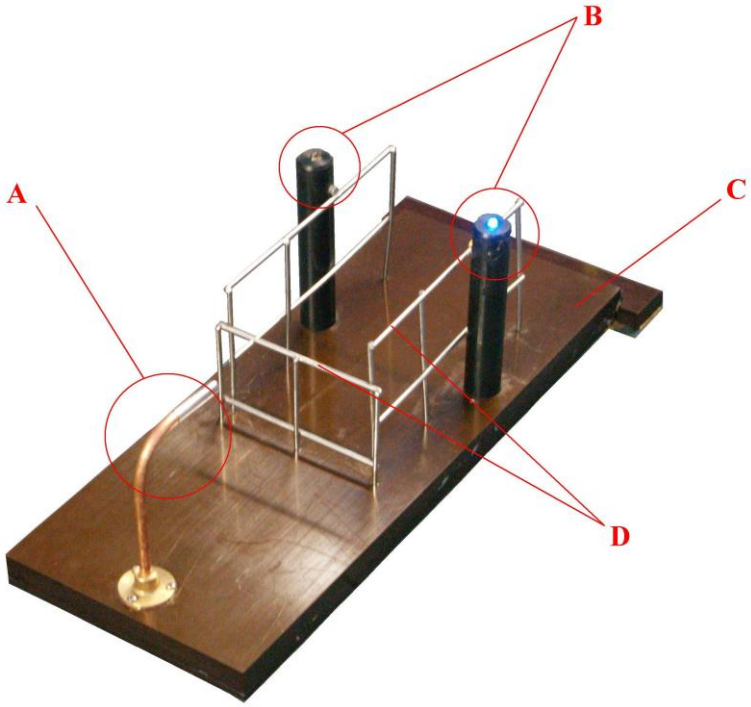

Fig 1.2.3. The central platform of the ferret arena. (A) The central spout (B) posts, when lit the LED indicates the beam has not been broken (C) the platform base (D) the fence, used to encourage desired body position.

Behavioural sessions were controlled by a custom made Visual Basic program (SITAFCv.2) run on a Viglen PC (Pentium 4 4GHz CPU, 1GB RAM) the PC served as the hub to the network (Fig 1.2.4). The PC, controlled by SITAFCv.2, communicated with the hardware in the arena via an audio interface system (MOTU 24 I/O). A trial was triggered when both the infrared beam between the fence posts (B in Fig 1.2.3) and infrared sensor in the centre spout (A in Fig 1.2.3) were broken. All responses were sent to the PC via the MOTU and upon the correct conditions being met the signal would be sent to the relevant module, again via the MOTU. At this point the animal would respond by licking at one of the peripheral spouts, the response would be sent back to the PC and logged. The process would then start again.

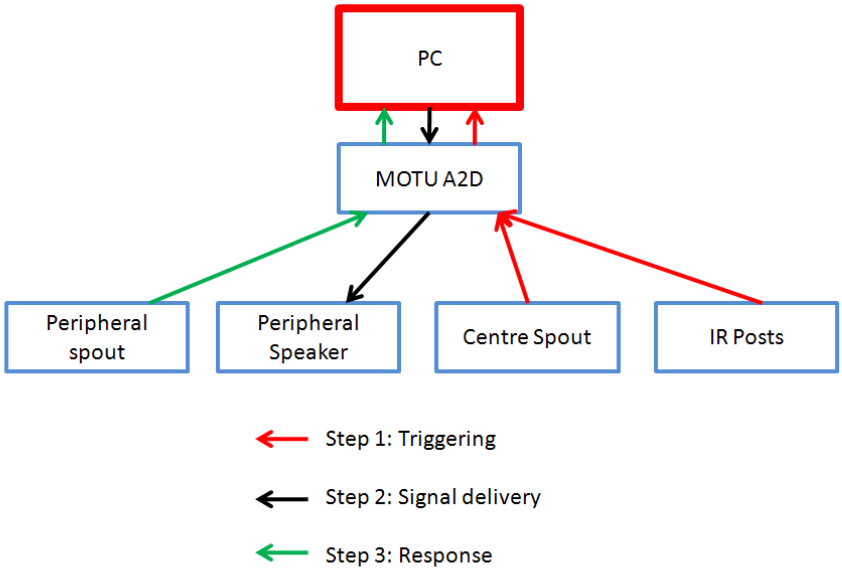

Fig 1.2.4. Schematic of the hardware used to collect animal behavioural data. Red boxes indicate the hardware outside of the sound attenuating chamber. Blue boxes indicate equipment inside the sound attenuating chamber. Arrows indicate the flow of information and the colour indicates the chronology of the flow.

1.2.2.3 Stimuli

Three different stimuli types were used in this study: broadband noise, band-pass noise and pure tones, each were generated differently. All noise stimuli were generated in Matlab, stored as a file and then accessed and presented via

the SITAFCv.2 behavioural software. The broadband noise file was a 30 second 1-48kHz flat noise-band. Two band-pass noise files were generated: one ½ octave and one 2 octave band-pass noise files centred at 10kHz. To create the band-pass noise files the broadband noise file was filtered, in Adobe Audition 1.0, using a Butterworth filter to the order 18, this created a roll off of 108 dB per octave. The filter edges were determined by an octave step of half of the bandwidth required e.g. for a half octave band at 10 kHz the lower edge was 8408 Hz and the upper edge 11892 Hz. 10kHz pure tone signals were generated online by SITAFCv.2. Signal durations and levels were controlled by the SITAFCv.2 behavioural software dependent on task demands. When noise durations were less than 30 seconds a random segment, of the required duration, was taken and gated with on and off with 10ms $\cos^2$ ramp. Pure tone signals were also gated with a 10ms $\cos^2$ ramp.

1.2.2.4 Task design

1.2.2.4.1 Localisation paradigm

To test localisation and train animals at sound localisation a simple paradigm was introduced. First the animal would trigger a trial by licking at the spout on the central platform and a water reward given. Upon triggering a trial a sound (broadband noise randomly roved between 65 and 75dB SPL) would be presented, in pseudo-random order, from one of the 12 surrounding speaker modules. The animal would then respond by licking at one of these modules. A water reward would only be delivered if the animal responded at the correct location. At this point the sound was terminated and the trial ended, the process was then repeated.

1.2.2.4.2 3-Location discrimination paradigm

As with the localisation paradigm the 3-location discrimination paradigm used was an approach-to-target task. A continuous broadband masker was presented from directly in front of the central platform (0˚). The animal triggered a trial by licking at the spout on the central platform. Upon triggering a trial the pure tone signal was delivered from one of three surrounding speaker locations: direct left (270˚), right (90˚) or straight ahead (0˚). As before the animal would then respond by licking at one of these modules, correct answers were rewarded, any response terminated the sound presentation and the process was then repeated.

Animals were tested using three different stimuli: pure tone signals, 1/2 octave band-and two octave band-pass noise, centred at 10kHz. The masker level remained the same for each of the conditions run for each animal. The signal level was varied using a method of limits approach. The signal level started at a high level to elicit good (≈90% correct) performance. For each behavioural session the signal level remained the same throughout the session. When the animal had performed consistently for two sessions (performance on both sessions within 5% of one another) the signal level was lowered by 5 dB for the following session and the process repeated. Once performance had fallen to close to d'=0 the behavioural block was ended. This process was repeated for all 3 signal conditions.

1.2.2.5 Training

1.2.2.5.1 *Shaping*

Ferrets were first trained to approach and lick the centre spout in the absence of sound to teach them how to trigger a trail. First the ferret was rewarded every time it approached and licked the centre spout, until the ferret regularly

left and then returned to the centre spout. Then, to encourage, the correct body position (the ferret entering the fence and facing the spout) reward was only delivered when increasingly closer approximations of good body and head position were achieved. This continued until reward was only given when body position was deemed acceptable.

1.2.2.5.2 Localisation training

Once "shaping" was complete animals were trained to localise sound sources. The animals were already trained to lick at the centre spout and this behaviour was used to trigger a trial. Upon triggering a trial a 30 second broadband signal was then presented from one of the 12 surrounding speakers. The animal did not need to be trained to approach the target module as they instinctively approached the correct location. This step of the training was simple, once the ferret had been shaped training was passed over to SITAFCv.2 and training parameters were entered.

To ensure that the ferret had enough time to approach the correct module, the stimulus duration was set at 30 seconds at the start of training. Animals were given 10 drops of distilled water at the centre spout for triggering a trial and then 10 drops for correctly localising a sound. Giving so many drops at the centre spout is not ideal as it means that half of the reward for each session is given for triggering a trial, this means that less behavioural data can be obtained. Also the ferrets continued licking as long as they were receiving water, for shorter duration signals the noise generated by this licking could mask the signal presented. For these reasons the number of drops delivered for triggering trials was systematically reduced (from 10 to 1) over 2-5 sessions at a rate dependant on performance. Next the probability of a reward for triggering a trial was reduced (from 1 to 0.1) again over several trials, dependant on performance. Animals would generally stop behaving once they had received a certain number of drops of water (females generally require less water). To

ensure that a minimum of 100 trials could be reached in each behavioural session the reward at the peripheral spouts was adjusted.

Once a satisfactory reward balance had been met animals were trained for ever decreasing signal durations. The duration was reduced when performance reached and remained at >90% for a minimum of three behavioural sessions. Animals were considered trained when they could localise one second noise bursts at >=85% correct. After this noise duration was reduced until the animals could localise 0.2 sec noise bursts with a consistent level of performance, i.e. roughly equivalent performance in two consecutive sessions. Once a ferret had been trained and tested on the localisation task they were then trained on the 3-location discrimination paradigm.

1.2.2.5.3 3-Location discrimination training

First animals were acquainted with the reduced number of response speakers and the noise masker. Animals were trained to localise from 3 locations (0˚, 90˚ and 270˚) for 4 sessions using 500ms noise bursts target stimulus, in the presence of a continuous broadband noise masker located at 0˚. At this point the target stimulus was changed to a 5 second, 10 kHz pure tone signal. Signal duration was then reduced to 2, 1 and then 0.5 second durations when performance of >90% correct had been reached on two consecutive sessions. Generally only two sessions at each of these durations was necessary.

1.2.2.6 Data analysis

1.2.2.6.1 Threshold estimation and fitting psychometric function

In order to gain an estimate of the threshold it is necessary to interpolate values at a given performance level. In order to interpolate these values researchers often fit sigmoidal psychometric to the data using a least squares minimisation

approach. The psychometric functions here varied in shape with some being linear under some conditions and others being sigmoidal in shape for other conditions. In order to avoid assuming the shape of the function or comparison of thresholds derived using different types of fitting algorithm all thresholds were linearly interpolated between the two nearest data points using intervals of 0.01 along the y-axis and the SNR at the threshold value taken.

1.2.2.6.2 Centre of mass (COM) analysis

Centre of mass (COM) gives an indication of both accuracy and size of error and allows the use of standard multivariate analysis of variance (MANOVA). To calculate COM values the response angles are simply converted into Cartesian co-ordinates. This yields a two dimensional representation of each data point from which means and standard deviations can be calculated. One problem with traditional localisation data analysis techniques is that the representation used is often not circular i.e. for instance often responses are considered on a continuum from 180˚ to -179˚. This creates an artificial break in the data (the boundary at 180˚) where a one degree change in response location, from 180˚ to -179˚, results in a recorded difference between the two responses as 359˚ as opposed to just 1˚. COM uses sin and cos (circular) functions therefore a shift in angle, as mentioned above, would result in a 1˚ change.

Consider localisation of a signal originating from 0˚, when 100% correct is achieved the mean x value will be 0 and the mean y value will be 1. One might expect localisation errors of an unbiased ideal participant to be symmetrical about the sound origin, though probably more densely concentrated close to the origin than at more disparate locations. Therefore with unbiased performance as the accuracy decreases, and errors increase, the value of x will remain 0 and the y value will decrease. Thus points close to the origin represent poorer accuracy. Conversely if the accuracy was at some sub 100% correct value and the errors were asymmetrical then the value of x would change in the

direction of the more concentrated errors. Roughly speaking this means that distance from the origin predominantly reflects localisation accuracy and the distance tangential to a straight line from origin to the unbiased 100% correct performance (e.g. x = 0, y = 1 for 0˚) predominantly reflects localisation bias.

1.2.2.6.3 Signal detection theory

In psychophysics Signal Detection Theory (SDT) is a means of quantifying the ability to discriminate between different signals. SDT's aim is to measure the acuity of an (assumed) underlying sensory representation removing interfering decision factors such as bias towards certain response choices (Macmillan, 2005). It does this by seeking to separate the 'sensitivity' to a variable of interest from a response 'bias' component which is unrelated to the stimulus. SDT is often used in simple yes-no tasks as well as categorisation/identification tasks (Macmillan, 2005). It can also be used in any forced choice task, and is appropriate in any situation where there are concerns about response bias. Horizontal localisation can be considered an identification task, where the listener's task is to identify the source of the sound. The simplest example of this is a left/right localisation task (essentially a left/right discrimination task).

In a forced choice task one must make a response. In SDT terms one must decide whether the perceived sensory variable falls into one category or another (i.e. did the sound originate from the left or the right). When a signal comes from the left there is an associated likelihood that the sound will also be perceived from the left, this can be represented by a distribution (Fig 1.2.5). In a two alternative situation the listener needs to choose a point along the decision axis (the abscissa in Fig 1.2.5) that separates the two possible locations, i.e. the "decision criterion" (vertical line in Fig 1.2.5). As we are interested not just in the perceived location of the sound but also the accuracy of this perception we must disambiguate when the perceived location was correct and incorrect. Stimuli can be perceived as coming from the left: either correctly,

true left (TL), or incorrectly, *false left (FL)* or from the right: correctly, *true right (TR)*, or incorrectly, *false right*.

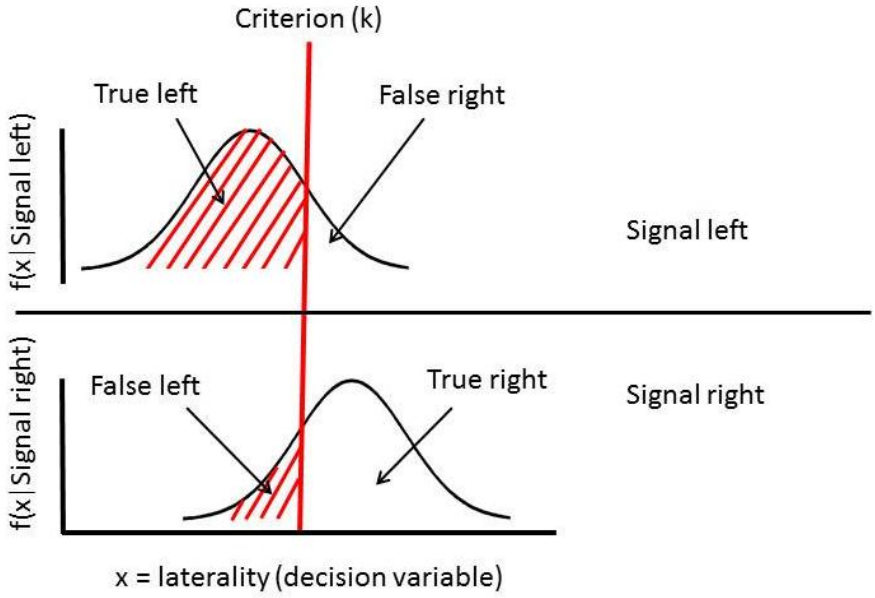

Fig 1.2.5. Underlying distributions of laterality for sounds originating from the right and left. The top curve shows the distribution of a sound originating from the left, values above the criterion k lead to false alarms. The bottom curve shows the distribution for sounds originating from the right, values above the criterion k lead to hits.

Two factors will determine the proportion of true left, false left, true right and false right responses. The first of these is the amount of overlap between the two distributions. If we hold the decision criterion and moved the two distributions closer together it would result in fewer correct (true left and right responses) and more errors (false left and right responses). In other words we have changed the sensitivity of the listener. Secondly, if we hold the position of the distributions and moved the decision criterion we would also change the proportion of true left, false left, true right and false right responses. Notice though that the sensitivity does not change (the distributions remain the same distance apart) we merely alter the point at which we make our classification. We have changed the bias of the listener.

We are interested in a measure free of bias, SDT provides such a measure: the distance between the means of the two distributions, known as *d'* (pronounced

'dee-prime'). We will now make the first of two assumptions: the first is that these are normal distributions. The second is that the two distributions have the same standard deviation. Under these assumptions the optimal decision criterion is precisely between the means (peaks) of the two distributions and changes in the distance between the two means will reflect changes in sensitivity.

Therefore to calculate the *d'*, we calculate distances between each respective mean and the criterion. For the left-signal distribution this is the z-score of the proportion of true left responses, or z(TL), where the sign indicates the direction from the mean to k. For the right distribution this is the z-score of the proportion of true right responses or –z(TR), the negative sign indicates the direction, from mean to k, whereas we want the direction from k to the mean z(TR) as this is in the same direction as z(TL) . As z(1-p)=-z(p) we can also write z(TR) as –z(FL). Therefore the distance between the two means (*d')* can be written as:

$$d' = z(TL) - z(FL) \qquad \textbf{eqn. 1.2.1}$$

In a more complicated case we might wish to measure differences in detectability between a number of stimuli. For a 3-location discrimination task the only change necessary is to add another distribution to our underlying distributions. Fig 1.2.6 displays how this could be represented for three sound sources left, middle and right (270^o or S_1, 0^o or S_2 and 90^o or S_3). To gain the *d'* we simply need to calculate the *d'* for each pair of distributions. From this we can also calculate the total *d'*, which in this case is the discriminability between left and right, by summing all of the *d'* values.

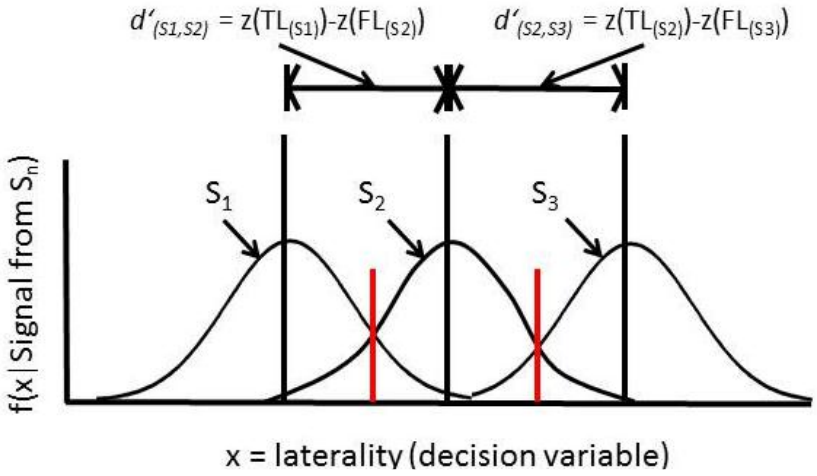

Fig 1.2.6. Underlying distributions for a 3-location discrimination. The left most distribution (S_1) represents a sound source on the left, the middle distribution (S_2) a sound source in the centre and the right most distribution (S_3) a sound source on the right. From the sensitivity between each source pair can be calculated as before (equations above the distributions). The total *d'*, the sum of the two equations, is the discriminability between the left and right sound sources.

As mentioned previously, *d'* gives a bias free measure of sensitivity but cannot be directly compared with data reported as % correct. To make these comparisons we require the data to be in % correct form but to also have the bias removed. To do this we simply set the criterion (*k*) to be precisely halfway between the means of the two distributions, the optimal point for the criterion (Macmillan and Creelman, 2005). This can then be converted back into a p(c) based value, yielding the bias free proportion correct measure $p(c)_{max}$:

$$p(c)_{max} = \Phi\left(\frac{d'}{2}\right) \qquad \textbf{eq. 1.2.2}$$

Confidence intervals can be estimated by adapting the variance estimate from Gourevitch and Galenters (1967) G statistic:

$$Var(\widehat{d'}) = \frac{P_{TL} - (1 - P_{TL})}{n_L(ordZ_{TL})^2} + \frac{P_{FL} - (1 - P_{FL})}{n_R(ordZ_{FL})^2} \qquad \textbf{eqn. 1.2.3}$$

Where P_{TL} and P_{FL} are the observed true left and false left probabilities, n_L and n_R are the number of trials presented to the left and right side respectively. The values $ordZ_{TL}$ and $ordZ_{FL}$ are the height of the normal density function at the inverse-normal transforms of P_{TL} and P_{FL}. Assuming that the sampling distribution of *d'* is normal this variance estimate can be used to find significance at any criterion, for example the 95% confidence is the interval spanning 1.96 standard deviations on either side of *d'*.

1.2.3 Results

1.2.3.1 Localisation training

Before behavioural 3-location discrimination testing began the animals were trained to localise sounds. A number of signal durations were used and the training criterion was such that animals had to complete two consecutive sessions at >90% correct for one second broadband noise. Fig 1.2.7 presents the % correct performance, at each test location for a one second noise burst stimuli. Performance at all 12 speakers was similar. Percentage correct did not vary beyond ±5% for each speaker location for three of the four animals (F:2-4). At posterior signal locations (150-210˚) F:1 demonstrated poorer performance, with performance below 80% correct at 180˚.

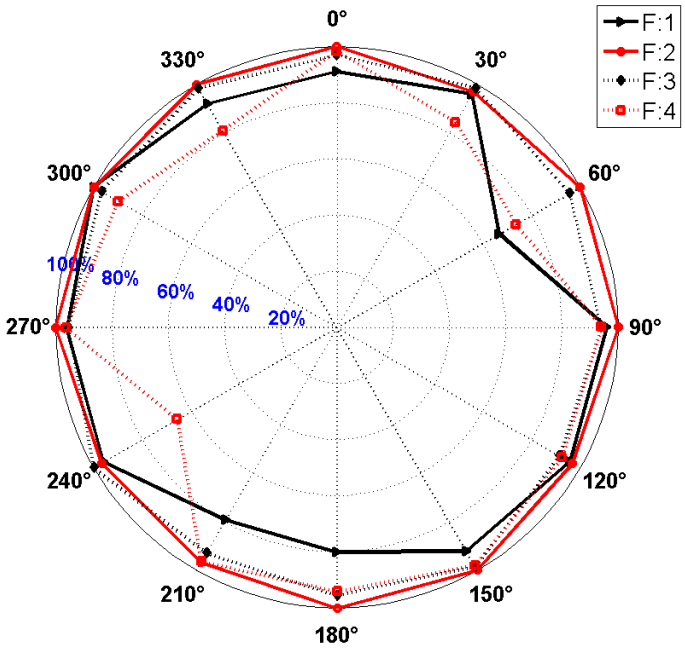

Fig 1.2.7. Localisation of one second noise bursts by four ferrets (F: 1-4). 12 locations were tested (labelled around the circumference).

Fig 1.2.8 displays 12 speaker localisation performance for 1 second broadband noise, in the form of a confusion matrix. For three of the four animals (F:2-4) errors were very few and did not appear systematic, the mean signed errors were 3.7, 5 and 2.6˚ respectively. F:1 made a larger proportion of errors (mean unsigned error of 6˚) and these were generally focused at locations close to the sound source location (Fig 1.2.8). For all animals errors were generally at speakers neighbouring the source locations.

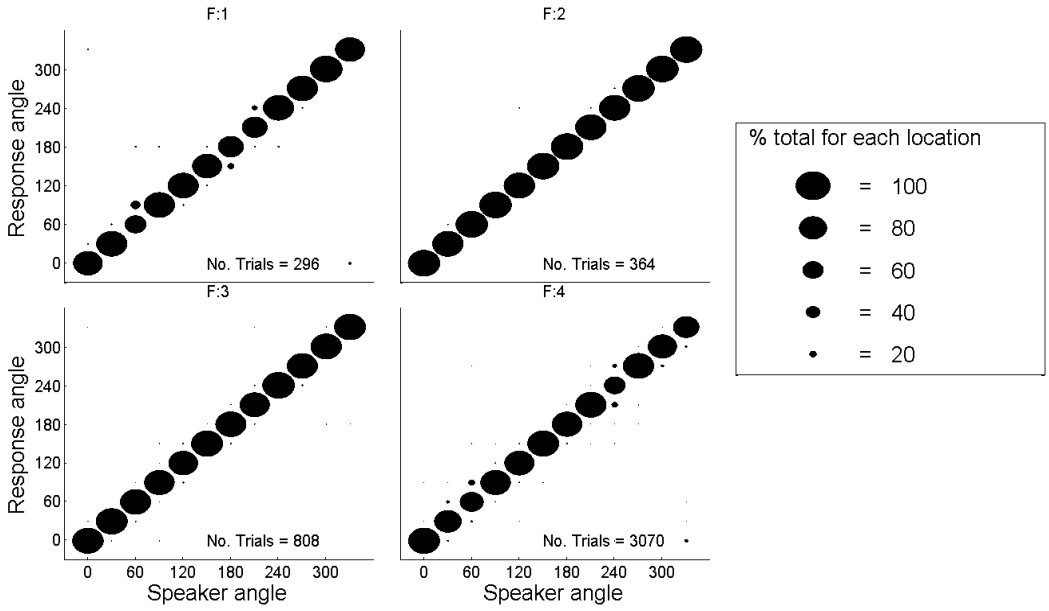

Fig 1.2.8. Localisation of one second noise bursts by four ferrets (F: 1-4). Circle size indicates % correct (see legend).

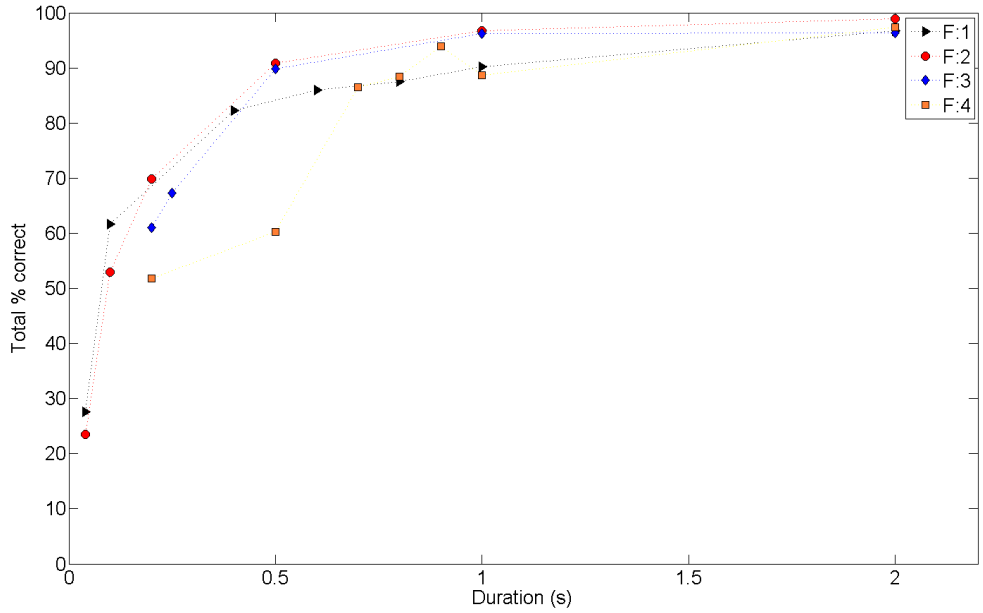

Fig 1.2.9. Localisation performance as a function of duration. Noise bursts of different durations (abscissa) were used to test localisation performance, plotted as the total % correct (ordinate) found by pooling data from all speaker locations. Legend indicates animal number, error bars are standard error.

Performance on shorter duration noises was measured once an animal had reached consistent performance on 1 second stimuli (Fig 1.2.9). For longer durations (1-2 seconds) % correct was relatively high (> 85% correct) for all animals. As the signal duration was reduced to below 1 second the % correct declined, though it remained relatively high for most animals for durations greater than 0.4 seconds (F:1-3). Not all animals were tested at shorter durations than 0.2 seconds. Those tested at the shortest durations (40ms) performed at ~50% correct.

In order to test if performance was similar at all speaker locations data was converted into Cartesian co-ordinates and analysed using COM (Fig 1.2.10). When localisation performance is high one would expect the COM to be close to the source location, therefore, to test for significant differences in the bias and accuracy between different spatial locations COM coordinates were centred on the same location (0˚). A MANOVA was then run on all of the COM data at that duration. Thus the statistical significance of differences in accuracy and bias were calculated for each duration tested across the 12 locations. As discussed in

the methods, to some extent, the accuracy and bias could be separated. With y-axis predominantly reflecting accuracy and x-axis predominantly reflecting bias for values centred at 0˚. The MANOVA itself returns the probability that the means of each location were the same for each dimension (x and y) and so it is possible to tentatively infer the role that bias and accuracy had in affecting the statistical differences found between locations. For the purposes of brevity I will refer to the two dimensions as though they were independent though when referring to, for example, the effect of rotational bias in reality it should read the "effect of the predominantly bias related dimension".

For animal F:1 performance across speakers at long durations (2 and 1 seconds) was relatively similar across locations. For 2 second noise bursts there were significant differences in the accuracy dimension (MANOVA, $p<0.05$) but these were not significant for bias. The accuracy at the rear speakers (210 and 180˚) was responsible for creating these differences (as discussed previously). For 1 second stimuli there was no significant difference between accuracy or bias for any of the 12 speakers. At most shorter durations (0.8, 0.6, 0.1 and 0.04 seconds) there were significant differences in performance at all 12 speakers in both bias and accuracy. With bias being particularly large at 60, 210, 240 and 270˚ at shorter signal durations. Below 0.2 seconds bias was so large at 210˚ that the mean speaker location was 240˚.

For animal F:2 both bias and accuracy were not significantly different (MANOVA, $P>0.05$) at longer durations (15 and 2 seconds). When the signal duration dropped below 2 seconds there were significant differences for both bias and accuracy at all shorter durations (1, 0.5, 0.2, 0.1 and 0.04 seconds). At the shortest durations errors were broadly distributed and accuracy low. For animal F:3 there were no significant differences between speaker locations at durations greater than 0.5 seconds (MANOVA, $P>0.05$) for either bias or accuracy. As the signal durations was lowered to 0.5 seconds differences in bias became significant but not differences in accuracy. Bias at 120 and 150˚ accounted for

the differences that created this difference. Also at 0.25 seconds bias also accounted for significant differences between speaker locations but not accuracy. Performance at 120 and 150˚ was again responsible for these significant differences with the animal demonstrating noticeable pull to 90 and 180˚ for the respective speaker locations. At the shortest duration tested (0.2 seconds) both bias and accuracy demonstrated significant differences, with performance between 60 and 150˚ being particularly poor (~45% correct).

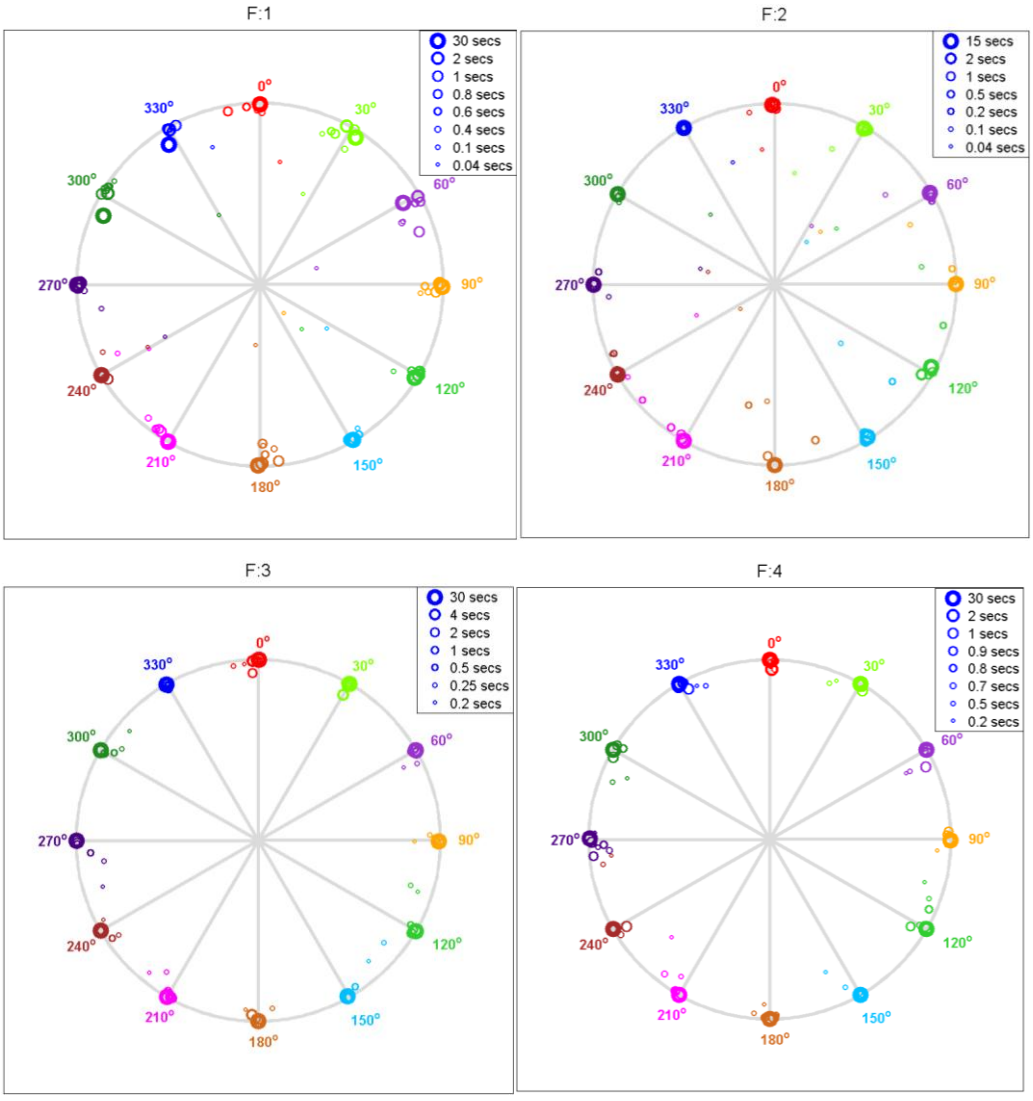

Fig 1.2.10. Localisation performance as a function of signal duration (plotted using Centre Of Mass, COM, see methods and text). Marker colour indicates the sound origin angle, marker size indicates sound duration (see legend).

For animal F:4 most longer durations (30, 1 and 0.8 seconds) demonstrated no

significant (MANOVA, p>0.05) differences between speaker locations. For 2 second duration stimuli bias at 0˚, pulling toward 30˚, created this significance despite bias and accuracy at all other speaker locations being relatively well matched. For 0.9 second stimuli bias at 90˚, pulling toward 60˚, created this significance and again performance at all other speakers was well matched. Below 0.8 seconds significant differences were found in both bias and accuracy for all durations tested (0.7, 0.5 and 0.2 seconds). These differences were due to variability in performance at a number of speaker locations as opposed to specific performance at one speaker location.

1.2.3.2 3-Location discrimination

Once animals were trained to criterion on a generalised localisation task they were then trained on a three-location discrimination task. The number of speaker locations was reduced, from 12 to 3 (0, 90 and 270˚)., a broadband masker was presented from directly in front of the animal (0˚)and the test signal was changed from a broadband to a pure tone signal (10kHz). Reducing the bandwidth of a signal leads to a reduction in localisation ability (Pierce, 1901; Terhune, 1974; Butler, 1986; Wightmann and Kistler, 1992; King and Oldfield, 1997; Eberle *et al.*, 2000) but the reduction in the number of speakers was intended to compensate for this.

Data were collected using the method of limits (see section 1.2.2.4.2). In brief, animals were tested in behavioural blocks, lasting ~2 weeks, with two test sessions per day. At the beginning of each block the signal-to-noise ratio (SNR) was high (>25 dB SNR) and it was gradually reduced. Once an animal had maintained a comparable performance level (within 5% correct) for two consecutive sessions the signal level was reduced by 5dB. Data were collected with a view to displaying the results in *d'* a Signal Detection Theory measure of performance (see section 1.2.2.6.3). Therefore a behavioural block was ended once performance was lower than threshold in *d'*. Deciding on the actual *d'*

value to use as threshold posed a problem because a subsequent step in this investigation was to test a number of data collection methods, including an adaptive staircase method. In addition these data were also to be compared with traditional signal detection data. The previously collected data had used an adaptive staircase method which approximated threshold around 70.7% correct, which equates to a *d'* = 1.0893. To ensure the data were comparable thresholds had to be set at this level and were collected to fit this level of performance.

Fig 1.2.11 displays 3-location signal-in-noise discrimination performance for the 4 test subjects. As expected *d'* increased with SNR for all 4 animals, demonstrating the dependence of localisation ability in the ferret. Thresholds were estimated using linear interpolation (see section 1.2.2.6.1). With this criterion across repeat mean thresholds were 9.3, 11.4, 7.5 and 7.6 dB SNR for the Left/Centre conditions for each ferret respectively (F1-F4). The corresponding thresholds for the Centre/Right conditions were 13.9, 23.7, 6 and 11.2 dB SNR. The means, across animals, were 9 and 13.7dB SNR for the Left/Centre and Centre/Right conditions respectively. This performance was poorer in the Centre/Right condition was reflected in three of the four ferrets. Three of the four animals (F1-F3) were tested on more than 1 behavioural block.

To test the reproducibility of these data the *d'* values at each SNR for each animal were compared across repeated behavioural blocks. Confidence intervals were overlapping in 61/70 measurements (the stars at the top of the plots in Fig 1.2.11 indicate non-overlapping confidence intervals), suggesting that the measurements obtained were reliable over time. Data were therefore pooled across behavioural blocks for subsequent data analysis in this chapter.

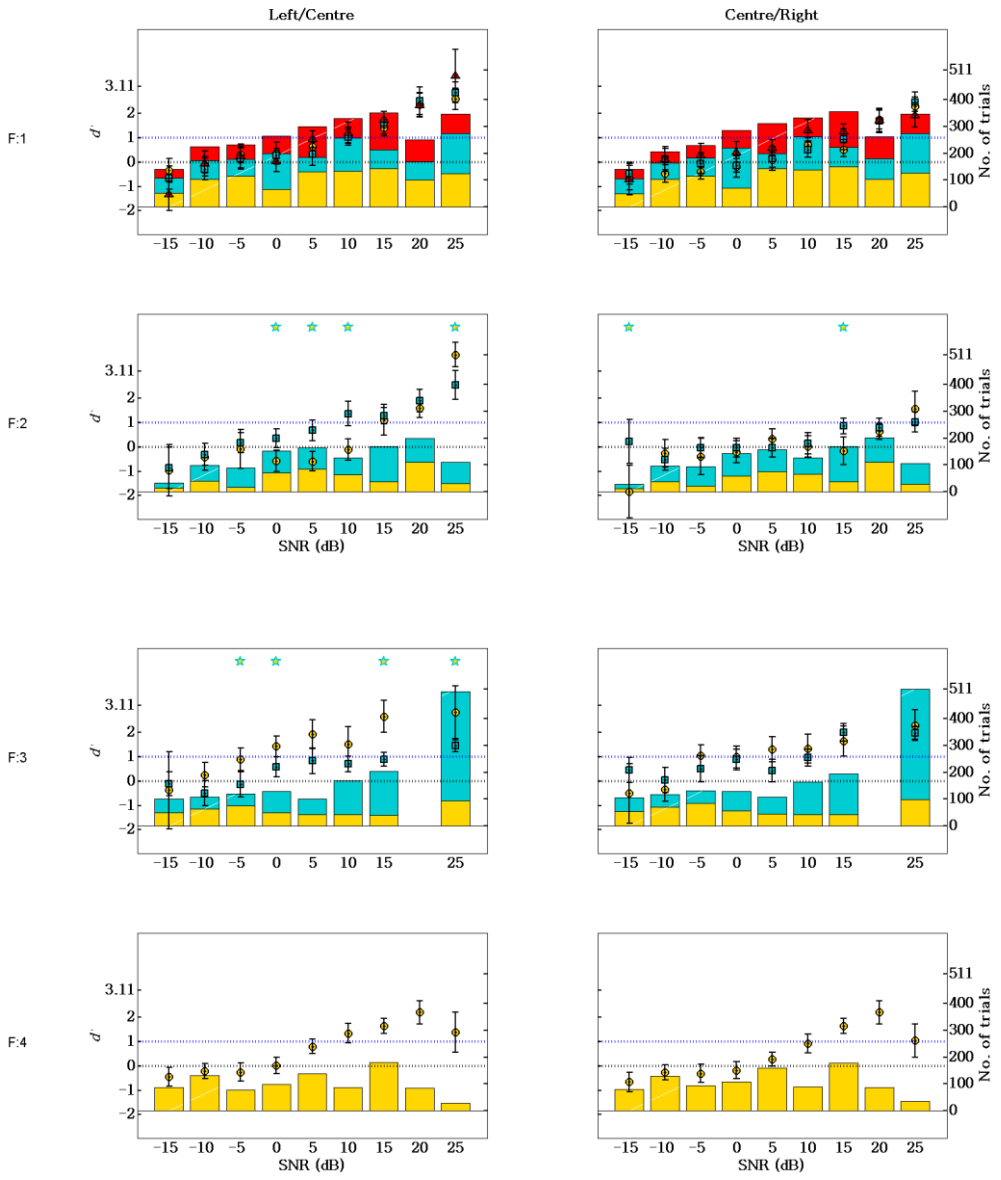

Fig 1.2.11. The effect of SNR on 3-location discrimination performance. SNR (abscissa) was systematically reduced for four ferrets (from top to bottom F:1-4) and performance between two speaker pairs (left/centre and centre right) measured using *d'* (see methods). Markers indicate *d'* values, error bars denote 95% confidence intervals and bars display the number of trials for each *d'* measured.

At high SNRs (~> 15dB) *d'* was significantly greater than 1 for the majority of data points (24/32). For a threshold of *d'* = 1 this would suggest a high SNR is required to reliably distinguish tones. At low SNRs (<-5 dB) *d'* was significantly less than chance (*d'* = *0*) for a number of data points (11/30). A negative *d'* is obtained when the proportion of false alarms is larger than the proportion of

hits. At chance performance one would expect data points to be either at *d' = 0* or evenly distributed around this value. This suggests at low SNRs responses become skewed toward the incorrect speaker (sounds from the left are assigned to coming from the centre and sounds from the centre as coming from the right).

To further investigate this data were plotted using the COM co-ordinates. The reduction of stimulus conditions changes the interpretation of these plots from those displayed previously. Evenly distributed errors for a given location would previously fall on a line between the between the speaker location on the perimeter and the centre of the COM plot. This is because errors could be evenly distributed either side of signal location. However, for the left speaker (270˚) errors could only come from the right (0˚) or directly opposite (90˚) and vice versa for the right speaker. This meant it would be difficult to use the skew of data points to understand any skew in the measured data.

In order to aid interpretation, data were simulated in which the probability of a correct response was held and the probability of an error at a given location was evenly distributed between the two remaining locations, thus estimating a zero skew condition. Deviations from these modelled data in a particular axis imply skew toward that response location in that particular axis. For example in Fig 1.2.12 (top left) we can look at the modelled data for 90˚, grey triangles, and compare it with the actual data, orange triangles. If points of modelled and actual data of a given level coincide then the actual data is free of any skew. This, however, is not the case and for a number of signal levels the actual data points are above, larger y values than, the modelled data demonstrating that the data is skewed toward 0˚. For lateral locations (90 and 270˚) points above the modelled data imply skew toward the centre and points below (or toward the opposing side) imply skew toward the opposing lateral location. For the central condition points either side of the modelled data demonstrate skew in that direction. One would expect some skew in these data as when errors are

made they are expected to be located close to the signal when the listener has some idea of the origin of a sound. If errors are evenly distributed across response locations one could assume that the listener did not have a clear sense of the origin of the sound and responded at random.

The skew of the data appeared to be smaller for sounds originating from 270˚, evident from the similarity in position between the modelled and actual data in Fig 1.2.12. (for example; top left and right, in many cases the grey and purple squares are coincident). Whereas for sounds originating from 0 and 90˚ data points appeared to be further from their associated model data points. A manova was performed on the COM coordinates of the actual and modelled data in order to test for significant differences. For each of the 4 animals (F:1, F:2, F:3 and F:4) 9 SNRs were tested (25 to -15 dB SNR in 5dB steps). At the 270˚ location 19/36 data points were significantly different from the modelled data ($p<0.05$). This is comparable to the 0˚ condition (20/36) and the 90˚ condition (22/36) suggesting that source location did not increase the likelihood of finding significant skew.

One might expect at high SNRs there would be significant differences between modelled and measured data as while the precise location might not be clear the sound could be localised to a rough area and hence mistakes would be located at adjacent speakers. As the signal level is reduced and performance gets closer to chance errors may become more evenly distributed. There is some evidence for this as at high SNRs (20-5dB SNR) 35/48 data points were significantly different ($p<0.05$), whereas only 20/48 were significantly different at low SNRs (0—15dB SNR).

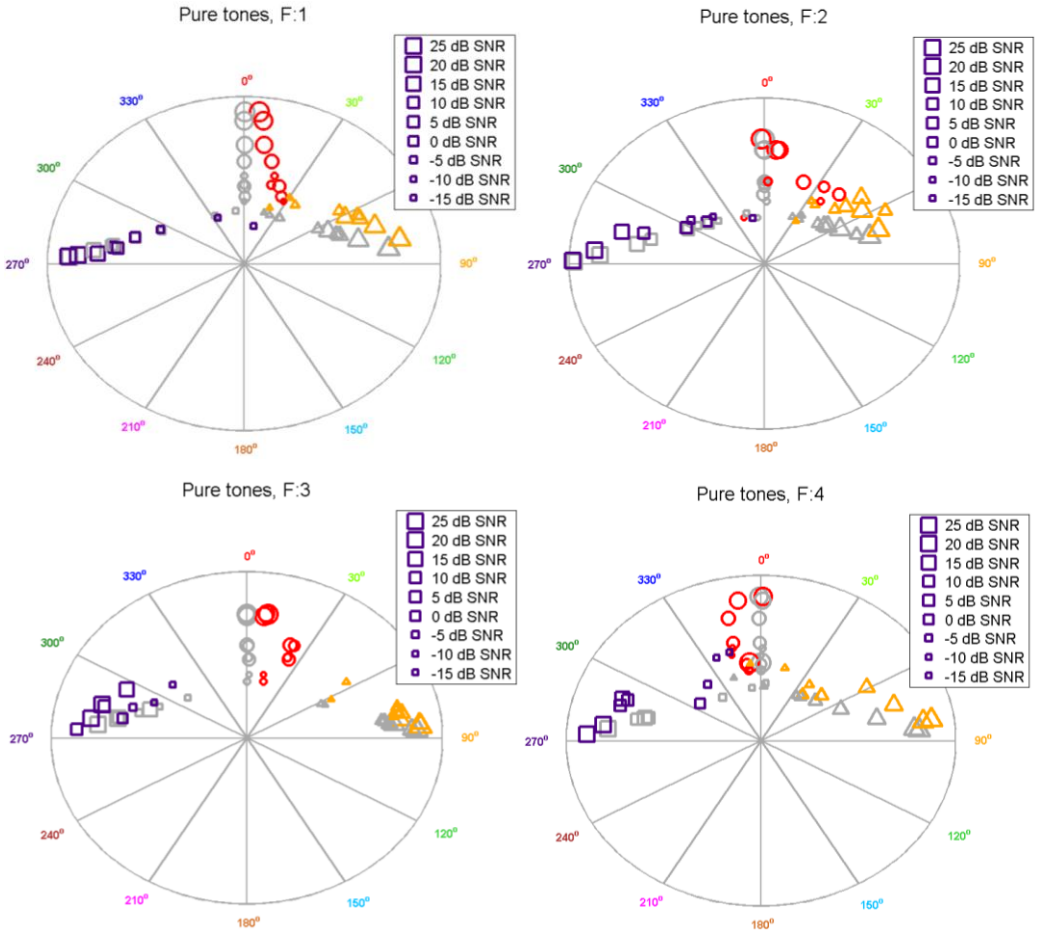

Fig 1.2.12. The effect of SNR on 3-location discrimination performance of pure tones with performance represented in COM coordinates. Marker shapes indicate the source location (squares - 270˚, circles - 0˚ and triangles - 90˚). Measured data are distinguished by colour (purple - 270˚, red - 0˚ and orange - 90˚). Skew-free data with equivalent % correct performance were also plotted in grey. Symbols at the denoted source location and on the perimeter of the circle reflect 100% correct performance. The distribution of errors can be inferred from the deviation of the symbols from the modelled skew-free data.

	270				0				90			
	F:1	F:2	F:3	F:4	F:1	F:2	F:3	F:4	F:1	F:2	F:3	F:4
25	0.17	0.00	0.01	0.13	0.00	0.74	0.00	0.12	0.00	0.07	0.00	0.07
20	0.07	0.00	0.00	0.71	0.00	0.39	0.10	0.95	0.00	0.00	0.05	0.23
15	0.00	0.00	0.00	0.00	0.03	0.33	0.15	0.00	0.00	0.00	0.00	0.01
10	0.02	0.02	0.00	0.03	0.01	0.07	0.00	0.04	0.00	0.01	0.00	0.01
5	0.00	0.57	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.03	0.26
0	0.01	0.82	0.50	0.69	0.00	0.00	0.00	0.72	0.00	0.10	0.00	0.94
-5	0.04	0.81	0.46	0.43	0.16	0.96	0.00	0.82	0.01	0.25	0.01	0.45
-10	0.86	0.55	0.16	0.00	0.01	0.02	0.02	0.01	0.01	0.16	0.00	0.04
-15	0.20	0.75	0.18	0.01	0.14	0.64	0.19	0.03	0.83	0.91	0.29	0.12

Table 1.2.1. Probability measured data points differed from modelled data points. Values are rounded to the nearest two decimal places. Orange background indicates significance ($p<0.05$)

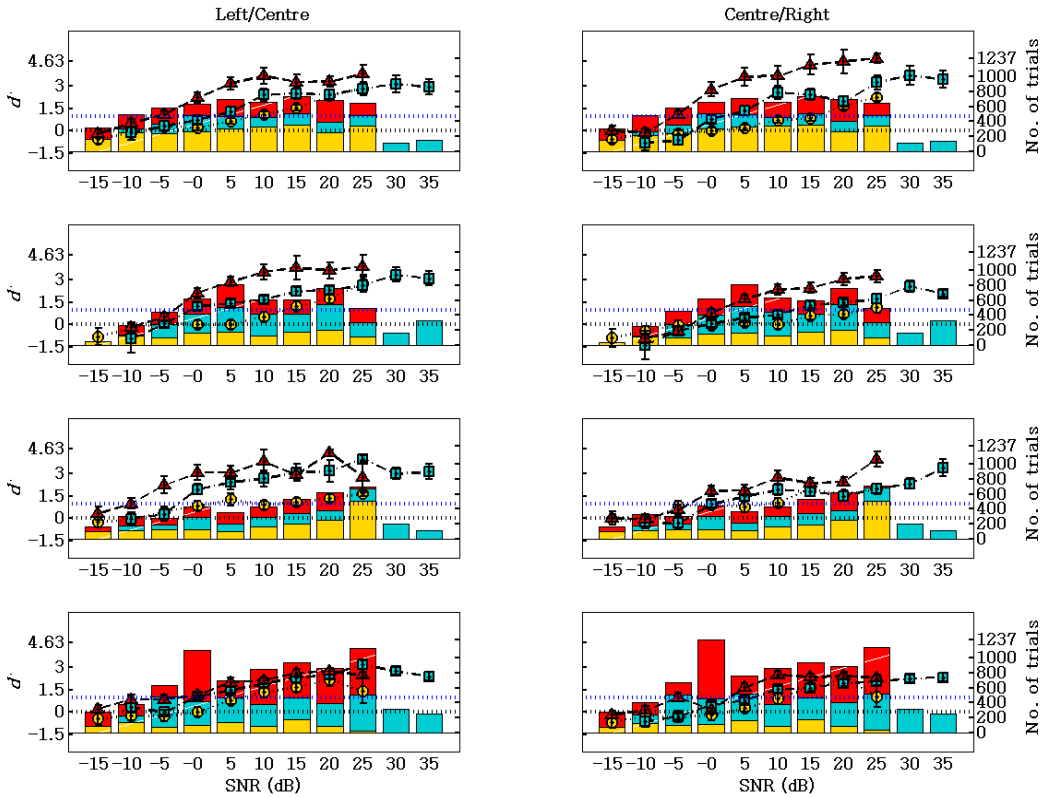

Fig 1.2.13. The effect of SNR and bandwidth on 3-location discrimination performance. Each row represents a subject with subject identifiers given on the far left. For each subject discriminabilty between two speaker pairs, Left/Centre (270 and 0˚, left column) and Centre/Right (0 and 90˚, right column), was measured. SNR (given on the abscissa) was varied for three signal bandwidths centred at 10kHz: pure tones (yellow), ½ octave flat noise band (blue) and a 2 octave flat noise band (red) and corresponding *d'* values calculated (markers and lines). The number of trials completed for each condition and at each SNR are given in the underlying bar charts.

Collecting 3-location discrimination data using pure tone stimuli highlighted some issues. Firstly the thresholds collected (mean thresholds of 2.08 and 6.87dB SNR, for left/centre and centre/right, respectively) were higher than expected. Secondly there was a general trend for confusion between the centre and right locations, evidenced by the skew toward each other in the COM data (Fig 1.2.12). This was, presumably, the source of the asymmetry in thresholds between the left/centre and centre/right thresholds. The task relies on the listener's ability to both successfully hear the sound and to then successfully localise it. For the pure tone data alone it is not possible to distinguish which variable impacted on the performance observed. It may have been the animals were unable to hear the pure tone stimuli and hence could not localise it (the

desired result) alternatively it could be that the animals could hear the stimuli but were unable to localise it. Detection of pure tone signal is believed to occur within an auditory filter (Moore, 1996). Widening the bandwidth of the signal to beyond that of an auditory filter while holding the overall signal level reduces the SNR within the original filter and essentially spreads the signal power across multiple filters, this has the effect of reducing detectability. Simultaneously widening the bandwidth of a signal improves localisation performance. Therefore, widening the bandwidth and keeping the spectrum level constant would be expected to improve performance if it was limited by localisation and impair it (further) if it reflected detectability.

Fig 1.2.13 shows 3-location discrimination performance using three different signal bandwidths: pure tones, ½ octave and 2 octave noise bands (all centred at 10kHz) for each animal (top to bottom: F:1-4). For each animal the *d'* values were generally graded, where the widest bandwidth condition (2 octave) led to the largest *d'* value at each SNR and the narrowest bandwidth condition (pure tone) led the lowest *d'* values. Another feature of the wider band stimuli was to yield a steeper psychometric function for most animals. For example for animal F:1 (top row of Fig 1.2.13) the *d'* functions for the 2 octave condition (red) were relatively flat above 5dB SNR with the majority of the change in the function occurring below 5dB SNR. Conversely the slope of the *d'* functions for the pure tone condition (yellow) changed gradually with SNR.

Thresholds in the pure tone condition were on average highest in the pure tone condition (mean across animals of 11.7 and 15.5 dB SNR) when compared to the half octave (0.1 and 6 dB SNR) and two octave (-4.8 and -0.8 dB SNR) conditions. For all bandwidths the thresholds were higher in the centre/right versus the left/centre locations, with no apparent reduction in the asymmetries in performance with increasing bandwidth. In addition the across animal standard error of means (the error bars plotted in Fig 1.2.14) were larger for each condition in the centre/right versus left/centre locations.

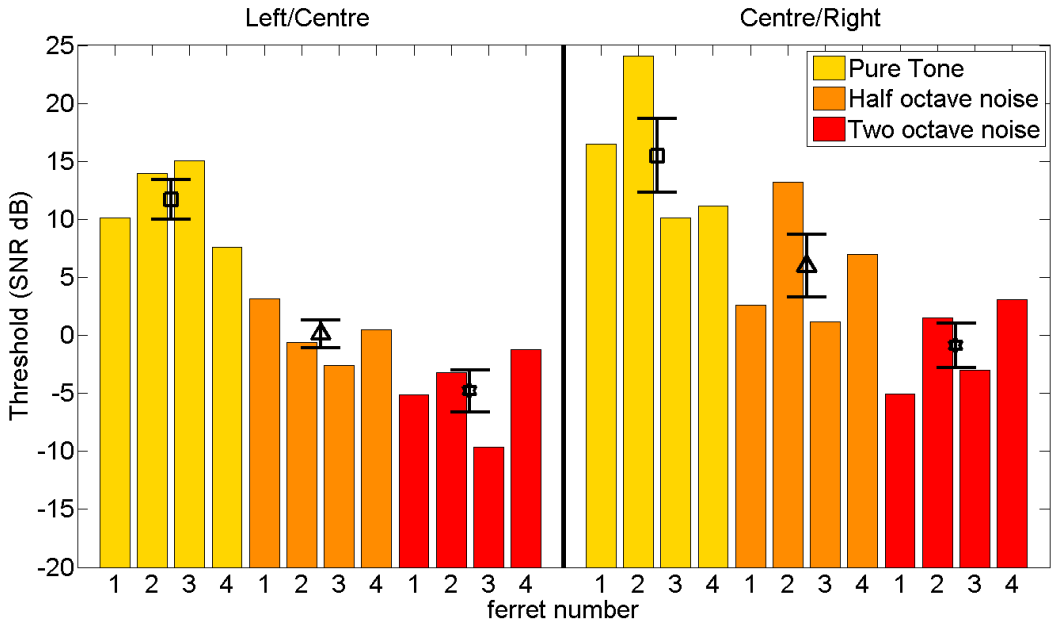

Fig 1.2.14. The effect of SNR and bandwidth on 3-location discrimination threshold. Thresholds were taken as the first SNR that fell significantly below *d'*=1. Colours indicate the stimulus bandwidth (see legend). Symbols indicate the across animal mean for each condition where squares, triangles and stars indicates the pure tone, half octave and two octave noise band conditions. Error bars are the standard error of means.

1.2.4 Discussion

Animals were successfully trained on a localisation task. For most of the animals (F:2-4) the COM for all durations remained close to the origin speaker, demonstrating that even at short durations errors were close to or evenly distributed around the source location. Previous study of ferret localisation, under similar condition, found ferrets could localise at 94% correct at one second duration (Kacelnik, et al. 2006), whereas, performance measured here was slightly less (>85% correct). For three of the animals performance, for each duration tested, was comparable (F:1-3). One animal demonstrated poorer performance at durations of less than 1 second (F:4). When excluding this animal from averages the performance can be considered similar to that found in previous studies. For instance at 500ms the average was 90% correct comparable with 86% correct found elsewhere and at 40ms the average was 52% correct compared with 59% correct . At 1 second no obvious problems

existed with the animals' localisation performance, though at shorter durations F:4 demonstrated poorer localisation performance than the other animals tested.

Human localisation acuity is superior to that of ferrets (Mills, 1957; Kavanagh and Kelly, 1987). It could be inferred from this that localisation performance in noise would also be poorer, in ferrets, when compared to humans. Studies of sound localisation in noise have shown that, in humans, performance is relatively unaffected until 0-6dB SNR for broadband signals (Good and Gilkey, 1996; Lorenzi, et al. 1999). MAAs in noise have also been recorded and demonstrate no degradation in acuity at -5 to 5dB SNR for pure tone signals (Jacobsen, 1976). As expected performance was poorer in ferrets, 3-location discrimination began to decline when the SNR was reduced from 25dB SNR (Fig 1.2.11). The mean tone thresholds between each speaker pair were estimated to be 9 and 13.7dB SNR (left/centre and centre/right , respectively). A Centre of Mass (COM) analysis was carried out on these data. Sounds originating from 0˚ were frequently confused with 90˚ and sounds from 90˚ frequently confused with 0˚. This led to 3/4 animals (F:1-3) demonstrating a reciprocal and significant skew between 0 and 90˚. This confusion between 0˚ and 90˚ led to asymmetries in performance between the two speaker pairs (left/centre and centre/right in Fig 1.2.12). It has been widely reported that detection occurs before localisation (Cherry, 1962; Egan, 1966; Jacobsen, 1976; Good, 1996; Lorenzi, 1998). But it was not clear how close to signal detection threshold these values were. It is possible these values reflected signal-in-noise detection thresholds alternatively localisation performance may have been affected at relatively high SNRs, leading to elevated threshold estimates.

In the study of signal detection it has been shown that as the bandwidth of a band-pass signal increases beyond the CB the threshold of the signal increases (Spiegel, 1981). This is also true of adding tones to a tone complex (Gäsler, 1954). Conversely increasing the bandwidth of a signal has been shown to

improve localisation performance both in silence (Boerger, 1965; Terhune, 1974; Butler, 1986) and in noise (Jacobsen, 1976; Lorenzi, 1998). The suggestion is that when increasing the signal bandwidth any improvements in performance should be due to improvements in localisation and not detection. Widening bandwidth improved performance leading to lower thresholds and steeper psychometric functions. Across animal mean thresholds were reduced by 16.5and 16.3 dB when increasing the bandwidth from pure tone to two octaves, for the left/centre and centre/right conditions respectively. In addition the slopes of the psychometric functions increased with increasing bandwidth, demonstrating that the increasing bandwidth made performance more robust to changes in SNR (Fig 1.2.13).

The improvements found with increasing bandwidth may be attributable to improvements in the localisation aspect of the task. Regardless of the source of this improvement it clearly demonstrates that 3-location discrimination of pure tone signals is far from optimal. This demonstrates a need to reduce task demand.

1.3 Two-location (left/right) discrimination in noise

1.3.1 Introduction

1.3.1.1 Considerations of using left/right discrimination

The previous findings suggest the 3-location discrimination thresholds measured the reduced ability to localise sound accurately (i.e. some aspect of localisation), rather than an inability to hear the sounds (i.e. signal detection thresholds). It was hoped that by reducing the complexity of the localisation aspect of the task, it would be possible to gain a measure closer to detection thresholds. To achieve this the number of speakers was reduced (2 speakers were used, at 90 and 270˚) and hence the spacing between the speakers was increased (from 90˚ to 180˚ separation).

One reason for carrying out a 3-location discrimination task was to have one condition where the signal and noise were presented from the same location (0˚). This would allow simple comparison of thresholds, at this location, when compared with the same data using a detection task (1-I 2-AFC). This was considered important as presentation of the signal from other locations (e.g. 90 and 270˚) could lead to effects that reduce the degree of masking and make comparison of thresholds difficult. As the proposed experiment will only use two sources at peripheral locations it is worth discussing the potential impact of this design. For instance, one well known effect is the binaural masking level difference (BMLD) where differences in the phase of a masker and signal lead to reductions in threshold (Hirsh, 1948; Licklider, 1948; Blodgett *et al.*, 1958; Jeffress *et al.*, 1962). Hence, when the masker is presented at 0˚ and a signal at either 90 or 270˚ the interaural phase difference in the two signals can lead to reduced detection thresholds. While this is certainly true at low frequencies where large BMLDs can be found, e.g. 15dB at 500 Hz, above 1500 Hz the BMLD

is only 3 dB (Durlach and Colburn, 1978). It should be noted that these results were obtained using headphones, whereas studies in the free-field result in much smaller MLDs (Saberi et al., 1991). Taken together these results suggest that at high frequencies (above the limit of phase locking at the auditory nerve of the ferret) the current paradigm would, at most, confer a 3dB advantage.

Another effect which may influence thresholds is spatial unmasking caused by the head related transfer function (HRTF). In humans sounds presented with equal amplitude and at an equal distance from the head are easiest to detect when presented from ~45˚ azimuth (Saberi *et al.*, 1991; Sabin *et al.*, 2005). This is due to the way in which sound interacts with the head and pinna creating frequency dependent amplifications and attenuations. In ferrets there are a number of noticeable differences that occur, for example differences in head size and position, shape and size of the ear and pinna. For the purposes of this experiment it is not necessary to consider the entire HRTF (or directional transfer function etc.). We are only interested in the difference in the signal level when presented ±90˚ when compared to 0˚, which is between 0 and 6dB less than the signal level at 0˚ for 10 kHz signals (Carlile, 1990; Campbell *et al.*, 2008, and raw data from Jan Schnupp). Therefore we would expect no advantage to be conferred by presenting the signal at ±90˚ when compared to 0˚, in fact we would expect an increase in threshold of between 0 and 6 dB. The ability of the ferret to perform a two-location discrimination paradigm was tested and contrasted with existing data on a traditional 1-I 2-AFC paradigm.

1.3.1.2 Optimising estimates

If the method produces comparable or lower thresholds than the 1-I 2-AFC then it would be worthwhile to refine the paradigm in order to reduce the time it takes to collect these data. Psychometric functions aim to measure performance associated with changes in strength of a particular stimulus feature e.g. changes in performance associated with changes in signal level. A

threshold is a point along this psychometric function which is equated to some limit of performance. Dependent on the aims of a study one may be interested in either or both of these measurements.

One implementation of the method of limits is to first present a signal with high signal level, where performance would be expected to be relatively high. A number of trials then follow at this signal level. Once performance is determined at this signal level the level is reduced by a set amount and performance again measured. This process is repeated until the full psychometric function has been collected, i.e. performance level falls below a predetermined threshold level. This method has been used to collect detection thresholds in the ferret (Kelly *et al.*, 1986). For the collection of thresholds this method is slow as it requires the entire psychometric function to be measured and an entire sessions worth of data is collected before a decision can be made to reduce the signal level or not. Alternatively with the method of constant stimuli a range of values are selected and then randomly presented from trial to trial. Traditionally a large range of values are selected to ensure the entire psychometric function is collected.

As a result of clinical constraints on time a number methods have been developed which aim to reduce the amount of presentations required to gain both psychometric functions and thresholds. The main focus of these methods is to reduce the set of stimulus levels presented and their repetitions by using information gained on previous trials to inform the choice of stimuli on the following trials. One such method is the adaptive staircase method. The staircase method can converge on a particular level of performance by selecting a rule which dictates the direction of change in signal strength. For instance, if the subject's response is incorrect then increase the signal strength by a set amount. If the subject responds correctly for two consecutive trials decrease the signal strength. This is referred to as one-up two-down as one incorrect answer is results in an increase in stimulus strength and two correct answers

are required to reduce the stimulus strength. This rule will converge on 70.7% correct performance (Levitt, 1971).

It has been shown in human psychophysical studies that adaptive procedures are not only more efficient but also yield lower thresholds than fixed methods such as the method of constant stimuli (Kollmeier *et al.*, 1988; Leek, 2001). While animal studies have used the staircase method to gain thresholds in animals (Evans *et al.*, 1992; Niemiec *et al.*, 1992), it is not clear whether this method produces higher or lower thresholds than other methods. This problem has been directly addressed in the mouse where it was found that tone detection thresholds were 24dB higher when using the staircase method compared to the method of constant stimuli. The method of limits has been shown to produce reliable thresholds in the ferret (amongst others) and the method of constant stimuli in other animals. These methods, however, can be inefficient, when compared to adaptive paradigms like the staircase method. Although, at present, it is unclear whether the staircase method is suitable for animal studies. In order to resolve these issues this study also aims to compare different data collection methods and identify which can be used to measure reliable thresholds quickly.

1.3.2 Methods

1.3.2.1 Training

At this point animals were trained on an approach to speaker task, initially on a 12 speaker localisation task and then on a 3-speaker discrimination task. To make the transition from a 3 to a 2-location (left/right) discrimination task animals were first acquainted with the reduction in response speakers (from three to two) by localising 500 ms noise bursts for 4 sessions in the presence of a continuous noise masker located at 0˚. After this the stimulus was then changed to a 0.5 ms duration 10 kHz pure tone. Animals were tested using this

stimulus until performance had reached >90% correct on two consecutive sessions. As the animals were already trained on 12-speaker localisation and 3-speaker discrimination generally only two sessions were required to reach this criterion. In all other respects (reward, stimulus and reward timing etc.) the conditions were as described for the previous experiments.

1.3.2.2 Experimental design

The ability of 4 ferrets to discriminate pure tones at ±90˚ degrees from midline was measured at different tone levels in the presence of a continuous noise masker, to obtain psychometric functions and thresholds. Three methods were compared, two of which were adaptive methods: the method of limits and the adaptive staircase method. In addition a non-adaptive method was also used: the method of constant stimuli. The method of limits data were collected by holding the signal at a single, initially high, level throughout the entirety of a behavioural session. Two sessions worth of data were then collected at this level, after which the signal level was reduced and a further two sessions of data were collected. This process was repeated until performance approached chance performance. For the method of constant stimuli 5 levels were selected on the basis of pilot measurements in each individual ferret. One sound level was chosen to deliver high levels of performance. Two sound levels were chosen close to, but above, threshold (the threshold gained via the method of limits) and two below threshold. These four values were at 5 dB intervals. Performance using these levels was then collected for 10 sessions. The adaptive track method was a 1-up 2-down staircase method using 3 rules. The first rule was designed to quickly approach threshold and also to test that the animal was attending to the task, a 1-up 5-down rule was used i.e. five trials correct before the signal was reduced or 1 before it was increased. The step size was set to 10 dB and only one reversal was needed before the next rule was implemented. The second rule was a 2-up 2-down design, with a step size of 5 dB for 5 reversals. The final rule was a 1-up 2-down design with 2 dB steps, this rule was

used for data collection proper and as such was used for the rest of the session. These three rules had the effect of quickly iterating toward threshold and then allowing for collection of data around threshold in the final rule.

1.3.2.3 Estimating threshold

The psychometric functions observed using *d'* were relatively linear in shape producing either a straight line of broken stick function. For this reason no fit was made to the entire psychometric function but a linear interpolation was made between measured points with 10000 points estimated between each measured data point. From this the nearest value to the threshold, *d'*=1 or 70.7% for both $p(c)_{max}$ and % correct, was taken as the threshold. The choice of 70.7% for both $p(c)_{max}$ and % correct was to allow comparison of thresholds to the 1-up 2-down adaptive staircase method which converges on 70.7% correct.

1.3.2.4 Comparing threshold estimates across data collection method

To assess which method yielded the best result with the fewest trials the data for each method were collated across behavioural sessions. The methods were contrasted by comparing threshold estimates and the standard deviation of these estimates, threshold was taken as *d=1* and estimated as set out in section 1.3.2.3. For each method the number of trials was varied (trails = 20, 50, 75, 100, 200, 300, 400, 500, 750 and 1000) and this number of trials was then split between the number of data points (SNRs). During data collection each method sampled the number of SNRs differently; the method of limits took a large portion of the psychometric function, the method of constant stimuli took 5 data SNRs and the staircase method took many SNRs but aimed to collect most of these around 70.7% correct.

In order to sample the data in a manner consistent with the way the SNRs were collected a number of considerations were made. For the method of limits the

trials were distributed evenly between all SNRs, for the method of limits the trials were distributed evenly between the five SNRs that were selected. For the staircase method the SNRs are determined not by the experimenter but by the subject. For this reason a probability distribution was created based on the actual number of trials collected for each SNR this was then used to determine how the trials were to be distributed across SNRs. Where the number of trials could not be split evenly across trials the remainder of trials was distributed using one of two methods. For the method of limits and constant stimuli the remainder of trials were distributed at random using a uniform distribution. For the staircase method the collected distribution was used. In this way the number of trials sampled for each SNR was made to reflect the actual conditions of data collection and hence allowed a comparison of how the estimates would change according to changes in the number of trials.

Once the number of trials to be sampled at each SNR was determined they were bootstrapped with replacement. This process was then repeated 500 times for each method and the mean and standard deviation of this estimate taken.

1.3.2.5 Comparing threshold estimates across behavioural paradigm

The aim of this experiment was to test if it is possible to collect detection thresholds using a left/right discrimination paradigm. To make this comparison it is necessary to collect detection thresholds in ferrets. Fortunately these data had already been collected "in house" and recently published (Alves-Pinto et al., 2012). These data were collected using a one interval two alternative forced choice paradigm (1-A 2-AFC) and using the same three data collection methods specified in section 1.3.2.2. In short, upon triggering a trial, a sound either would or would not be presented (50-50 presentation rate). If a sound is presented the ferret must respond on the right hand side (90˚) if no sound was presented the ferret must respond at the left hand side (270˚). Both a masker sound (the same as used here) and a tone (at the same frequency as used here)

were presented from directly in front of the animal (0˚). The difference between this task and the left/right discrimination task is that the location of the tone presentation (at the sides for the left/right discrimination as opposed to straight ahead) and the nature of the task. In the left/right discrimination task the animal approached the sound source whereas in the 1-A 2-AFC the animal assigned an arbitrary tag right/yes, left/no sound.

For both the 1-A 2-AFC (or Yes/No) and left/right discrimination data an adaptive staircase method was used (1-up, 2-down rule). This estimates performance at 70.7% correct, therefore to compare like for like it was necessary to compare performance at this performance level (also because this is the format the 1-A 2-AFC data were received). For this reason thresholds were taken as 70.7% correct performance (Levitt, 1971). For the method of limits and constant stimuli data it was possible to provide a bias free measure of % correct performance, $p(c)_{max}$. Therefore method of limits and constant stimuli thresholds were defined as 70.7% correct in $p(c)_{max}$. This was estimated by fitting functions to the psychometric functions in $p(c)_{max}$ and estimating threshold from this. For both data sets (Yes/No and Left/Right) a number of psychometric functions were collected in separate behavioural blocks. All data were collated and sampled with replacement taking 300 trials, this process was repeated 100 times.

1.3.3 Results

1.3.3.1 Individual psychometric functions

Psychometric functions were measured on the two-location discrimination (left/right) task (Fig 1.3.1). The method of limits was used to collect these data and two repeats were taken for each animal in separate behavioural blocks. For ¾ animals (F:1, F:2 and F:4) *d'* values appeared to decrease slowly with level. For animal F:3 *d'* remained high (above *d'*=3) at high SNRs (>5db SNR) and

decreased rapidly below -10dB SNR. For ¾ animals (F:1, F:2 and F:3) *d'* for both psychometric functions did not fall to d=1 until -15dB SNR, demonstrating the animals were still able to discriminate locations even at negative SNRs. Data points were often reproduced across behavioural blocks, evidenced by the similarity of the two functions for each animal in Fig 1.3.1 (error bars indicate 95%, two tailed, confidence intervals and stars indicate non-overlapping confidence intervals). 28/37 data point pairs measured had overlapping confidence intervals and hence were considered similar. Due to the similarity in psychometric functions the data were pooled across behavioural sessions.

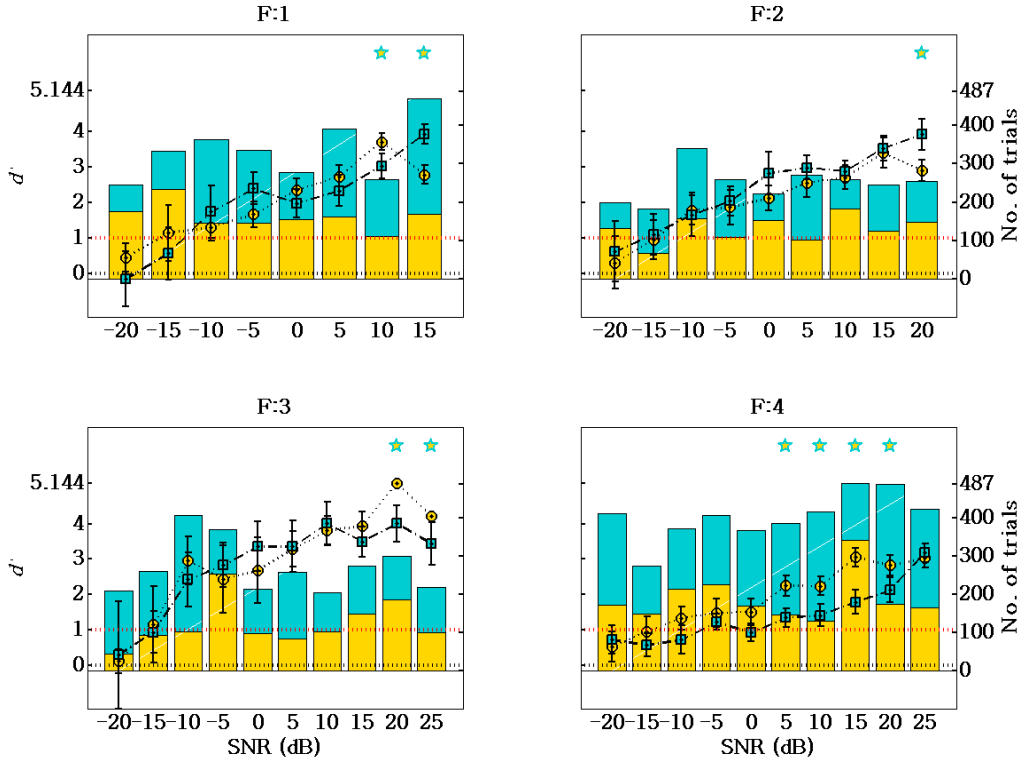

Fig 1.3.1. The effect of SNR on left/right discrimination performance (expressed in *d'*). *d'* values are displayed using markers with corresponding 95% confidence intervals. The number of trials collected for each block/SNR combination are displayed in the underlying bar graphs. Marker and bar colours indicate the behavioural block numbers 1 and 2 (yellow and green, respectively). Black and red dotted lines indicate *d'* = 0 and 1, respectively. Stars indicate a significant difference in *d'* between SNR measures across the two behavioural blocks.

Data were pooled by taking all of the trials from each behavioural block and then analysing them together. For ¾ animals (F:1, F:2 and F:3) *d'* remained high (*d'*≥3) at high SNRs (≥ 10dB SNR), as SNR was reduced below this reductions in

d' were observed. For animal F:3 *d'* values remained reasonably high (*d'*≥2) until the SNR was reduced below -10dB SNR. Animal F:4 demonstrated a gradual reduction in *d'* with SNR. In general the animals were reasonably robust to reductions in SNR when the SNR was high until a point was met where sensitivity began to decrease more rapidly.

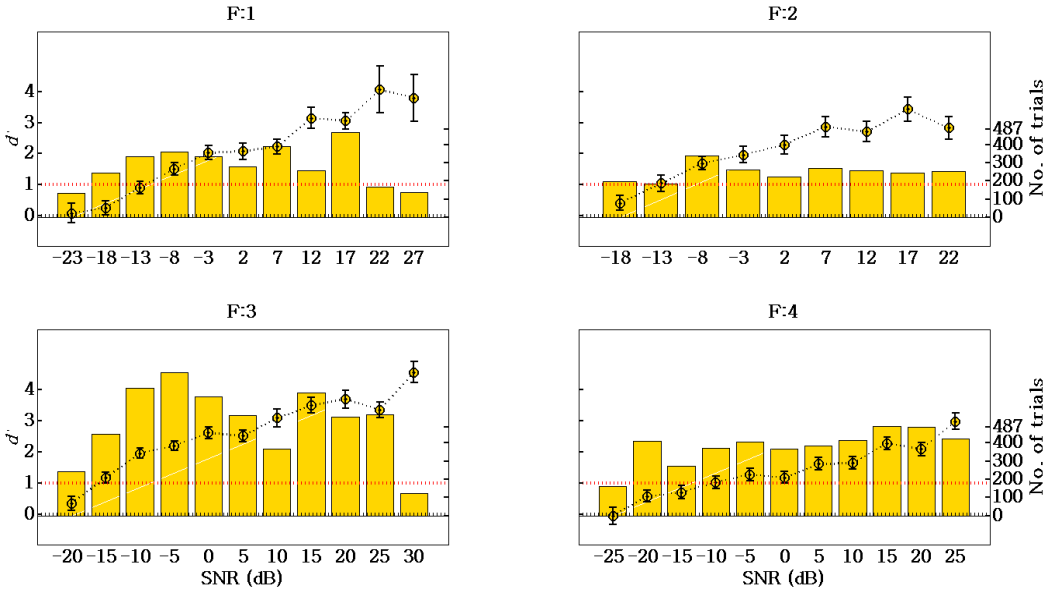

Fig 1.3.2. The effect of SNR on the collated left/right discrimination performance (expressed in *d'*). Data were collated across behavioural sessions (plotted in Fig 1.3.1). *d'* values are displayed using markers with corresponding 95% confidence intervals. The number of trials collected for each block/SNR combination are displayed in the underlying bar graphs. Black and red dotted lines indicate *d'* = 0 and 1, respectively.

Thresholds were estimated by linearly interpolating between adjacent data points of the psychometric function and finding the SNR where *d'* = 1. It is well known that *d'* is invariable to the number of the number of alternatives in forced choice procedures (Green and Birdsall, 1964; Green and Swets, 1959). This allows comparison of the 2 and 3-location discrimination thresholds. Thresholds were low when compared to the left/centre and centre/right pure tone conditions in the 3-location discrimination paradigm. Threshold values of -11.5, -12.6, -15.5 and -8.74dB SNR were obtained for animals F:1-4, respectively. For each condition, in the 3-location task, the animals ability to discriminate between two locations (left/centre and centre/right) was

measured, albeit while attending to the other speaker pair simultaneously. The improvements in sensitivity demonstrate that some element of the previous task, e.g. a smaller angle between speaker pairs or attending to three instead of two locations, increased task difficulties and elevated thresholds. Reducing the speaker locations and increasing the angle between speakers resulted in a large difference between the mean 3-location thresholds (3.75 and 8.75 dB SNR for left/centre and centre/right, respectively) and the mean left/right threshold (-12.1dB SNR).

1.3.3.2 Comparing data collection methods

The left/right discrimination paradigm produced reliable behaviour with relatively low thresholds. However, collecting data using the methods of limits is time intensive. With a minimum of two sessions measured for each SNR and with ~10 SNRs measured, this meant that the minimum amount of time taken for data collection of a single threshold was 10 days, though it often took longer. In order to reduce the time taken to collect discrimination thresholds alternative paradigms were investigated. The method of constant stimuli, as applied here, looks to measure five SNRs: two above threshold, two below threshold and one at some arbitrary SNR which evokes high performance. As the whole psychometric function is not measured, it is quicker and will measure the slope of the psychometric function around threshold but will only sample one value at higher SNRs. The staircase method adaptively reduces SNR within a single session and iterates quickly to threshold and then remains around this SNR for the remainder of the session. This will mean that SNRs far removed from threshold will be under-sampled but a very good estimate of threshold should be gained.

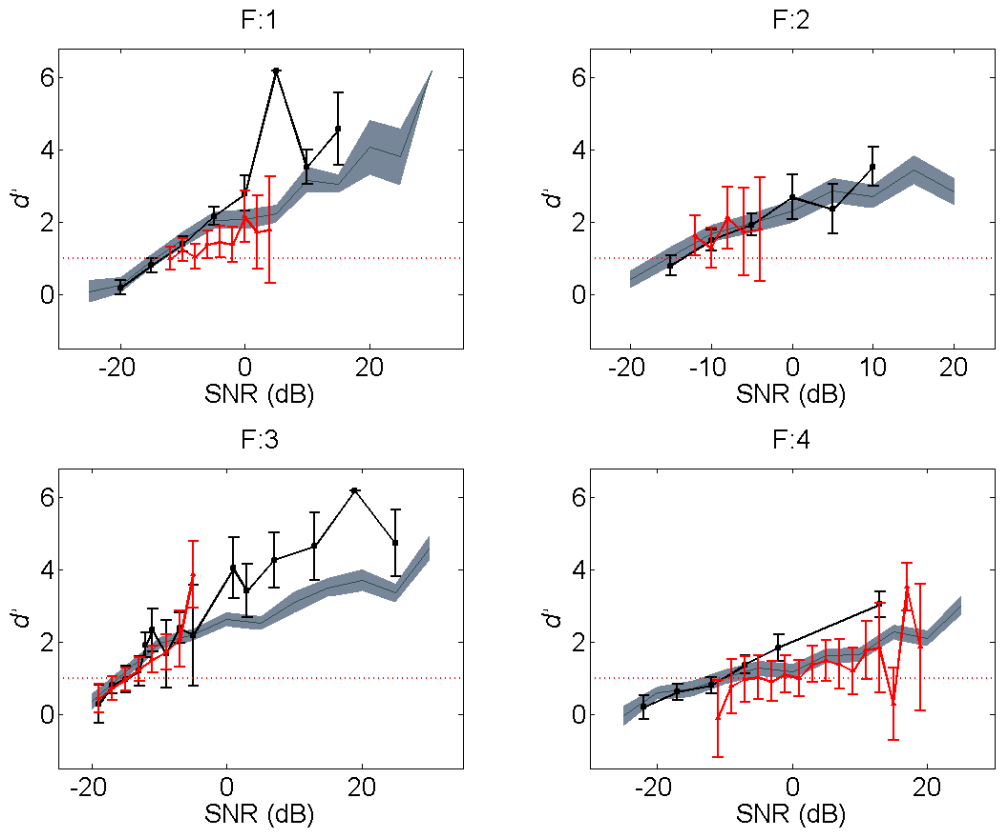

Fig 1.3.3. The effect of method on the psychometric functions. Data collected using each method were collated across behavioural block and the psychometric function plotted. Line colour indicates the data collection method: dark grey lines show the *d'* for the method of limits where the shaded grey area is the 95% confidence intervals, black lines and error bars show the method of constant stimuli *d'* and 95% confidence intervals (respectively) and red lines and error bars show the staircase method *d'* and 95% confidence intervals (respectively). Dotted grey line indicates threshold (*d'*=1)

The psychometric functions measured for each method were contrasted for each animal (Fig 1.3.3). The reader should note that the staircase method was designed to approach 70.7% correct performance and not to reach a threshold in *d'*. Also worthy of note is that for the method of constant stimuli the same five data points were not always used in both behavioural blocks, these were decided on by performance at the beginning of each behavioural block. As the method of limits had been successfully applied in previous studies this was used as the point of comparison (dark grey lines). For an alternative method to be deemed consistent with this performance it needed to yield similar *d'* values at the same SNR. When collecting data using the method of constant stimuli (black lines) the measured *d'* values were mainly consistent, though sometimes

inconsistent, with the method of limits data. At high SNRs *d'* values were frequently higher, for example for animals F:1, F:3 and F:4 a number of data points at high SNRs yielded larger *d'* values then the method of limits. However, at the four lowest SNRs (those points surrounding threshold) the method of constant stimuli provided a close fit with the method of limits data where all data points had overlapping confidence intervals. For the staircase method there seemed to be no dichotomy in the results dependent upon SNR. At most SNRs *d'* values were similar to those collected using the method of limits. For two of the animals F:1 and F:4 a number of *d'* values were lower than the method of limits at comparable SNRs.

The main aim of testing alternate methods was to determine which method could produce an accurate threshold measure in the shortest possible time. Therefore the number of trials needed to yield a threshold estimate was modelled by bootstrapping these data (for methods see section 1.3.2.4). In short trials were evenly distributed between the SNRs tested in a manner consistent with their data collection method. Threshold was then estimated via linear interpolation between the generated *d'* values (see 1.3.2.3). Fig 1.3.4 displays both the individual animal results (top and middle rows) and the mean results of the bootstrap (bottom plot) for each method.

The method of limits demonstrated a large change in threshold estimate when increasing the number of trials from 20 to 75 trials. This was evident for all animals in the individual data (top left) and hence in the mean data (bottom, square marker). The method of limits also demonstrated a majority of the change in threshold estimate had occurred when the number of trials was varied from 20 to 75 trials. This was also evident in both individual (top middle) and mean data (bottom, triangle marker). By contrast the amount of change over this range was modest for the method of constant stimuli when compared to the method of limits (mean change of 3db compared with 17.3 dB, respectively). Threshold estimates produced using the staircase method

required more trials in order to stabilise for both individual (top right) and mean data (bottom, diamond marker). Below 100 trials changing the number of trials could either increase or decrease the threshold estimate and even when a large number of trials was used the threshold estimate could vary (F:3, top right, diamond marker).

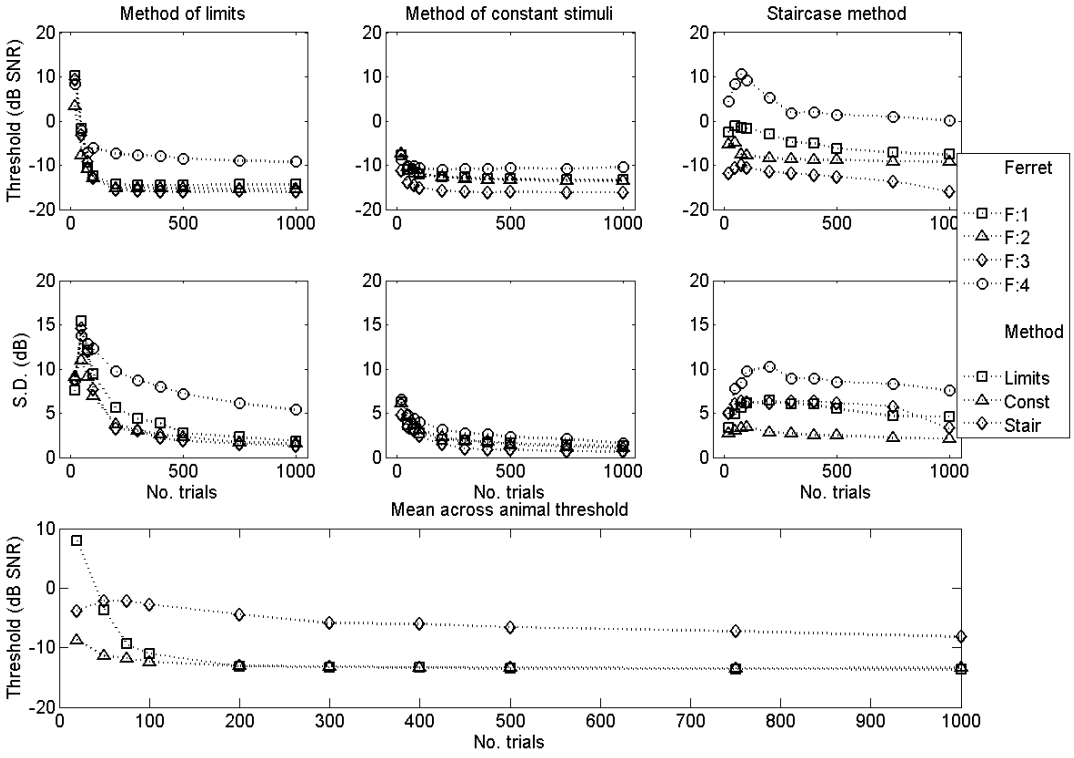

Fig 1.3.4. The effect of varying the number of trials on the bootstrapped threshold and standard deviation (S.D.) for each data collection method. Top panels – Threshold for each animal and each method. Middle panels – S.D. for each animal and each method. Bottom panel – The across animal mean threshold for each method.

Standard deviation in the thresholds estimates mirrored the changes in the threshold estimates. For both the method of limits and method of constant stimuli standard deviation reduced rapidly with increasing numbers of trials and were relatively stable when increasing the number of trials above 100 (middle left and right). For the staircase method three of the four animals (F:1, F:2 and F:3, middle right, square, triangle and circle, respectively) still demonstrated high variability in threshold estimates even when 750 trials were used.

There was a good fit between threshold estimates for the method of limits and constant stimuli. The across animal mean thresholds were almost identical when 200 or more trials were used (bottom, square v.s. triangle, respectively). By contrast the staircase method produced much higher estimates of threshold (bottom, diamond). Similarity of the across animal means does not ensure the similarity of the two methods as individual differences across methods could be cancelled out. The differences between the individual functions for each method were calculated. When more than 200 trials were used the across animal mean absolute difference between the method of limits and constant stimuli was never greater than 2dB (~1dB when 1000 trials were used). By contrast the across animal mean absolute difference between the method of limits and staircase method was always >5dB and as large as 8dB when 200 trials were used.

In general threshold estimates varied little, when more than 100 trials were used, for all three methods. The method of limits and staircase method demonstrated a smaller effect to the number of trials before this point. The staircase method, however, continued to change as more trials were added. There was good agreement between the threshold estimates for the method of constant stimuli and method of limits when 200 or more trials were used, whereas the staircase method elevated thresholds relative to the other two methods.

1.3.3.3 Assessing the ability of the task to measure detection thresholds

The main aim of this study was to test if an approach to target paradigm could be used to measure detection thresholds. The most obvious way to assess this is to compare the thresholds gained using a signal detection paradigm (Yes/No paradigm, see section 1.3.2.5) with an approach to target paradigm (the left/right discrimination paradigm applied here). Fortunately the signal detection data had already been collected “in house” using the same

behavioural arenas and equipment, though using a different population of animals. To aid comparison the data were sampled and thresholded in an identical manner across the two behavioural methods (see section 1.3.2.5).

Fig 1.3.5 displays the thresholds gained using each paradigm and for each method for two separate populations of ferrets. For all three data collection methods the Yes/No paradigm produced higher mean thresholds (-5.7, -4.9 and -0.1 dB SNR) than the Left/Right task (-10.9, -10.8 and -6.3 dB SNR, for the method of limits, constant stimuli and staircase respectively).

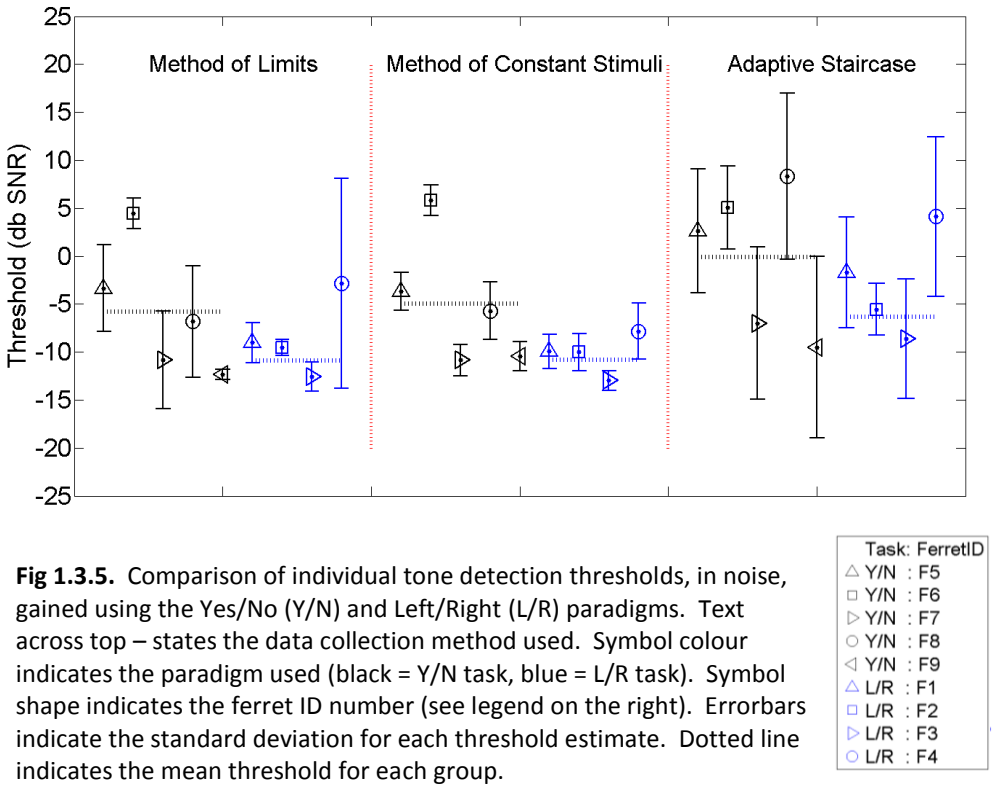

Fig 1.3.5. Comparison of individual tone detection thresholds, in noise, gained using the Yes/No (Y/N) and Left/Right (L/R) paradigms. Text across top – states the data collection method used. Symbol colour indicates the paradigm used (black = Y/N task, blue = L/R task). Symbol shape indicates the ferret ID number (see legend on the right). Errorbars indicate the standard deviation for each threshold estimate. Dotted line indicates the mean threshold for each group.

A two-way ANOVA, with paradigm and data collection method as factors, revealed a significant main effect found for both group and method ($F= 305.7$, $p< 0.05$ and $F=32.6$, $p<0.05$, respectively) and a significant interaction was also found ($F=10.5$, $p<0.05$). This demonstrated that the Left/Right paradigm produced significantly lower thresholds than the Yes/No paradigm that could not be attributed to the differences across method. For both paradigms the method of limits and constant stimuli produced similar mean thresholds, whereas the staircase method produced elevated thresholds. A Tukey post-hoc

test ($p<0.05$) revealed that no significant difference existed between the method of limits and method of constant stimuli, the staircase method was significantly higher than both within each paradigm.

1.3.4 Discussion

The Left/Right paradigm produced lower thresholds than measured previously using a 3-location paradigm. Reducing task difficulty by reducing the set of potential sources and increasing source separation improved performance. Data collected in separate behavioural blocks (with at least two weeks between the collection of the same data points) was consistent across behavioural blocks, demonstrating the reliability of the psychometric functions measured using the method of limits. Three data collection methods were compared these produced relatively similar psychometric functions (Fig 1.3.3). However, the method of constant stimuli overestimated *d'* values at high SNRs relative to the two other methods. It is possible that this difference can be attributed to the differences in the task. With the method of constant stimuli 5 SNRs were selected at the beginning of the behavioural block. SNRs were randomly selected on each trial, 4 SNRs were presented at a relatively low SNR and one at high SNR. This means from the perspective of the subject carrying out this task the majority of the signals would be difficult to perceive with the occasional signal being very easy to perceive. For the method of limits one SNR was used for each session and so the task difficulty would not vary during the session. For the staircase method SNR presentation was designed to decrease dependent on performance, therefore task difficulty would remain relatively stable across the session. It is possible that for the method of constant stimuli the majority of stimuli in a session required a large degree of focus and so when the easiest stimuli was presented (the high SNR designed to yield good performance) it was easy to detect.

The number of trials needed to collect a threshold with each method was contrasted by bootstrapping the data. When less than 100 trials were used the number of trials largely affected threshold estimates for the method of limits when compared with the method of constant stimuli. This is presumably due to the distribution of trials across SNR. The method of constant stimuli only used 5 SNRs for each threshold estimate and the number of trials was evenly distributed amongst these 5 points. Whereas the method of limits sampled the entire psychometric function (10-11 SNRs) and hence sampled a fewer number of trials at each SNR. Both the method of limits and method of constant stimuli required ~200 trials before threshold estimates began to asymptote. At this point both provided similar threshold estimates although the method of constant stimuli had more repeats at the SNRs sampled. The similarity of the thresholds across these two behavioural methods again highlights the stability of the thresholds measured using the left/right discrimination task.

Thresholds for the staircase method were estimated here using the *d'* psychometric function where it is intended to collect samples around 70.7% correct and not a threshold value in *d'*. In this implementation the difference between the thresholds using this method and the method of limits and constant stimuli was between 5-8dB when more than 200 trials were used. By contrast when comparing the measured thresholds at 70.7% (Fig 1.3.5) the difference was <4dB. This demonstrates that this is potentially an inappropriate way to compare the staircase method to the other thresholds although regardless of the measure compared, whether in *d'* or % correct, this method still produced elevated thresholds relative to the other two methods.

Finally the results from the left/right discrimination paradigm were contrasted with results from the 1-I 2-AFC detection paradigm. The left/right discrimination task produced lower thresholds than the 1-I 2-AFC detection paradigm. This suggests that this paradigm measures a better estimate of detection thresholds. Presenting stimuli from left and right would be expected

to either not change the signal level or decrease it by 6dB (Carlile, 1990; Campbell *et al.*, 2008, and raw data from Jan Schnupp). Therefore, the difference in thresholds could not be attributed to acoustical factors caused by presenting signals from 90 and 270˚. In addition to this BMLDs are very small at high frequencies (at most 3dB) and could become smaller in free-field situations (Durlach, 1978; Saberi *et al.*, 1991). Also the signal is presumably monaural for the Left/Right task but binaural for the Yes/No task, binaural listening conditions conferring approximately a 3 dB advantage. Finally localisation ability declines 20dB above detection threshold in humans, though this is for a greater number of sources than used here. Even if Left/Right discrimination is only affected within a few dB of threshold the acoustical factors should have meant that the thresholds on the Left/Right task were higher and not lower. The fact that thresholds were lower and the data less variable using the Left/Right task suggest it is an appropriate method for measuring the limits of sound detection. Add to this that it easier to train animals on and takes less time (in house observations) suggest it is a more appropriate way to collect detection thresholds than the 1-I 2-AFC detection paradigm.

1.4 Auditory filter functions

1.4.1 Introduction

The auditory system has long been conceived of as an overlapping bank of band-pass filters covering the audible range. The perceptual correlate of these filters are referred to as auditory filters (Fletcher, 1940). Fletcher measured auditory filter bandwidths via the detection of pure tone signals in the presence of band-pass noise. By gradually increasing the bandwidth of the noise signal, but holding the spectrum level, the signal detection thresholds were progressively reduced, as more noise enters the auditory filter. Once a given bandwidth, the critical bandwidth (CB), is reached the thresholds no longer increase with increasing bandwidth, as the energy is being added in frequency regions beyond that of the filter. While CB is a measure of the frequency integration of auditory filters, using a band-widening method limits the ability to glean additional information about the filter, one of the problems being that large changes in bandwidth are necessary to produce small changes in threshold. This is because it should take at least a doubling in bandwidth to produce a 3 dB change in threshold.

To counter this and other problems, the idea of using notched-noise was introduced (Patterson, 1976). This is where a notch in the spectra of the noise is introduced, centred at the frequency of the tone to be detected. The notch width is then progressively increased. This encourages the listener to attend to the auditory filter centred on the tone (where the notch is present) as opposed to switching to another filter, as this would offer the greatest SNR (Patterson, 1976). Using this technique, Patterson was able to estimate the frequency response of the auditory filter as an actual shape. This can be done by assuming that the threshold of a signal at a given frequency is determined by the amount of masker energy passing through an auditory filter (centred at that frequency), known as the “power spectrum” model. More rigorously, the power of a signal (P_s) at masked threshold is given by:

$$P_s = K \int_{-\infty}^{\infty} N(f)W(f)df$$ **eq. 3.6.1**

Where *f* is frequency, *N(f)* is the long-term power spectrum of the noise masker and W(f) is the weighting function which describes the auditory filter. *K* is a constant which represents the efficiency of the detector for a given subject at a given frequency for a given masker. The power spectrum model is valid in most listening situations although there are a number of conditions where it is violated (Moore and Glasberg, 1987). For example co-modulation masking release (Hall *et al.*, 1984), profile analysis (Green, 1988), informational masking (Watson, 1987), the overshoot effect (Zwicker, 1965) and dip listening (Kohlrausch and Sander, 1995). While this may seem like a daunting list of exclusions in reality the power spectrum model is valid in many listening situations.

From equation 3.6.1 two of the variables are known: the power of the signal at threshold can be measured and the noise spectrum used when measuring this threshold. The *K* value is a constant, and so by measuring the threshold under a number of different noise functions this can be solved. The filter values themselves are derived using a model of the auditory filter shape, a number of forms have been proposed, such as the roex function (Patterson *et al.*, 1982), the gammatone function (Patterson *et al.*, 1988) and the gammachirp function (Patterson and Irino, 1997). Using one of these functions to fit the data allows estimation of the filter shape, which in turn can be used to estimate other values such as thresholds in different listening conditions and equivalent rectangular bandwidths (ERBs).

As yet, there are no published data to describe the auditory filter shapes in the ferret. This data is only currently available for the cat, guinea pig, mouse and chinchilla (Evans, 1972; Pickles, 1975; Niemiec *et al.*, 1992; May *et al.*, 2006) and no clear consensus has been found on whether auditory filter widths are wider

in animals then in humans. This has ramifications for the study of neural frequency tuning, as the perceptually measured auditory filter tuning presumably reflects the underlying neural frequency tuning. This phenomenon provides an opportunity to further test the Left/Right method for collecting detection thresholds in more complex listening situations. It is also an opportunity to collect data that can be used in conjunction with neural data to understand frequency tuning and neural encoding within this species. Thus, the frequency resolution of the ferret auditory system was probed using the Left/Right method and notched-noise maskers.

1.4.2 Methods

1.4.2.1 Data collection

The method of constant stimuli was used to collect thresholds, with 5 signal levels presented during the course of each session. In order to select the signal levels of interest, such that one level gave >80% performance, two levels were above threshold and two below (covering a 15 dB range in 5 dB steps), it was necessary to first gain an estimate of threshold. Initially, a number of signal levels were presented, using the method of constant stimuli, within a single session where a 50 dB range would be covered in 10 dB steps. These parameters were presented for a minimum of three sessions. This yielded a sparse psychometric function, demonstrating the rough level range where performance reached threshold. The 5 levels of study were then selected from this. Signal levels used were between 0 and 94 dB SPL. Masker levels were 20±2 dB SPL for all notch-widths and frequencies. A minimum of four sessions were collected for each masking condition.

1.4.2.2 Noise maskers

Two band-pass noise stimuli were generated by creating two equal bandwidth band-pass noises with a given frequency separation. To ensure that frequencies presented were restricted to the desired range, the band-pass noises were created by summing sinusoids. So, for a given frequency range, e.g. 5 kHz to 10 kHz, a tone was generated for each frequency (at every integer Hz value), each with a random phase, between and including 5 kHz and 10 kHz and then summed. These were all scaled in exactly the same way to give constant spectral density (flat spectra). The two noise bands were centred on the frequency of interest (the centre frequency of the auditory filter of investigation) and separated by the required notch width. The bandwidth of the individual noise band was varied dependent upon the centre frequency of interest. The bandwidths of the two noise bands were selected such that, when there was no separation between the two bands, the combined bandwidth was wider than an estimated auditory filter width.

The notch-widths were selected, in pseudo-random order, from the set: 0, 0.0313, 0.0635, 0.125, 0.25, 0.33*, 0.5, 0.66*, 0.75, 1, 1.5, 2, 3, 4 octaves. Performance on previous trials dictated the selection of the notch widths. For small notch-widths there were occasions where between two notch widths no change had occurred, in which case any notch-widths in between were not collected. If, at large notch-widths, a large enough proportion of change had occurred to sufficiently describe the filter, no larger notch-widths were measured.

3.6.2.3. Roex function fitting procedure

The rounded-exponential (roex) function is a function designed to stereotype the shape of auditory filters. The roex function takes the form:

$$W(g) = (1 - r)(1 + pg)\exp^{-pg} + r \qquad \textbf{eq. 3.6.2}$$

Where *p* is a parameter determining the slope of the filter skirts, *r* is a parameter determining the cut-off of the filter (where it flattens out, effectively controlling the dynamic range) and *g* is the normalised deviation from the centre of the filter (deviation from the centre frequency divided by the centre frequency of the filter):

$$g = \frac{(f_c - f_l)}{f_c} \qquad \textbf{eq. 3.6.3}$$

Where f_c is the centre frequency and f_l, the frequency of the lower edge of the high frequency noise band (the noise band with the higher centre frequency). Using the power spectrum model (eq. 3.6.1), we can use our auditory filter (eq. 3.6.2) to estimate the masked threshold for each notch-width. Before this can be done, however, it is necessary to provide three paramaters: *K*, *r* and *p*. In order to find the optimal parameters for the roex function a least squares minimisation was performed. Unfortunately, one problem with this is that a number of local minima can exist within parameter space and when performing the minimisation it is possible for the minimisation algorithm to become "trapped" in these local minima. I could find no systematic way of estimating initial parameters within the auditory filter literature. To reduce the probability of this problem occurring, parameter space (*K*, *r* and *p*) was sampled at a number of points where $0 < r < 1$, $0 < p < 500$ and $0 < K < 200$. In total 288 points were calculated. These parameters were used to find the roex filter using eq. 3.6.2 for each parameter set. A different noise spectrum was associated with each notch-width tested. Each noise spectrum was filtered using each set of filter estimates, yielding the estimated noise energy within a given auditory filter:

$$Ns_{ij} = K_i \int_{-\infty}^{\infty} N_j(f)\, W_i(f)\, df \qquad \textbf{eq. 3.6.4}$$

For *i* = 1 to 288, each of the possible parameter sets and *j* = 1 to n, i.e. the number of notched-noise conditions. Where N_j is the actual noise spectrum used and Ns_{ij} is the noise spectrum within the estimated auditory filter, W_j is the roex function and K_j the constant for this parameter set. Parseval's theorem states the sum (or integral) of the square of the waveform is proportional to the sum (or integral) of the square of its Fourier transform. The spectrum is presently in squared pressure units therefore the summed noise spectrum (given in the equation above) can be converted into SPL using the following formula:

$$Nest_{ij} = 10log10\frac{Ns_{ij}}{(20 \times 10^{-6})^2} \qquad \textbf{eq. 3.6.5}$$

Where $Nest_{ij}$ is the SPL of the noise within the estimated auditory filter. The denominator refers to the squared reference pressure of 20µPa. At this point, we have an estimate for the noise level within an estimated auditory filter, but we do not know the necessary signal to masker ratio within a filter at threshold. It is common to assume that the signal level necessary to perceive the sound will be equal to the energy within the auditory filter, and this was the assumption used here, thus the noise level within the estimated auditory filter serves as an estimate of the signal level at threshold. Using this threshold estimate, we can then compare the estimated threshold using these fitted parameters to the empirically obtained threshold using an error function:

$$err_i = \sqrt{\left(\sum_j Nest_{ij} - \theta_j\right)^2} \qquad \textbf{eq. 3.6.6}$$

In the 3-dimensional parameter space there is now an error value associated with each parameter set tested. This can be plotted as error surfaces in parameter space (Fig 1.4.1).

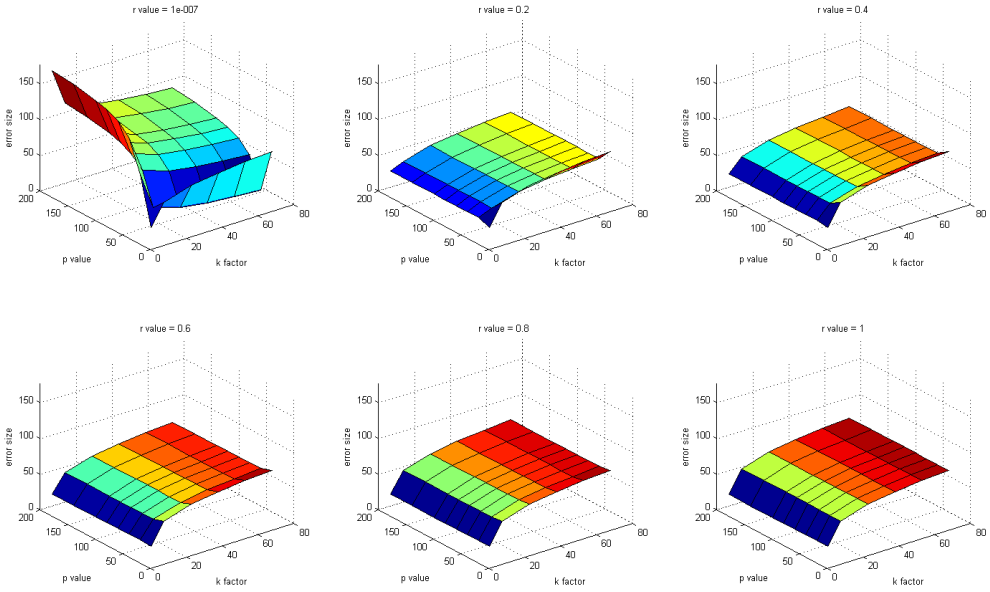

Fig 1.4.1. An example "snap shot" of error size in parameter space. Before parameter sets were selected for the minimisation algorithm a sparse sampling of parameter space was taken to find the approximate locations of local minima (see text).

The first measurement of the errors (the 288 points derived using eq. 3.6.6 and the error surfaces in Fig 1.4.1) acted as a broad snap shot of errors in parameter space. This highlighted regions where fits between the modelled data and the measured data were similar, indicating an appropriate estimate of the auditory filter. Parameter space was split into 10 cubes (2x5) and, within each cube of parameter space, the local error minima taken. The effect of running 10 minimisation searches, spread over parameter space, was to limit the influence of the starting parameters. It was hoped that any sub-optimal fits, found by iterating into a local minima, could be ignored subsequently and the best fits used.

Under certain listening conditions it might be beneficial to listen to a filter which was not centred on the target frequency if the signal to noise ratio was better than that for a filter centred on the tone frequency. The centre frequency of each filter estimate was allowed to vary within ± 10% of the centre frequency. This led to a number of estimates of the function, representing thresholds at different notch widths. If the total error (eq. 3.6.6) was lower for a different

centre frequency, then this frequency was used as the centre frequency of the filter. Normally, the best filter shape, for a filter at centre frequency, would be found using the least squares regression. Using the amendment mentioned allowed a different shape to be fitted at a different centre frequency if it produced a lower error in the fitting (eq. 3.6.6).

1.4.3 Results

1.4.3.1 Measured thresholds

Thresholds with different noise notch widths were tested at 4 frequencies (Fig 1.4.2). Overall it is clear that there were substantial effects of notch-width for all target frequencies, thresholds dropped by as much as 35dB with increasing notch-width. This clearly demonstrates that the animals are processing sound in band-pass filters. It also demonstrates that the Left/Right method itself can be used to measure these changes.

For those animals tested at 1 kHz (F:3 and F:4) the difference between threshold when the notch width was 0 Hz and 4 octaves was 25.2 dB and 30.8 dB respectively. A majority of the decrease in threshold occurred between 1 and 2 octaves, whereas, after 2 octaves the decrease in threshold became shallower on the octave scale. The threshold did not reduce immediately for animal F:3, with very little change in threshold until the notch width was >1 octave, followed by a steep decline in threshold.

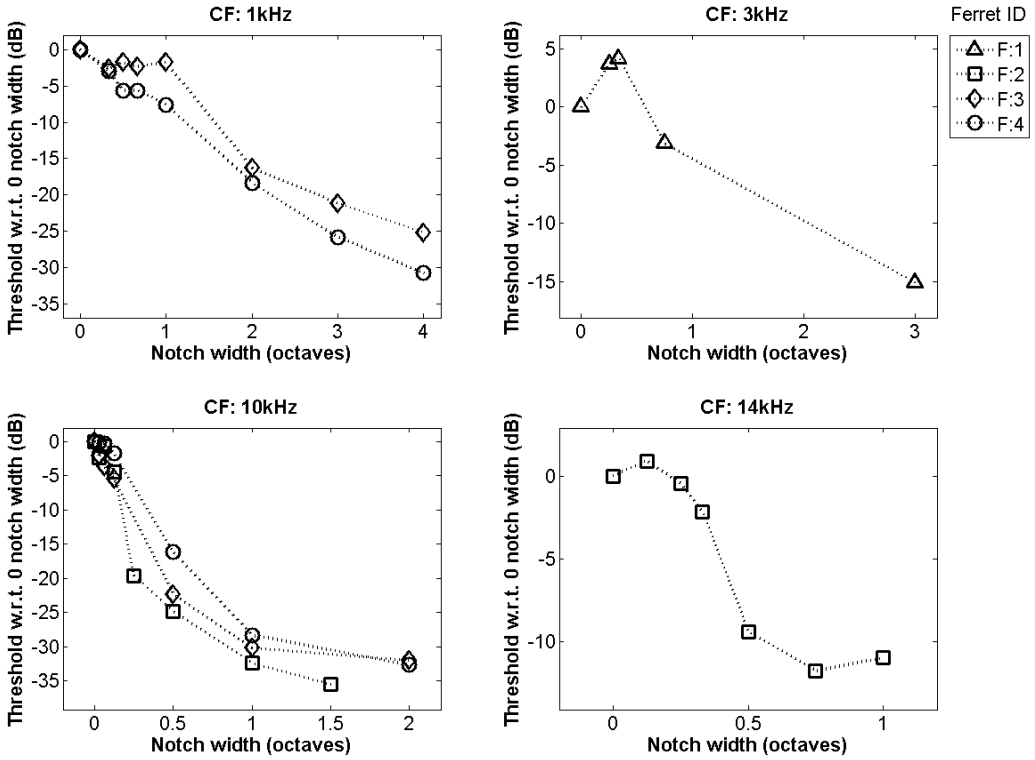

Fig 1.4.2 Detection thresholds with increasing notch-width. Thresholds are plotted with respect to the threshold when there was no notch. Centre frequency (tone frequency) is displayed as the title. Symbols represent the animal tested.

Only one animal, F:1, was tested at 3 kHz. As the notch width was increased from zero, the threshold actually increased by 5dB and then began to drop. At the largest notch width tested, threshold only dropped by 15 dB. Three animals: F:2, F:3 and F:4 were tested at 10 kHz. Thresholds decreased sharply as the notch width increased, with large reductions, > 15dB, having occurred by 0.5 octaves. Thresholds continued to decrease sharply beyond this and then began to reduce less. Animal F:2's thresholds began to reach the performance floor before 1 octave, while the other two animals began to reach performance floor at 1 octave notch width. One animal was tested at 14 kHz (F:2). At the largest notch width tested the threshold only dropped by 12dB. Thresholds began to reduce at 0.33 octaves, though after 0.5 octaves the majority of the change had occurred.

1.4.3.2 Fitting roex functions

In order to gain the weighting of the auditory filters, roex functions were fitted to these data. Fig 1.4.3 shows an example of the fitting procedure. The first example is that of animal F:3 measuring the auditory filter at 1 kHz. A number of minima were found in parameter space when fitting functions (Fig 1.4.3, top left). The *r* values (see methods) for these fits were all restricted to $r<0.001$. A cluster of fits were localised within the parameter range $0<p<8$, $0<r<0.0002$ and $0<K<100$ (fits 1:7). These produced poor estimates of thresholds and hence poor estimates of the psychometric function (Fig 1.4.3, top left) this in turn meant the error between the actual and modelled data were large (Fig 1.4.3 top right). The fits within this cluster appear to overestimate the width of the auditory filter (Fig 1.4.3 bottom left) leading to large ERB values (Fig 1.4.3, top right).

Fit 10 produced a better fit to the measured data (Fig. 1.4.3, top left) and hence yielded a smaller error and also lower ERB values (Fig. 1.4.3, top right). The parameters here were markedly different from those fits described previously, with relatively large *r*, similar *p* but large *k* (>100). Fits 8 and 9 iterated to the same minima, this gave the best fit of the psychometric function of all (Fig. 1.4.3, top left). The error was half that found for fit 10 and the ERB 200Hz smaller. Overall, the majority of the minima found resulted in filters that were too broad to result in fits which accurately matched the slope of the threshold function. This demonstrates the need to sample a number of starting parameter values before performing the minimisation search for the most appropriate fit.

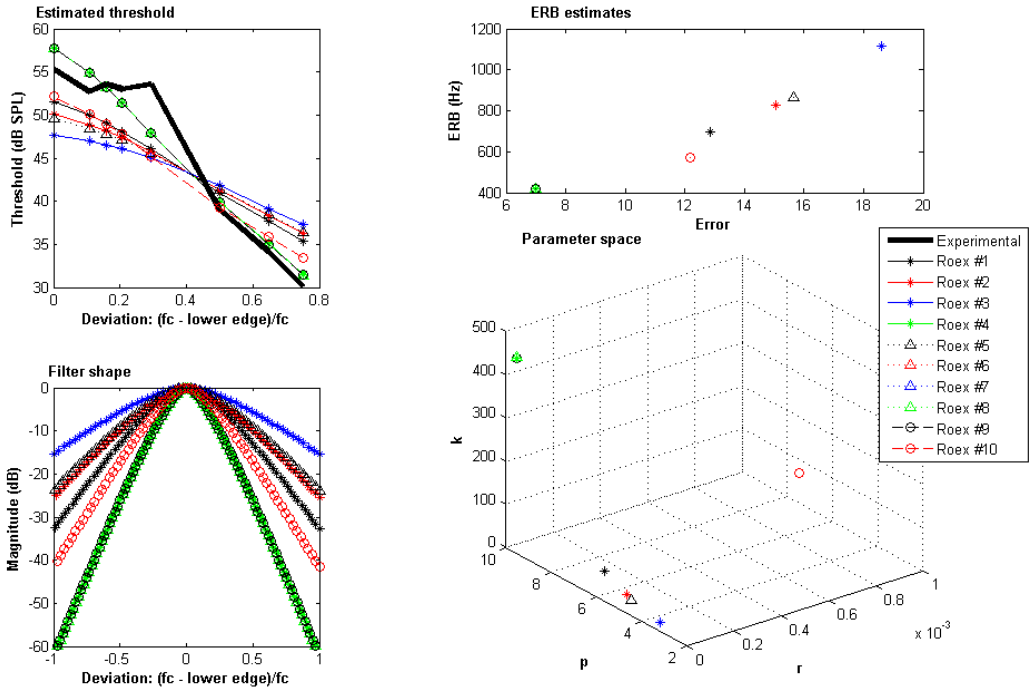

Fig 1.4.3 Roex fitting procedure for F:3 at 1kHz. Top left – estimated and observed thresholds. Bottom left -auditory filter shapes corresponding to the estimated thresholds. Top right – the error size (total dB difference in function from top left) versus the estimated ERB estimate. Bottom right – The parameters of the local minima found for each regression undertaken, fits could take very different parameter values and so a number of fits were necessary to avoid selecting inappropriate fits (see text).

The second example I have included is for the fitting procedure for F:4 at 10 kHz. Five local minima were found in parameter space (Fig 1.4.4, bottom right). Fits 4 to 7 iterated toward minima with low *p* values, resulting in a shallow slope gradient in the filter shape (Fig 1.4.4, bottom right). The *r* values were also relatively low, resulting in a narrow cut-off bandwidth. The *K* values for this cluster were small (<100). The error in this cluster of fits was large, due to underestimating the slope (*r* value) of the function, which resulted in broad filter widths near the peak of the filter (Fig 1.4.4, top right and bottom left). Fits 1, 2 and 3 all iterated to the same point with a large *r* value, moderate *p* value and moderate *K* value (Fig 1.4.4, bottom right). Due to the large *r* value, the cut-off bandwidth was relatively narrow and hence the skirts of the filter become visible at ~-50 dB magnitude on the filter shape plot (Fig 1.4.4, bottom

left). The ERB width was similar to that obtained for fit 7, as was the error (Fig 1.4.4, top right). The best fits were those obtained for fits 8, 9 and 10 which also iterated onto the same minima, with a very large *K* value, moderate p values and very small r values. The small *r* values produced steep slopes, resulting in a narrower ERB (Fig 1.4.4, bottom left and top right). Again the majority of the minima found resulted in filters that were too broad to give accurate fits which matched the slope of the threshold function. Overall this step of the fitting process demonstrated the care that needs to be taken when fitting roex functions and validated the use of an additional step in the fitting process.

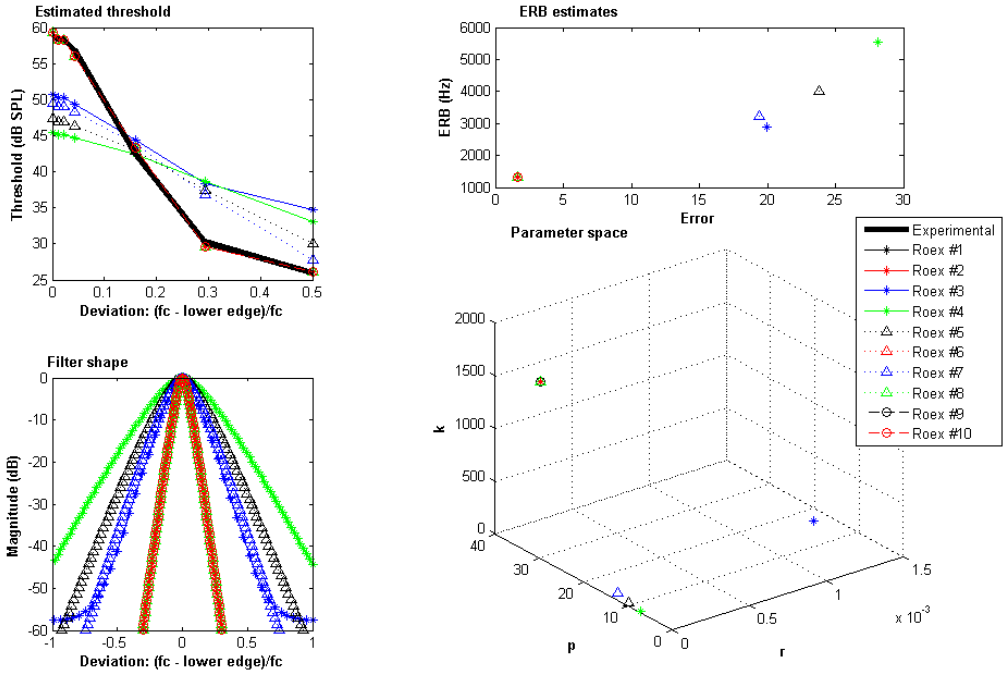

Fig 1.4.4. Roex fitting procedure for F:4 at 10kHz. Top left – estimated and observed thresholds. Bottom left -auditory filter shapes corresponding to the estimated thresholds. Top right – the error size (total dB difference in function from top left) versus the estimated ERB estimate. Bottom right – The parameters of the local minima found for each regression undertaken.

1.4.3.3 Fitted auditory filter shapes

The best fits and actual data for the 1 kHz data are displayed in Fig 1.4.5. For F:3 the measured thresholds at narrow notch-widths changed relatively little at smaller deviations. This resulted in the difference in the slope of the roex estimated thresholds at small deviations when compared with the measured thresholds. The mean of the error for each modelled data point was taken and was still relatively small at 0.87 dB. For F:4 the mean error was 0.5 dB.

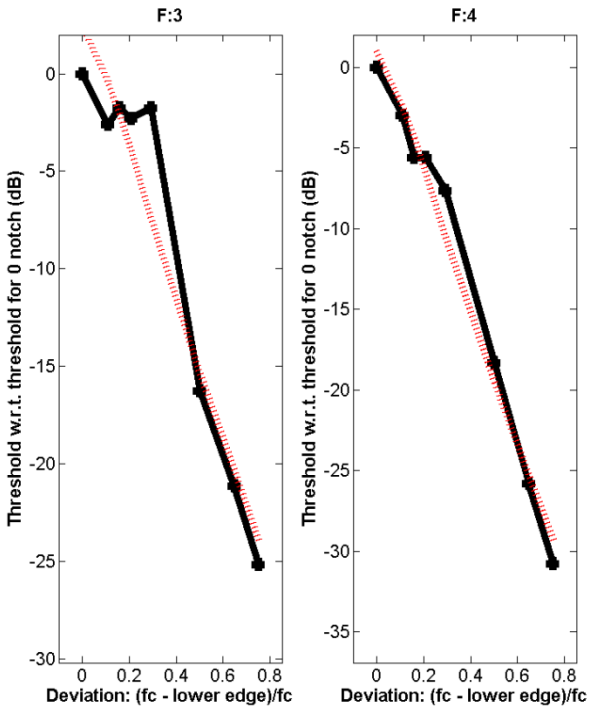

Fig 1.4.5
Measured (solid black line) and modelled (dotted red line) thresholds for increasing noise notch-width at 1kHz.

At 3 kHz one animal was tested. For F:1 the measured thresholds increased slightly and then began to decrease, with increasing notch width (Fig 1.4.6). The resulting fit still reasonably represented the data, with a mean error of 0.9 dB.

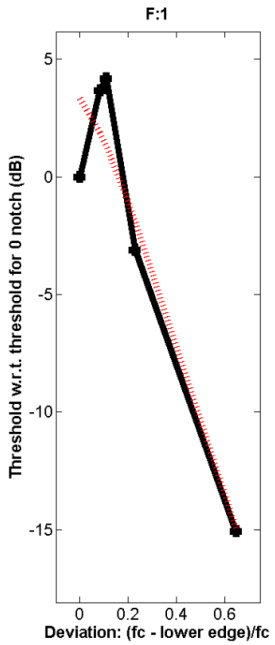

Fig 1.4.6. Measured (solid black line) and modelled (dotted red line) thresholds for increasing noise notch-width at 3kHz.

Three animals were tested at 10 kHz. All three demonstrated reductions in threshold at relatively narrow notch widths (Fig 1.4.7). The best fit for F:2 fits the data well at low deviations with larger error at higher deviations, and the mean error was 0.8 dB. Relatively good fits were found for F: 366 and F:4: the modelled filter fit the data well at all deviations measured, resulting in a low mean error of 0.27 and 0.23 dB for each animal respectively.

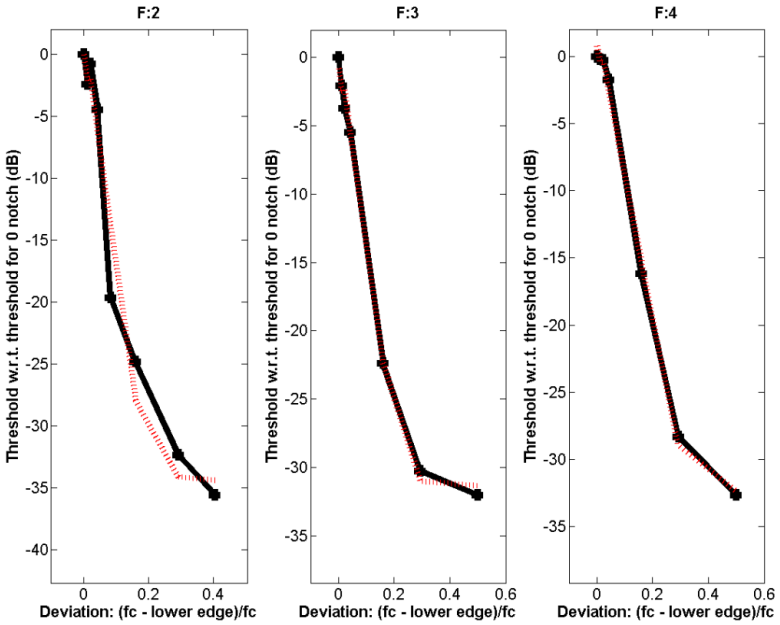

Fig 1.4.7. Measured (solid black line) and modelled (dotted red line) thresholds for increasing noise notch-width at 10kHz.

One animal was tested at 14 kHz, animal F:2. At small deviations, little change was seen in the thresholds (Fig 1.4.8). The modelled thresholds at larger and smaller notch widths deviated from that of the actual data but a reasonable fit was still gained, with a mean error of 0.48 dB.

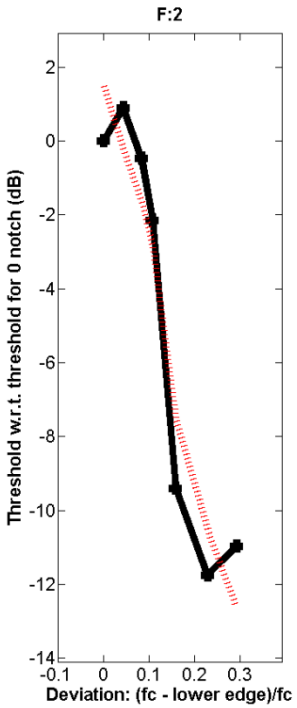

Fig 1.4.8. Measured (solid black line) and modelled (dotted red line) thresholds for increasing noise notch-width at 14 kHz.

Fig 1.4.9 displays the modelled auditory filters and the associated ERB estimates for each filter. The filters themselves are plotted in terms of magnitude and deviation, i.e. the centre frequency less the notch-width between the two band pass noises divided by the centre frequency. As notch width was scaled by centre frequency, some comparison of width across frequency could be made. At 1000 Hz the filter widths were comparatively wide relative to those measured at higher frequencies. Q10s have been recorded at the level of the auditory nerve in the cat, guinea pig and, more recently, the ferret (Sumner, 2010). Cat and guinea pig Q10s at 1 kHz are approximately between 2 and 4, with ferret Q10s being marginally smaller at ~2. The two animals tested here at 1 kHz (F:3 and F:4) had Q10s of 2.32 and 1.2 respectively,

demonstrating that the behaviourally measured Q10s were a little lower than those recorded at the auditory nerve for the same species.

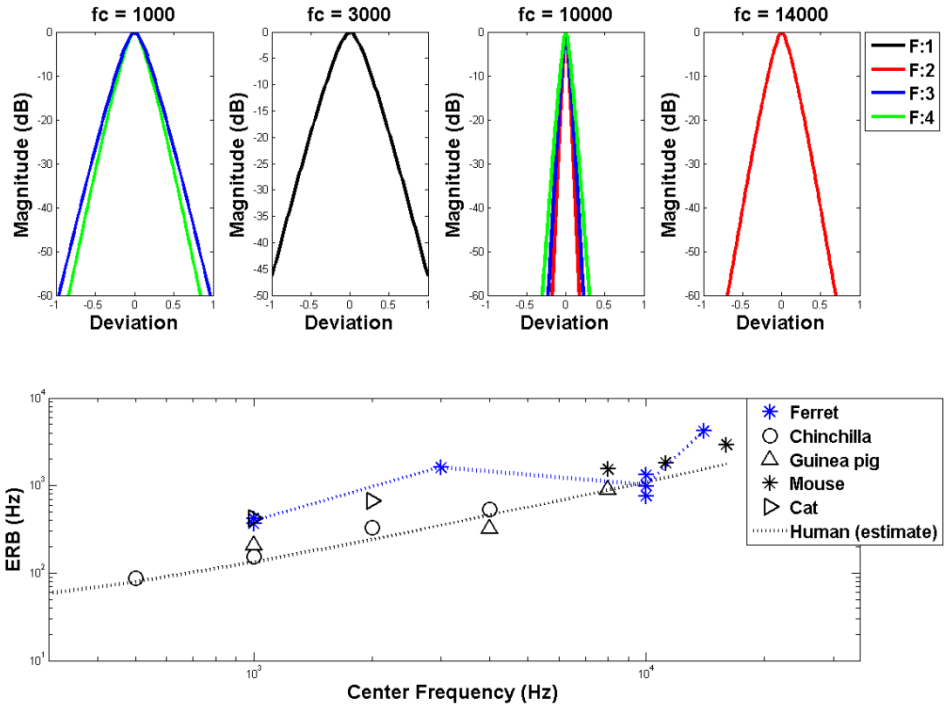

Fig 1.4.9. Modelled auditory filter shapes and the associated ERBs. Top row – Estimated filter shapes for each frequency (separate panels, the title is the centre frequency). Colours indicate the ferret being tested. Bottom row – ERBs for a number of species including the chinchilla (Niemiec *et al.*, 1992), guinea pig (Evans *et al.*, 1992), mouse (May *et al.*, 2006), cat (Pickles, 1975) and estimated data for humans (Glasberg and Moore, 1990). Blue stars reflect the ERB estimates based on the filter fits displayed in the top panels.

At 3 kHz, the measured auditory filter was a similar width, in deviations, to those measured at 1 kHz. Ferret Q10s of auditory nerve fibers are approximately between 2 and 4. The behaviourally measured Q10 here was 1.6, again slightly lower than the physiological data. In the cat and guinea pig the Q10s are between 3 and 9 and 3 and 8 for each species respectively (Pickles, 1975; Evans 1992). A number of filters were modelled at 10 kHz. These were all narrower, in terms of deviation, than those measured at lower frequencies. Q10s of auditory nerve fibers in the ferret were between 3 and 8. Those measured behaviourally here were 11, 8.5 and 6.4 with a mean of 8.67. These values were at the upper limits of the Q10s found in auditory nerve fibers,

whereas at the lower frequencies discussed, behavioural Q10s have been found to be lower than those found physiologically. The filter measured at 14 kHz was also narrower than those measured at 1 and 3 kHz (Fig 1.4.9). Auditory nerve fiber Q10s, with CF at this frequency, range between 3 and 8 in the ferret. The filter Q10 behaviourally here was 2.8 close to the lower bound of those measured physiologically. The modelled filters generally yielded Q10s close to or below the lower bounds of those measured for auditory nerve fibers with similar CFs.

Another measure of filter widths is the ERB. This data has been gathered in a number of species including the guinea pig, mouse, chinchilla, cat and human (Pickles, 1975; Glasberg and Moore, 1990; Evans *et al.*, 1992; Niemiec *et al.*, 1992; May *et al.*, 2006). At most frequencies ERB measurements reflected the trends seen for Q10s. At 1 kHz, the mean ERB was similar to that of the cat at 530Hz, although this was markedly higher than that of the chinchilla, guinea pig and human. There are no data at 3 kHz for other animal species, however, performance can be compared to that of humans using ERBs. At 3 kHz the human ERB is 348.5 Hz the ERB measured here for the ferret is considerably larger (1606Hz).

At 10 kHz the ERB was relatively narrow, with a mean of 1025 Hz, similar to that of humans (1104 Hz). The mouse data which exist for frequencies above and below this, suggest that ferret filters are narrower than the mouse over this frequency range. At 14 kHz the ERB was again higher than those measured in humans, with a value of 4208 Hz, compared with 1536 Hz for humans. At most frequencies measured, the ERB widths in the ferret were relatively broad when compared with other species, with the exception of at 10 kHz, where filter widths were comparable to that of humans.

1.4.4 Discussion

The Left/Right method was successfully applied to measure auditory filter widths. As would be expected, increases in notch width surrounding a given frequency led to improvements in the detection of signals at those frequencies. One potential issue for the Left/Right method is its application at low frequencies. This is because at low frequencies ILDs become smaller, increasing the difficulty of the task in the spatial dimension and potentially reducing thresholds. One point of note from the across species comparison, of ERBs at different centre frequencies (Fig 1.4.9), is that for each animal species the relationship between ERB and centre frequency tended to broadly reflect that measured in humans. A similar degree of increase in ERB with centre frequency was witnessed in these animal data. With the exception of the data at 10kHz (discussed later) this can also be deemed the case for the ferret data presented here. This suggests that even at low frequencies, where ILDs are small, the paradigm appears appropriate for collection of auditory filter width data.

Consistent with findings in other species, the relationship of thresholds with noise notch-width was consistent with the power spectrum model. For most of the frequencies measured (excluding 10 kHz) and on a linear scale, the trend was such that the higher the frequency, the wider the filter. This is demonstrated by the increasing ERB size with frequency (Fig 1.4.9), as is observed in other species. The ERBs measured were 53, 54, 10 and 30% of the centre frequency for 1, 3, 10 and 14 kHz respectively. For the majority of frequencies, again with the exception of 10 kHz, this is larger than those measured in other species. For instance, in humans, ERBs are typically between 11 and 17% (Moore, 2003). ERBs measured in the cat, a taxonomically similar species, are 42 and 33% of the centre frequency at 1 and 2 kHz, when measured using a conditioned avoidance task. Measurement in other species, such as the guinea pig, mouse and chinchilla, are lower than those observed here (Evans *et al.*, 1992; Niemiec *et al.*, 1992; May *et al.*, 2006). These behaviourally measured

ERBs are also slightly higher than physiological measures such as Q10s of auditory nerve fibers in the ferret. For some species, for instance the guinea pig, ERB widths measured in auditory nerve fibers are similar to those measured behaviourally (Evans *et al.*, 1992). In other species, such as the cat, they have been found to be higher (Pickles, 1979).

Performance at 10 kHz seemed anomalous, when compared with performance at other frequencies tested. The filter widths were considerably narrower than those observed at other frequencies resulting in ERBs comparable to those of humans. Furthermore, the behavioural data yielded Q10s at the upper limit of those measured in auditory nerve data. All three animals tested yielded similarly low ERBs at this frequency, suggesting that the measurement itself was relatively stable. The implication is that the ERBs measured at this frequency were smaller than should be expected in the ferret and the detection process was optimal, given the similarity to the upper bounds of the auditory nerve Q10s. The frequency range around 10 kHz is exceptional in some respects, as this corresponds to the most sensitive part of the ferret audiogram (Kelly *et al.*, 1986). Though there is no evidence to suggest that level sensitivity in certain spectral regions relates to sharpness of frequency tuning. A more likely candidate to explain this exceptional performance is training. Before measuring auditory filter widths at a number of frequencies, animals had previously only been exposed to 10 kHz stimuli. All of the work that precedes this study is carried out at 10kHz, and hence the animals have had much exposure in testing conditions with this stimulus. For three of the animals, comparable masking conditions were measured with an interval of over a year. The thresholds measured in these conditions were all within 2dB of one another, therefore no detectable improvement in performance could be measured over this time course. Presumably, if frequency selectivity had improved during this course of time, thresholds should reduce, as auditory filters would be narrower and hence more of the noise should be filtered out. This suggests that the additional

training given at 10kHz was not responsible for the narrow auditory filters measured at this frequency.

An additional step was added to the fitting process of auditory filters in this study, where parameter space was roughly sampled and the fitting process applied to a number of starting parameters. This demonstrated the starting parameters largely influenced the eventual fits and hence highlights the fact that a number of local minima (in terms of error) existed within parameter space. The method of sparsely sampling parameter space before fitting aided the process of finding the most optimal fit.

2 Electrophysiology

The second project in this thesis aimed to further our understanding of frequency integration in the auditory system. Frequency is the first auditory feature the auditory system extracts. Its importance to the auditory system is presumably reflected by the inherited tonotopic arrangement of many subcortical structures and primary auditory cortex (Pickles, 1982). The connectivity from cochlea to cortex is complicated, involving branching and convergence of inputs from multiple areas at multiple levels (Fig 2.1.1). This is reflected by differing sensitivity to sound duration, amplitude, direction and pitch being found at different subcortical levels (Palmer, 2007 and discussed in this literature review). By contrast, sensitivity to visual information such as line orientation and binocular disparity only arise in the primary visual cortex and as a product of the connectivity from LGN, as opposed to pre-processing by subcortical structures (Hirsch and Martinez, 2006). This means that auditory cortex could be considered more akin to a higher order cortical area (King and Nelken, 2009). This would suggest that we might expect to observe complicated interactions between different auditory features.

The aim of this study was to probe the relationship between frequency and space in the hope of learning about the importance of frequency integration on the coding of spatial information (in the auditory cortex). In section 2.2 an initial attempt is made to gain a broad picture of how these two features might interact. This serves to inform whether an interaction does exist and how best to investigate it further. In section 2.3 a simplified approach is applied which allows a more quantitative investigation of the interaction to be applied. The following section serves as a basic literature review of the ascending auditory system and how it processes information before and at the level of the auditory cortex.

2.1 Introduction to the processing of sound by the ascending auditory system

What follows is a description of the ascending auditory system, tailored to focus on the nuclei of the ascending auditory system that are important for the encoding and transmission of ILD information. The review begins at the auditory nerve and moves up through the auditory system stopping at the auditory cortex, as this is the area under investigation. The aim is to understand how ILD and frequency information are processed ultimately leading to a discussion of the implications of frequency integration on the encoding of ILD information in the auditory cortex. Fig 2.1.1 is a schematic drawing of the ascending auditory system leading up to the auditory cortex.

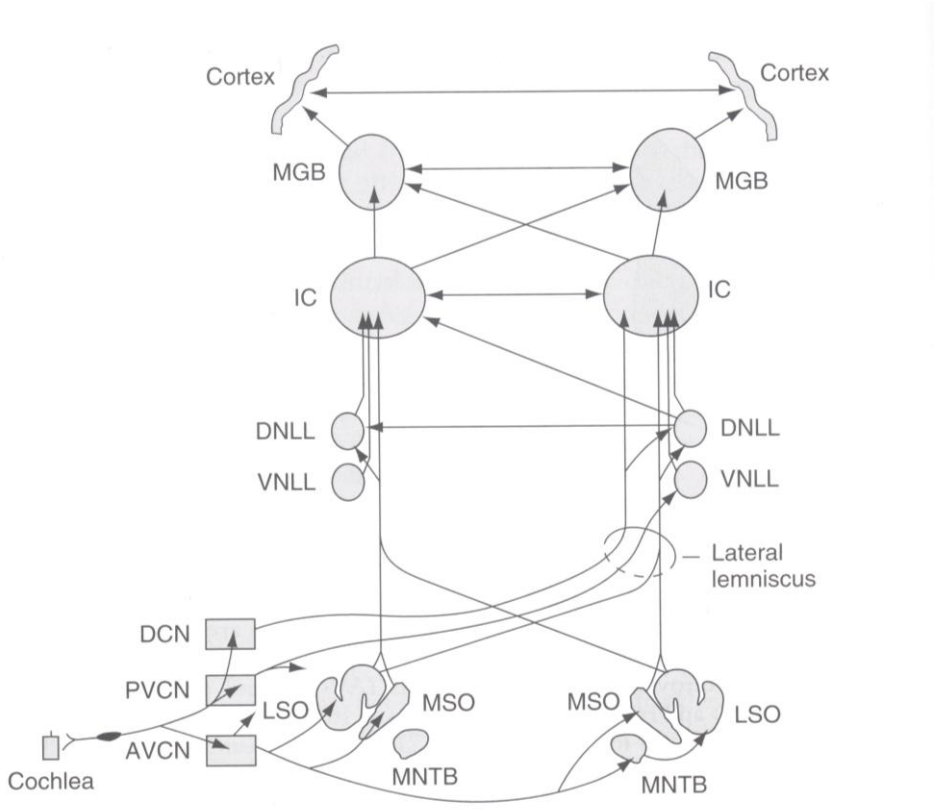

Fig 2.1.1. The main ascending auditory pathways. AVCN, Anterior ventral cochlear nucleus, PVCN, Posteroventral cochlear nucleus, DCN, Dorsal cochlear nucleus, MNTB, Medial nuclei of the trapezoid body, LSO, Lateral superior olive, MSO, Medial superior olive, VNLL, Ventral nucleus of the lateral lemniscus, DNLL, Dorsal nucleus of the lateral lemniscus, IC, inferior colliculus, MGB, Medial geniculate body. Taken from Pickles, 2008.

Before a sound is encoded in the brainstem it needs to be transduced from sound into a useable signal for the brain. Sound is a pressure wave that results from vibrations of particles, of a given medium, through which the sound wave passes. This wave enters and travels through the external auditory meatus, or ear canal, until it reaches and vibrates the tympanic membrane, or ear drum. The tympanic membrane is connected to the auditory ossicles consisting of the malleus (hammer), incus (anvil) and stapes (stirrup). The osiccles transmit the vibrations of the tympanic membrane to the oval window. Here the vibrations become motion of the basilar membrane of the cochlea. The cochlea is the organ involved in transducing sound into the electrical impulses used by the brain. The inner hair cells of the cochlea respond to the motion of the basilar membrane converting this motion into neural spikes which innervate the auditory nerve.

2.1.1 The auditory nerve

The auditory nerve transmits information from the cochlea to the cochlea nucleus. Auditory nerve (AN) fibres can be defined anatomically and physiologically. Anatomically they can be divided into two categories: those that connect with inner hair cells (IHCs) and those that connect with the outer hair cells (OHCs, (Pickles, 1982). In the cat 95% of AN fibers connect with IHCs these have bipolar cell bodies, myelinated cell bodies and axons (Spoendlin, 1978) and are referred to as Type I. The remaining 5% are those that connect with OHCs, these are not myelinated, are monopolar and are referred to as Type II (Pickles, 1982). Retrograde labelling has shown that both fibre types send axon exclusively to the cochlea nucleus (Ruggero *et al.*, 1982).

Physiological responses in type I fibres to sound are always excitatory (Pickles, 1982). Fibres can be classified into one of two categories: high and low spontaneous rate. The low discharge rate group typically fire at less than 20 spikes per second (sp/s) with the majority firing at ~0.5 sp/s, whereas the high

discharge group have a mean discharge rate of 60-80 sp/s with a maximum of 120sp/s (Evans, 1972; Liberman and Kiang, 1978). These also differ in their response to sound. Both types generally vary sigmoidally in firing rate with sound level, but low spontaneous rate fibres have higher thresholds and their spike rate varies over a larger dynamic range than high spontaneous rate fibres which saturate rapidly with increasing sound level (Sachs *et al.*, 1974). Temporal responses to tones, demonstrated by the post-stimulus time histogram (PSTH), show a sharp onset response which drops rapidly within 10-20ms, and then gradually declines after this (Kiang, 1965). The frequency tuning of AN fibers acts like a band-pass filter. The frequency yielding the lowest threshold marks the frequency known as the characteristic frequency (CF), an example of AN tuning curves can be seen in Fig 2.1.2.

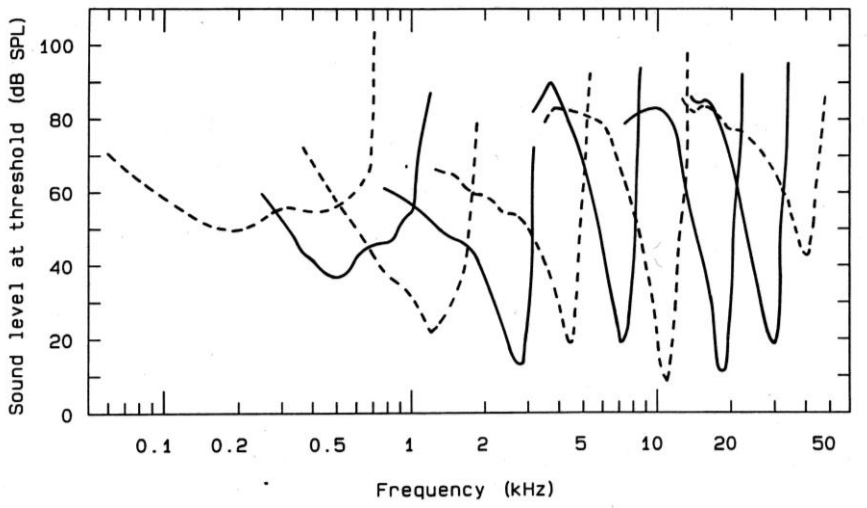

Fig 2.1.2. Tuning curves of single auditory nerve fibers in the guinea pig. Taken from Moore, 2003 (redrawn from Palmer, 1987).

When considered on a logarithmic frequency scale AN fibers with low CFs, 1kHz and below, have relatively symmetrical tuning curves and as CF increases they become increasingly asymmetrical (Palmer, 1987). Frequency tuning can be measured in a number of ways, one such measure is the quality factor or Q factor. The Q factor is a ratio of the resonant frequency, in this case the characteristic frequency, divided by the bandwidth of the unit. The bandwidths

of AN fibers vary in a monotonic way with level, becoming narrower with decreasing level. Usually the Q10 is measured, this refers the bandwidth at 10dB above threshold. An example of AN Q10 values measured in the guinea pig can be seen in Fig 2.1.3.

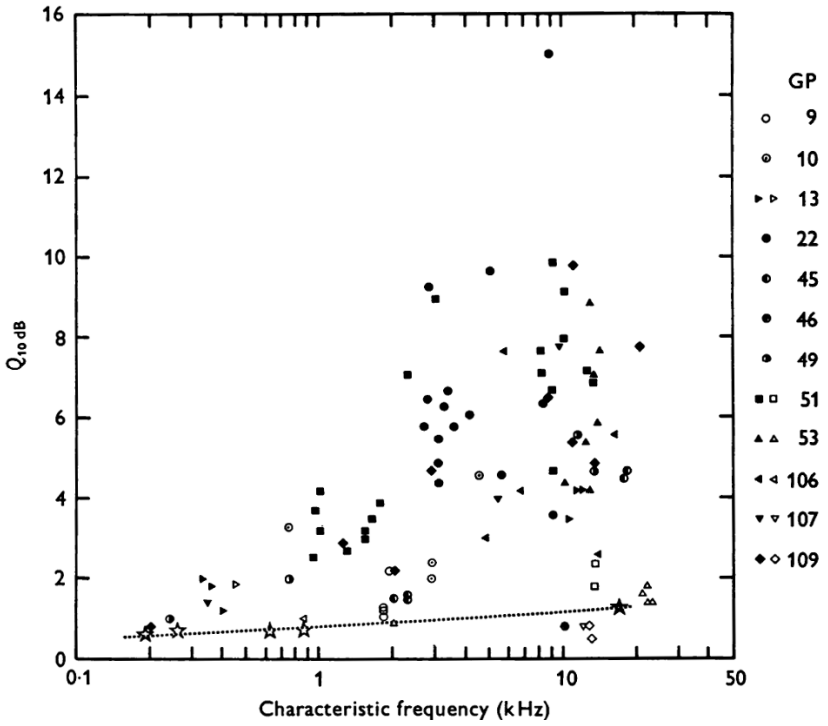

Fig 2.1.3. Q10 values for auditory nerve fibers in the guinea pig. Legend refers to the animal number. Taken from Evans, 1972.

Auditory nerve fibers with low characteristic frequencies (100-300 Hz) showed little variability in the Q10 values yielded. Fibers with higher characteristic frequencies (~10kHz) demonstrated more variability in Q10 values, ranging between 0.5 and 15. Generally Q10 values increased with increasing characteristic frequency.

Frequency is represented at different points along the cochlea with high frequencies nearer to the oval window, i.e. the point where the signal first reaches the basilar membrane, and lower frequencies represented at increasing distances from the oval window (Von Békésy, 1960). The time at which the signals reaches each point along the basilar membrane is affected by this

distance so that the lower the frequency the longer the latency between stimulus presentation and excitation on the basilar membrane.

An estimate of the travel delay along the BM can be seen in the firing of AN fibers to their characteristic frequency in the guinea pig (Palmer and Shackleton, 2009, Fig 2.1.4).

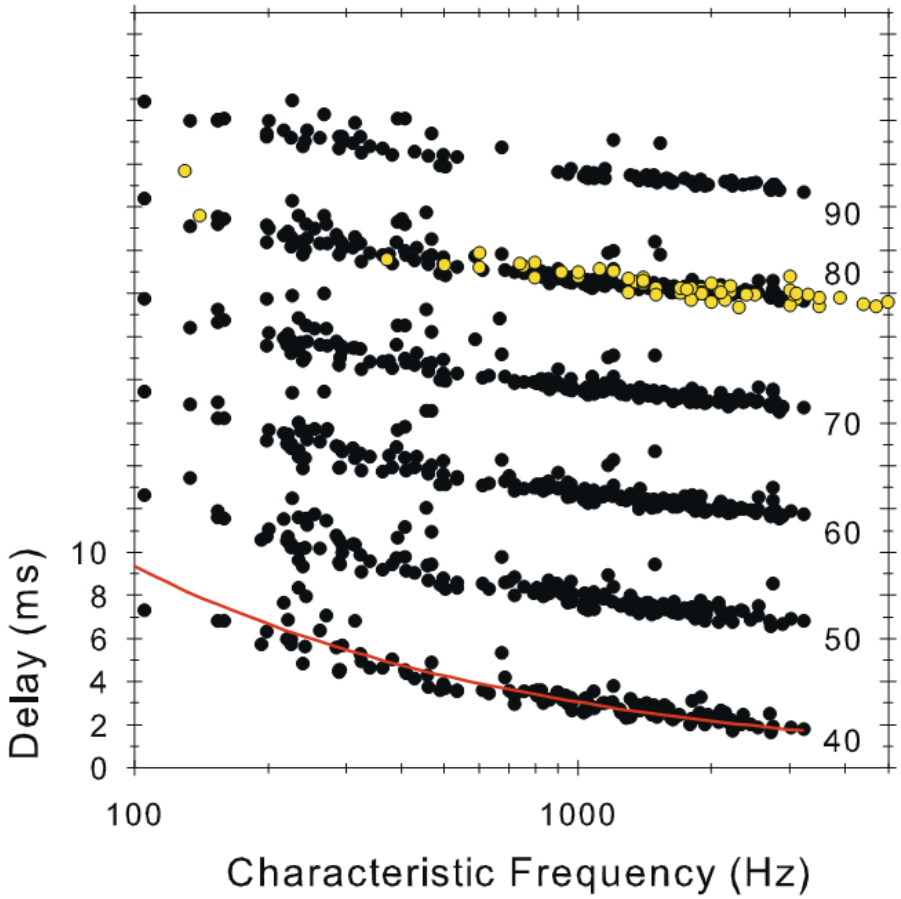

Fig 2.1.4. The delay in spiking of auditory nerve fibres as a function of characteristic frequency and signal level (dB SPL). See text for description. Taken from Palmer and Shackleton, 2009.

For each fibre a number of tones were presented at different frequencies. Frequencies were presented such that the delays across the entire frequency response area were measured. This yields a delay, measured in phase cycles, for each frequency measured. By taking the gradient of this slope a delay, in seconds, was computed. The difference in delay across frequency for each fibre is negligible when compared to the relative differences in delay for fibres with

different characteristic frequencies. As such the delay (in ms) can be considered representative of the actual delay for a single pure tone presented at the characteristic frequency. The black dots are the delays found in the AN fibers in this experiment, the yellow dots those found in a previous experiment and the red line is an equation fitted to both guinea pig and chinchilla data. Predictably the longest delays are found at lower frequencies, as the signal has to travel further along the basilar membrane. At the lowest signal level shown here we can see that from 200 and 3000Hz the change in delay is relatively small ~5ms and at increasing frequencies the rate of change in delay decreases.

Temporal information can be important for the perception of sound in a number of ways. One example is the computation of ITDs in the temporal fine structure or envelope of sounds for use as a localisation cue. The ITD information is derived at the medial superior olive (MSO, discussed later) but in order for this information to be derived it is necessary for the temporal information to be present in the inputs to MSO. In order to test the ability of cells to follow the temporal fine structure of sound signals the phase locking of these cells can be measured. Phase locking is (approximately) the correlation between the firing rate and the phase of the stimulus waveform. Phase locking of the AN has been measured in a number of species. The timing of spikes is completely random above 6kHz in most species (Rose *et al.*, 1967; Johnson, 1980; Sachs and Young, 1980; Palmer and Russell, 1986; Hill *et al.*, 1989), although phase locking goes beyond 10kHz in Barn owls (Köppl, 1997).

An example of the phase locking in the guinea pig AN is displayed in Fig 2.1.5. The phase locking was measured for tones presented at the characteristic frequency of each individual AN fibre. Between 200 and 1000 Hz the synchronisation (vector strength, see Goldberg and Brown, 1969) is stable at ~0.8. At higher frequencies the synchronisation begins to drop until the temporal fine structure is no longer represented at frequencies above ~3kHz.

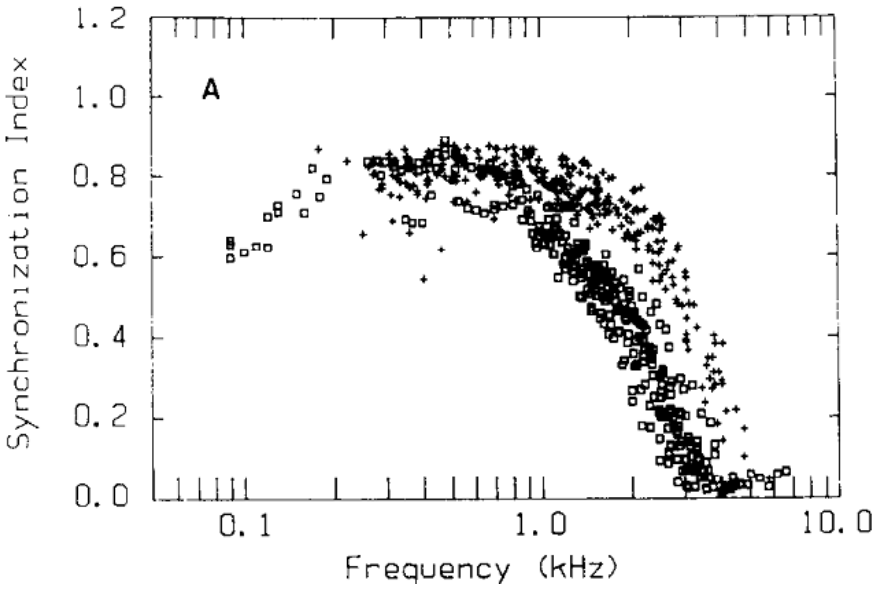

Fig 2.1.5. Phase locking of the auditory nerve as a function of frequency. Open symbols are taken from the guinea pig (Palmer and Russel, 1986), Cross symbols are for the cat (Johnson, 1980), Filled squares are from another study of the guinea pig (Harrison and Evans, 1979). Figure taken from Palmer and Russel, 1986.

2.1.2 Cochlear nucleus

Both the diversity of cell types and physiological responses measured increase in the cochlear nucleus (CN) relative to the AN. The AN branches both caudally and rostrally (see Fig 2.1.6) as it enters CN. The rostral branch connects to cells in the anteroventral cochlear nucleus (AVCN), while the caudal branch connects to both the posteroventral division of the cochlear nucleus (PVCN) and the dorsal cochlear nucleus (DCN). Cytoarchitectural differences exist in differing quantities within different areas of the CN. Large and small spherical cells, globular and small cells are all present in AVCN (Osen, 1969). Multipolar, octopus, small and globular cells are present in PVCN (Osen, 1969). Giant, pyramidal, small and granular cells are all present in DCN (Osen, 1969).

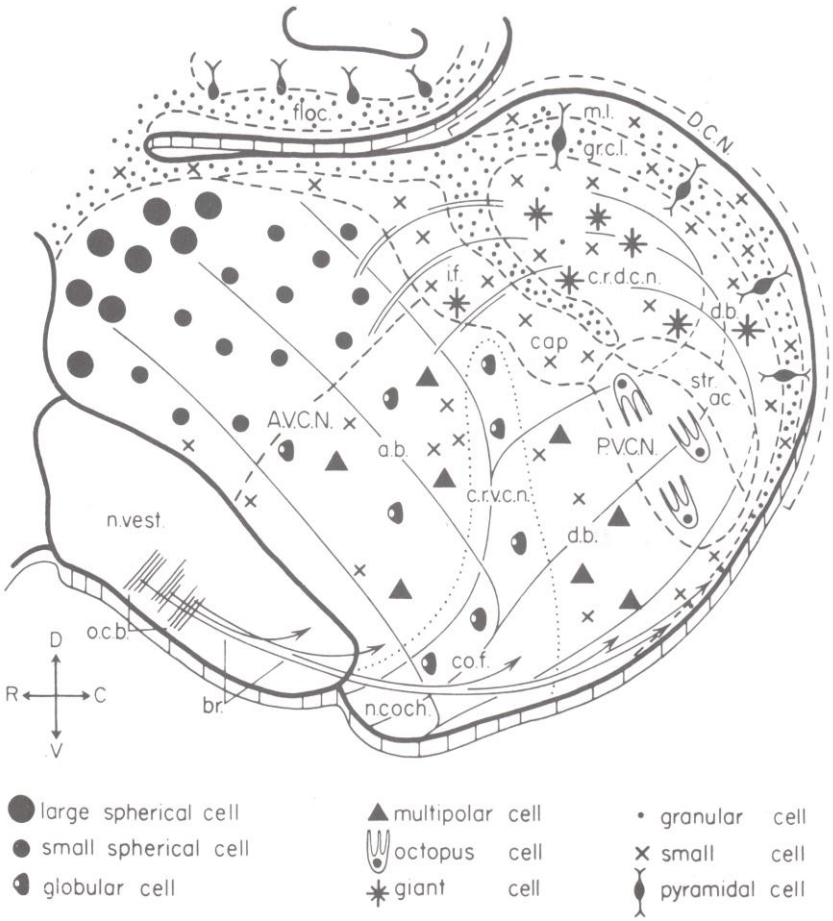

Fig 2.1.6. A cytoarchitectural map of the cochlear nucleus in the cat. Symbols indicate specific cell types (discussed in the text). The predominant cell types for each region are drawn. Taken from Pickles, 1982.

Electrophysiological classification of the response of CN cells to sound has been conducted principally on the basis of PSTHs to pure tones. They are split into at least 4 (further subdivision is possible) classes (Pfeiffer, 1966). The first class is *primary-like cells* these are similar to those of the AN with an initial peak, a rapid decline followed by a gradually declining sustained response. These cells are known to be contacted by few, very large auditory nerve synapses, ensuring a largely faithful relay of the incoming auditory nerve inputs. These units are found throughout the ventral cochlear nucleus (VCN), the responses have been associated with anatomically verified spherical bushy cells (Pickles, 1982). Another class are the *onset cells*. These cells demonstrate a sharp peak in the PSTH at the onset of a sound followed by a decline to no response or a very low

sustained response. This response type is found throughout CN, responses of this type has been associated with in octopus cells (Rhode *et al.*, 1983), and arises partly from the integration of a large number of very small inputs from many small auditory nerve synapses. *Chopper units*, which are also contacted by a relatively large number of auditory nerve synapses, fire repeatedly and regularly during tone burst stimulation. The rate of this repetition is intrinsic to the cell and not related to the period of the pure tone. These responses are found throughout the CN and not restricted to one particular division, though they are strongly represented in PVCN and the deep layers of DCN (Godfrey *et al.*, 1975; Rhode and Smith, 1986). *Pauser and buildup cells* demonstrate an initial onset followed by a period of no activity then followed by a gradual build-up of activity. Fusiform (pyramidal) cells of this type have been found (Godfrey *et al.*, 1975; Rhode and Smith, 1986). Thus overall there is a remarkable correspondence between different cell types of response to sound, cell morphology, and location within the CN.

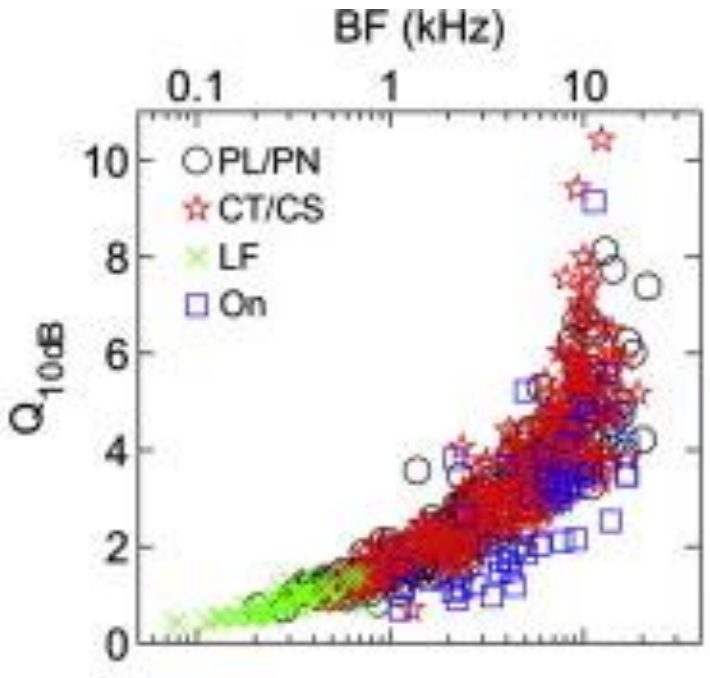

Fig 2.1.7. Q10 values for cochlear nucleus cells with respect to their best frequency (BF). Black circles are primary type response cells (Primary-like and Primary-Notch), red stars are chopper response cells (Chopper-transient and Chopper-sustained), green crosses are low frequency cells and blue squares indicate onset units. Taken from Sayles and Winter, 2010.

The tuning width of cochlear nucleus cells in the guinea pig appears to be slightly broader than in the auditory nerve (Fig 2.1.3 and Fig 2.1.7). The majority of the units from the auditory nerve data have higher Q10s than those found in the cochlear nucleus. The shapes of receptive fields are also known to be more complex in the cochlear nucleus when compared to the auditory nerve (Pickles, 1982).

2.1.3 Superior olivary complex

The superior olivary complex (SOC) can be subdivided into 9 separate subdivisions the: dorsolateral peri-olivary nucleus (DLPO), dorsomedial peri-olivary nucleus (DMPO), dorsal peri-olivary nucleus (DPO), lateral superior olivary nucleus (LSO), lateral nucleus of the trapezoidal body (LTB), medial pre-olivary nucleus (MPO), medial superior olivary nucleus (MSO), medial nucleus of the trapezoid body (MNTB) and the ventromedial peri-olivary nucleus (VMPO, , Harrison and Evans, 1979). The superior olivary complex is believed to be the initial site for encoding ITD and ILD as it is one of the earliest sites for convergence of from the two ears (Goldberg and Brown, 1969; Boudreau and Tsuchitani, 1970). It is also the site of origin of descending inputs to the cochlea (the olivocochlear system). The two nuclei of interest are the MSO and LSO therefore the discussion here shall focus on these areas.

2.1.3.1 Lateral superior olive

The LSO is believed to calculate ILDs (Boudreau and Tsuchitani, 1970). The duplex theory states that ILDs are useful at high frequencies whereas ITDs are useful at low frequencies. This would lead one to assume that in a centre for ILD processing one might expect high frequency sound to be better represented than low frequency sound. This idea holds as the LSO is biased toward representing high frequencies (Tsuchitani and Boudreau, 1966; Guinan *et al.*, 1972; Tsuchitani, 1977). In order to make interaural level comparisons the LSO

would need some mechanism for doing so. This can be explained by the nature of the inputs received by the LSO. LSO cells receive excitatory inputs from the ipsilateral ear through the spherical bushy cells of the ipsilateral AVCN (Stotler, 1953; Warr, 1966; Osen, 1969; Cant and Morest, 1977; Smith *et al.*, 1993). LSO cells also receive inhibitory inputs from the ipsilateral MNTB (Harrison and Warr, 1962; Warr, 1972; Spangler *et al.*, 1985; Friauf and Ostwald, 1988; Smith *et al.*, 1991; Smith *et al.*, 1998). The inhibitory ipsilateral inputs from MNTB are excited by globular bushy cells of the contralateral AVCN (Harrison and Warr, 1962; Warr, 1972; Smith *et al.*, 1991). As discussed the AVCN receives its inputs from the AN of each ear and therefore the LSO receives converging inputs from each ear from an inhibitory contralateral AVCN input via MNTB and an ipsilateral AVCN input. This can be viewed as a simple subtraction in order to compute ILDs, e.g. when the signal is more contralateral the response will be more excitatory and when the signal is more ipsilateral the response will be more inhibited. This is borne out in the responses of LSO cells giving a rate/ILD function, an example of which can be seen in Fig 2.1.8 (Tollin and Yin, 2001).

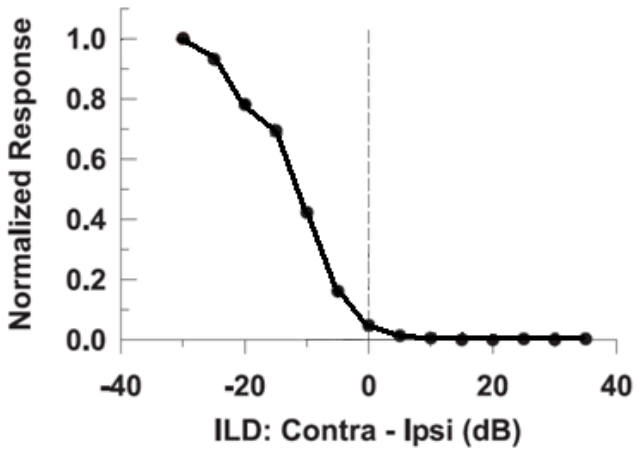

Fig 2.1.8. The average ILD tuning of LSO cells in the ipsilateral side of LSO. The firing rate is normalised relative to the maximum firing rate. Note that at the level of LSO ipsilateral sounds evoke an excitatory response. This switches over hemisphere at the level of the inferior colliculus.

The contralateral pathway is understandably longer than that of the ipsilateral pathway as it needs to travel from the opposite side of the head. This should in turn mean that the arrival times of signals from the left and right ear should differ leading to an inability to make level comparisons. Despite this difference in pathway length inhibitory responses can arrive before that of excitation when the ears are simultaneously stimulated (Tsuchitani, 1988; 1994). This is believed to be due to at least two reasons (Kuwada *et al.*, 1997); the first is that the axons of the contralateral globular bushy cells are 3 times larger in diameter than the ipsilateral spherical bushy cells (Warr, 1966; Warr, 1972). Furthermore the axons of the inhibitory ipsilateral MNTB cells are twice as large as the spherical bushy cells of the ipsilateral AVCN (Joris *et al.*, 1998). The second is that globular bushy cells make calyceal-type endings, known as Endbulbs of Held, on MNTB neurons (Warr, 1966; Irving and Harrison, 1967; Warr, 1972). A recent study has demonstrated that on average contralateral inputs reach the LSO within 200μs of ipsilateral inputs (Joris *et al.*, 1998). This is a relatively small difference when you consider that the contralateral inputs are functionally effective in inhibiting the ipsilateral excitation for ~1 to 2ms (Sanes, 1990; Wu and Kelly, 1992; Joris and Yin, 1995; Park *et al.*, 1996; Irvine *et al.*, 2001).

Another requirement in making ILD computations is that the frequency tuning of the contralateral and ipsilateral inputs be matched (Tollin, 2003). The tonotopic gradients of the LSO, MNTB and AVCN are well established and anatomical studies have demonstrated that inputs particular tonotopic regions of AVCN and MNTB project to corresponding regions in the LSO (Warr, 1966; Elverland, 1978; Spangler *et al.*, 1985; Glendenning and Baker, 1988; Smith *et al.*, 1998). Furthermore electrophysiological study of the contralateral and ipsilateral evoked frequency response areas has revealed that characteristic frequency from each ear are well matched (Boudreau and Tsuchitani, 1968; Caird and Klinke, 1983; Tsuchitani, 1997). Fig 2.1.9 demonstrates the ipsilateral and contralateral tuning of an LSO neuron. The heights of the lines indicate the number of spikes for each frequency/level combination. The ipsilateral tuning

has two troughs around 2 and 11kHz, reflecting two areas of sensitivity these are closely matched in the contralateral ear in both frequency and threshold. While the tuning is well matched the bandwidth of inhibitory contralateral inputs are, on average, slightly wider than excitatory bandwidths.

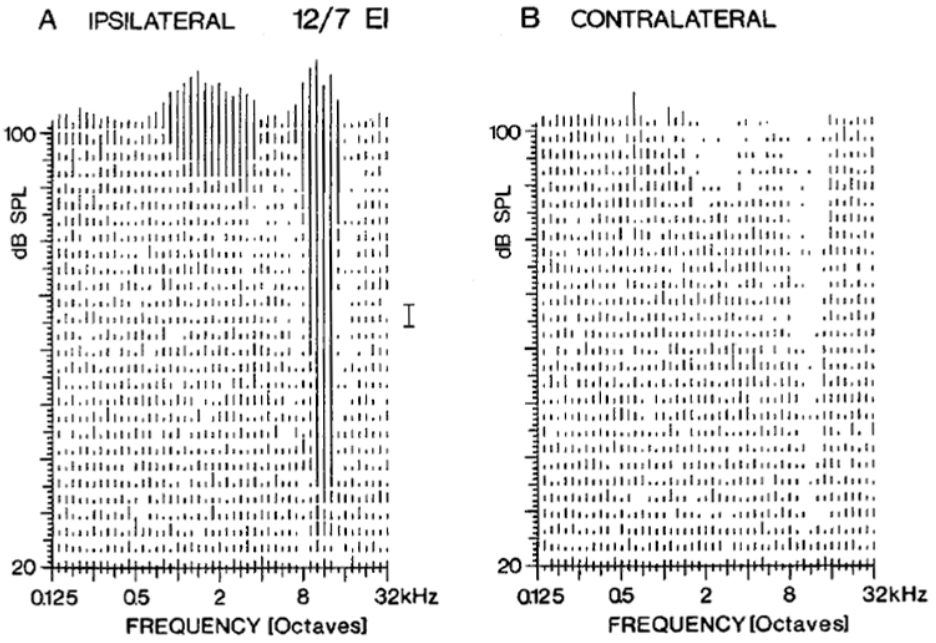

Fig 2.1.9. Monaural frequency tuning of cells in the LSO. The length of bars indicates the number of spikes, the longer the bar the more spikes were evoked. Taken from Caird and Klinke, 1983.

2.1.3.2 Medial superior olive

The MSO is believed to be the earliest and primary site for ITD computation (Goldberg and Brown, 1969). As might be expected from the duplex theory and conversely to the LSO, there is a bias toward representing low frequencies in the MSO (Guinan *et al.*, 1972). MSO receives direct innervations from the ipsilateral and contralateral AVCN (Pickles, 1982), for 75% of cells both of these inputs are excitatory (Goldberg and Brown, 1969). The long held belief has been that delay lines are responsible for a form of cross correlation between contralateral and ipsilateral inputs (Jeffress, 1948). In brief the hypothesis is that by introducing a number of delays to the input from one ear (along a number of delay lines) these delayed copies can then be individually compared

with the input from the other ear. Spiking is integrated for each copy of the signal giving a cross-correlation value at a number of delays. The delay line that yields the best cross-correlation is the one with the delay matching the ITD and this way a number of channels exist, each with their own best delays when a particular channel is very active the listener will know that this is the delay in the signal. There is evidence to show these delay lines do exist though it is not clear whether the delays created are sufficient to explain the available data (Joris and Yin, 2007).

While the mechanism which creates sensitivity to ITDs is not clear what is not in question is that MSO cells are sensitive to different ongoing ITDs (Goldberg and Brown, 1969; Moushegian *et al.*, 1975; Langford, 1984; Yin and Chan, 1990). Fig 2.1.10 demonstrates the change in the evoked spike rate of an example MSO neuron with changes to the ITD. As can be seen there is a peak in firing rate close to 0 ITD with other peaks at ~±2000 μs, the latter being outside of the physiological range of the ITD created by the guinea pig head.

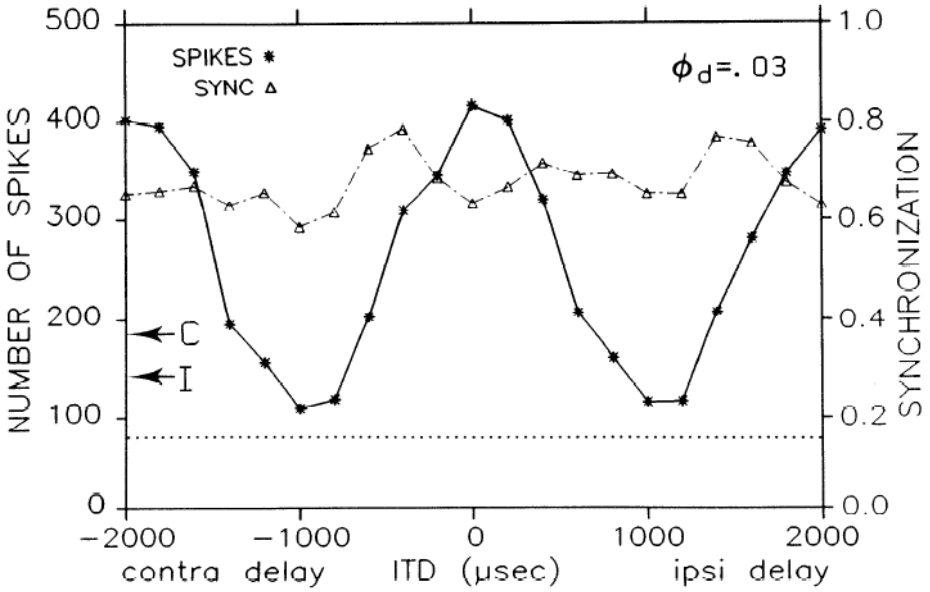

Fig 2.1.10. Interaural time delay curves for spike counts and synchronisation. Arrows indicate the monaural rate of the contralateral (C) input and ipsilateral (I) input. Taken from Yin and Chan, 1990.

2.1.4 Inferior colliculus

The Inferior Colliculus (IC) is generally partitioned into three major subdivisions: the central nucleus, dorsal nucleus and external nucleus (Morest and Oliver, 1984). While the boundaries and number of subdivisions change with species generally three subdivisions have been defined in a number of species: the central nucleus, the dorsal cortex and the paracentral nuclei. The following discussion will not focus on specific areas but treat the IC as one for the sake of ease.

The IC is believed to have the most diverse connections of any auditory structure (Winer and Schreiner, 2005). It receives projections from most regions of the cochlear (Oliver, 1984; 1987), from a number of regions of the SOC (Glendenning *et al.*, 1992), from each lateral lemniscus nuclei (Saint Marie *et al.*, 1997), and from every cortical area (Winer *et al.*, 1998). In addition the IC projects to most nuclei that project to it and also send connections bilaterally to the MGB (Huffman and Henson Jr, 1990). The IC is a key structure in both the ascending and the descending auditory system and is a major site of converging inputs in the auditory system.

As discussed previously, in the LSO neurons ipsilateral inputs excite and contralateral inputs inhibit, known as IE cells. Conversely in the IC contralateral inputs excite and ipsilateral inputs inhibit, known as EI cells. Each hemisphere of the IC receives ipsilateral input from the nucleus of the lateral lemniscus, contralateral inputs from the dorsal and posteroventral cochlear nucleus, and both bilateral input from the LSO and MSO (Pickles, 1982). The contralateral projection from LSO have been shown to be larger than the ipsilateral projection leading to an increased likelihood of finding cells with EI interactions (Fuzessery and Pollak, 1985; Semple and Kitzes, 1987). However it is an oversimplification to state the EI interactions encountered in the IC are solely attributable to the input of IE units from the contralateral LSO cells, as there is

evidence to suggest that some EI interactions are created at the IC by combination of excitatory inputs from various nuclei and inhibitory inputs from the dorsal nucleus of the lateral lemniscus (Sally and Kelly, 1992; Klug *et al.*, 1995; Kuwada *et al.*, 1997). Regardless of the circuitry leading to the EI interactions EI interactions are most common for binaural units at high frequencies, whereas EE type interactions are more prevalent at low frequencies (Semple and Aitkin, 1979; Irvine and Gago, 1990). As with the LSO cells sensitive to ILDs are typically broadly tuned and monotonic (Irvine, 1987; Irvine and Gago, 1990). There is not a satisfactory consensus to suggest that either the encoding of ILD or ITD are refined in the IC. One study has addressed the changes in ILD tuning when signal bandwidth was increased showing cells were more likely to show spatial sensitivity to noise than to pure tones (Aitkin and Martin, 1987). Similarly spatial acuity appears to be improved with increasing bandwidth (Sterbing *et al.*, 2003).

For the guinea pig the three subdivisions defined are the central nucleus (CNIC), this is covered dorsally by the dorsal cortex (DCIC) and laterally and rostrally surrounded by the external cortex or ECIC (Ramon and Cajal, 1909; Morest and Oliver, 1984). The CNIC receives inputs from the cochlear nuclei, superior olivary complex and the lateral lemniscus (Beyerl, 1978; Brunso Bechtold *et al.*, 1981; Druga and Syka, 1984). The DCIC receives inputs from the DCN, AVCN, dorsal nucleus of the lateral lemniscus and the nucleus sagulum (Syka *et al.*, 2000). The ECIC receives inputs from the somatosensory system, the CNIC and the nuclei of the lateral lemniscus (Syka *et al.*, 2000). Syka et al. (2000) investigated the frequency tuning of different subdivisions of IC in the Guinea Pig.

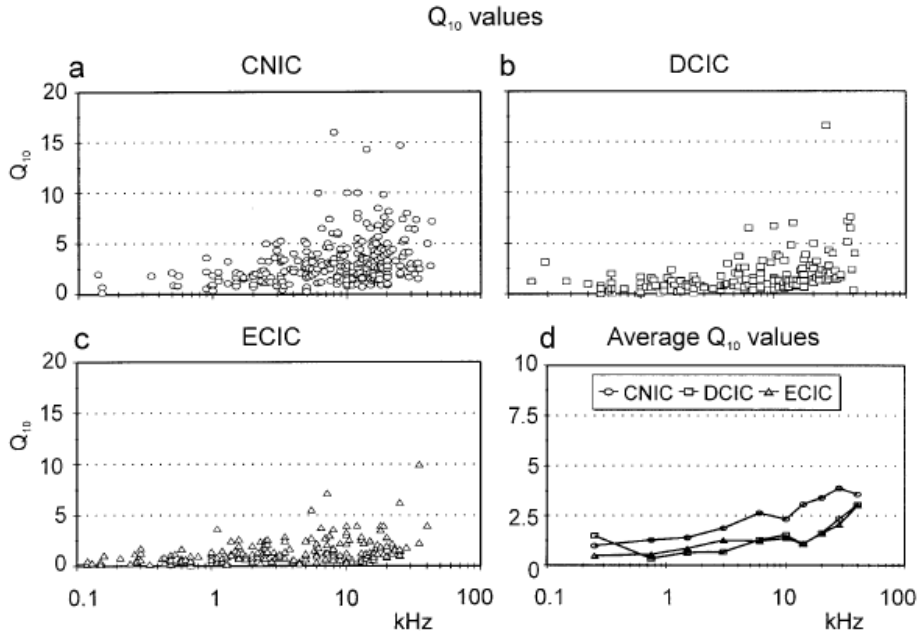

Fig 2.1.11. The relationship between Q10 values and characteristic frequency in IC of the guinea pig. Taken from Syka et. al 2000.

Fig 2.1.11 demonstrates the Q10 values yielded by cells from the different subdivisions of the Guinea Pig IC as well as the average Q10s for each subdivision. The CNIC gives larger Q10 values both on average and the range of possible Q10 values. This is evident by the wider spread of points on subplot a and the larger average values found in subplot d. DCIC and ECIC appear to yield similar Q10 values to one another. In relation to the cochlea nucleus all subdivisions of IC give smaller Q10s, demonstrating that in the Guinea Pig and using a similar anaesthetic regimen the 10dB tuning width is larger in the IC when compared to the cochlear nucleus. This is particularly evident at higher frequencies. For example the average Q10 value at 10kHz for the CNIC (the area with the largest Q10s and hence the narrowest tuning) is 2.5. In the cochlea nucleus Q10s fall between 4 and 10 and hence the average has to be much larger than those in the IC (Fig 2.1.6). Therefore even the area with the narrowest tuning in the IC demonstrates a large increase in tuning width when compared to the cochlea nucleus.

2.1.5 Medial Geniculate Body

As with most other nuclei there are a number of schemes for anatomically defining the subdivisions of the Medial Geniculate Body (MGB). One such scheme suggested suggests three subdivisions: the ventral, dorsal and medial divisions. The ventral division appears to receive only auditory inputs the majority of which come from the ipsilateral central nucleus of the IC (Calford and Aitkin, 1983). Both the dorsal and medial divisions receive multisensory inputs (Pickles, 1982). The dorsal division receives afferents from the region medial to the brachium, the superior colliculus, the pericentral nucleus of the IC and the somatosensory system (Harrison and Howe, 1974; Calford and Aitkin, 1983). The medial division receives inputs from the IC, lateral tegmental system and the somatosensory system (Harrison and Howe, 1974; Calford and Aitkin, 1983).

As with the IC cells in the MGB are broadly tuned to spatial location and are predominantly tuned to contralateral locations (Samson *et al.*, 2000). The spatial tuning of MGB neurons is also dependent on bandwidth of the signal with noise stimuli more frequently resulting in spatial tuning than pure tones (Clarey *et al.*, 1995).

2.1.6 Auditory Cortex

The auditory cortex can be functionally subdivided into “primary” and “belt” areas. Primary areas are functionally defined by strong responses to tones, relatively short latencies and a tonotopic gradient. Belt areas are more responsive to broadband stimuli, have longer latencies and either demonstrate a poorly defined tonotopic gradient or demonstrate none at all. The physiological data gathered here are recorded in the primary auditory cortex belt areas are believed to receive some of their inputs from the primary areas and so can be considered “downstream” of primary areas. For this reason the

discussion of the ascending auditory areas will stop at primary auditory cortex (AI).

2.1.6.1 Spatial tuning

As with in the IC and MGB the AI cells are generally tuned to the contralateral hemisphere (Middlebrooks and Pettigrew, 1981; Rajan *et al.*, 1990; Brugge *et al.*, 1996; Middlebrooks *et al.*, 1998; King *et al.*, 2001; Schnupp *et al.*, 2001; Stecker and Middlebrooks, 2003). Similarly AI cells are sensitive to ILDs on the contralateral side of midline or 0 ILD (Kitzes *et al.*, 1980; Phillips and Irvine, 1981). Hemispheric tuning is present in both the firing rate and first spike latency of cortical cells with larger firing rates and shorter first spike latencies for more contralateral sounds (Stecker and Middlebrooks, 2003; Nelken *et al.*, 2005). An example of the representatively broad spatial tuning of cortical cells is provided in Fig 2.1.11.

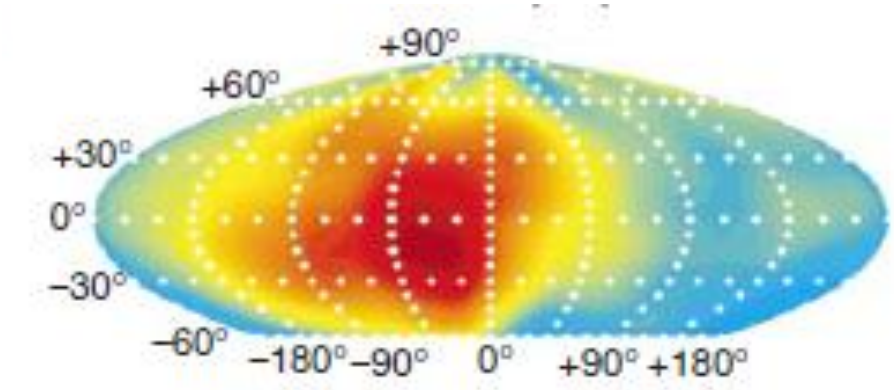

Fig 2.1.12. The spatial tuning of an auditory cortical neuron. Taken from Mrsic-Flogel, 2005.

Recordings were made from the left hemisphere auditory cortex. Azimuth is plotted along the abscissa, elevation on the ordinate. Positive azimuthal values indicate sounds toward the ipsilateral side (left) and negative values toward the contralateral side. The heat map indicates the scaled firing rate where dark red is the maximum firing rate and dark blue the minimum. When a cross section is taken, so that azimuth is considered at one elevation, the unit responds at high

rates for an area larger than 90˚ at most elevations. Though this particular unit appears to be most sensitive to locations between 0 and 45˚. The majority of ipsilateral space fails to evoke much of a response demonstrating the unit is essentially hemispherically tuned.

2.1.6.2 Average binaural level (ABL)

Often ILD functions have been measured by holding the level in one ear and varying the level in the other ear. While this is useful in a mechanistic way to assess the degree of facilitation or inhibition gained it does not reflect the changes in acoustics caused by movement of a sound source in space. As a sound source is moved around a listener in azimuth, at a fixed distance, from 0 to 90˚ the ILD increases in size, however the average of the level at the two ears (known as the average binaural level or ABL) remains approximately constant. Therefore as the sound source moves around the listener the sound level in one ear (ipsilateral) is increased and the sound level in the opposite ear is reduced essentially holding the ABL but adjusting the ILD. In order to study ILD processing under more realistic listening conditions a number of studies have investigated the effect of changing ABL on the encoding of ILD (Semple and Kitzes, 1993; Nakamoto *et al.*, 2004; Campbell *et al.*, 2006).

In terms of behavioural performance ILD lateralisation and high frequency localisation ~20dB above threshold is relatively unaffected by changes in level. Similarly at similar sound levels one might expect to see similar invariance in the ILD tuning of cortical cells. The results are not as simple as this. Increases in ABL can produce an increase or decrease in the ILD tuning width (Semple and Kitzes, 1993; Zhang *et al.*, 2004). This is attributable to non-monotonic relationship between firing rate and ABL (Semple and Kitzes, 1993). Though changes in ILD sensitivity are difficult to interpret it is clear that the rate code to ILD is not level invariant in primary auditory cortex.

2.1.6.3 Binaural interaction classification

Originally to distinguish between the different types of responses found in the SOC two classes were defined (Goldberg and Brown, 1968). Cells with an evoked excitatory response from each ear were classified as EE and cells with one excitatory and one inhibitory input cells were classified as EI (Goldberg and Brown, 1968; 1969). This classification has been inherited in the study of other auditory areas higher up the auditory system. A bewildering array of classification types have been suggested since (Irvine, 1986; Kelly and Sally, 1988; Nakamoto *et al.*, 2004). In attempting to partition responses into statistically verifiable binaural classifications it has been shown that ABL/ILD responses are continuous and do not appear to fall into discrete classes (Campbell *et al.*, 2006). Despite this it is sometimes helpful for the purposes of discussion to keep some form of classification.

2.1.6.4 Binaural response at different frequencies

One study has measured changes in the ABL/ILD responses with changes in frequency (Kitzes, 2008). The author measured ABL/ILD response maps in cat auditory cortex with tones at a number of frequencies, where the frequency was matched in both ears. Fig 2.1.13 is an example unit from this study. The cell was defined as EI with a CF of 7.5kHz. The grey scale maps demonstrate the firing rate for a given ipsilateral/contralateral level combination, each panel is the receptive field at a different test frequency. The size of the binaural level receptive fields varied with frequency with RFs close to CF demonstrating the largest receptive field. This is also evident in the spikes versus frequency plot for a given ABL/ILD combination (Fig 2.1.13, bottom right). The rate for most ABL/ILD combinations was highest at or close to 7.kHz, the CF of the cell.

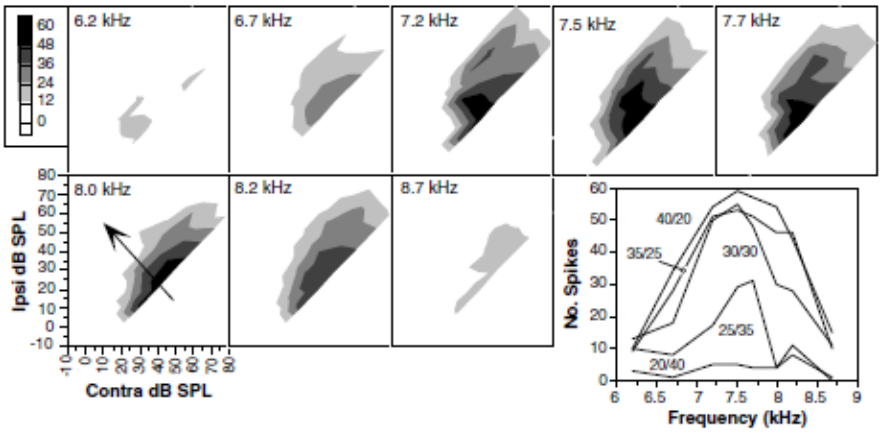

Fig 2.1.13. The ILD tuning of a cortical neuron to tones presented at different frequencies. Taken from Kitzes, 2008.

2.2 Ipsilateral frequency response areas

2.2.1 Introduction

Increasing the bandwidth of a signal leads to improvements in localisation performance in mammals (Terhune, 1974; Brown *et al.*, 1980; Butler, 1986; Martin and Webster, 1987). Similarly the neural representation of high frequency spatial cues has also been shown to be more sensitive at broader bandwidths (Aitkin and Martin, 1987; Rajan *et al.*, 1990; Clarey *et al.*, 1995; Sterbing *et al.*, 2003). Exactly how additional spectral information is integrated to change the spatial representation is unclear. Models of ITD lateralisation compute ITD cues in independent frequency channels and then apply a weighting function to these channels to produce accurate estimates of lateralisation performance (Stern *et al.*, 1988; Shackleton *et al.*, 1992). It has been suggested that ILD cues might also be processed in discrete frequency channels and then integrated (Middlebrooks, 1992).

There is some biological evidence to suggest that binaural information is computed over restricted frequency ranges. It is believed that binaural level comparisons are predominantly made at nuclei lower down in the ascending auditory system at the LSO (Pickles, 1982). It is well known that Frequency Response Areas (FRAs) become more complex after the auditory nerve. Experiments in the guinea pig using urethane anaesthesia have shown decreases in Q10 values between lower and higher levels in the auditory system (see section 2.1). Therefore at lower levels, like the LSO, frequency tuning is relatively narrow leading to binaural interactions being computed over narrow frequency ranges. Moving up the auditory system the frequency tuning becomes broader suggesting greater frequency integration. As the ILD comparisons primarily occur at LSO, where binaural interactions occur over narrow frequency ranges, later auditory nuclei will inherit this ILD encoding. At

the level of cortex frequency tuning is broader but the ILD comparisons have already been made over narrow frequency ranges suggesting that the ILD computations should be restricted to discrete frequency ranges within the more broadly tuned cortical cell.

Recent work has suggested that at high frequencies neural spatial sensitivity can be accurately predicted by a linear combination of discrete frequency bands for each ear (Schnupp *et al.*, 2001). Schnupp et al. (2001) presented random chord sequence stimuli to collect binaural frequency-time receptive fields (FTRF), i.e. the evoked response at a given time by a given frequency. These were generated independently for each ear such that a short tone pip (20 ms, 5 ms rise/fall time) could be presented in any one of 60 frequency bands between 0.5 and 32 kHz where each frequency band was 1/10 octave wide. On average between 2.4 and 4.8 pips would be present during a 5 ms time interval, the tone onsets were statistically independent. The two independently generated sequences were then presented simultaneously and responses of AI cells recorded.

Spike triggered averaging (STA) was used to generate the FTRF for each ear. The stimuli were presented binaurally but the STA was only correlated with the monaural input within one frequency band. Therefore the measured response to a given frequency in a given ear was averaged across all other contralateral and ipsilateral stimuli, which included tone pips at frequencies in different frequency bands. Therefore this average includes the across frequency contributions of both contralateral and ipsilateral sounds to the response in each frequency band. The across frequency interactions are linearly combined to give the STA. The relative monaural contribution, whether facilitatory, inhibitory or otherwise, was captured within the FTRF. The authors first measured individualised Directional Transfer Functions (DTFs), which are similar to HRTFs but exclude non-directional contributions like the ear canal. DTFs were then used to simulate the acoustic energy over time reaching each ear for

spatial stimuli. This was then filtered by the FTRF for each ear, linearly weighted across frequency and summed across the two ears to yield a predicted Spatial Receptive Field (SRF). This linear method for predicting SRFs produces strong correlation between predicted and observed SRFs, it is particularly good at predicting SRFs for units classified as EI units (Mrsic-Flogel *et al.*, 2005).

Similar to this Euston and Takahashi (2002) directly measured ILD tuning of pure tones at different frequencies along the frequency response area of IC cells in the Barn Owl. They then linearly combined the tone responses across frequency and used this to predict the ILD response of cells to broadband sound. They found good correspondence between predicted and independently measured ILD tuning for broadband stimuli (Euston and Takahashi, 2002). This demonstrates that even when across-frequency interactions are ignored good predictions of ILD tuning can be made. This suggests that ILD could be computed over discrete frequency ranges.

The hypothesis that ILD is computed over discrete frequency ranges in the auditory cortex, even after spectral integration has occurred, was tested by investigating the frequency range over which ILD information is computed at different points along the frequency responses area. At the level of cortex and at high frequencies the majority of cells are hemispherically tuned with excitatory contralateral inputs and inhibitory ipsilateral inputs. To test the frequency range over which ILD could be computed contralateral tones were presented at one frequency and ipsilateral tones at a range of frequencies such that the inhibitory ipsilateral frequency response areas (iFRAs) were measured. One would predict that if ILD functions were calculated over narrow frequency ranges, the area of the inhibition in the iFRA would be small and restricted to the contralateral and neighbouring frequencies. As the contralateral tone frequency changes one would expect the centre of the inhibitory region to shift in tandem. Otherwise we might expect the iFRA to match the contralateral frequency response area (cFRA) regardless of the contralateral sound. Simpler

paradigms could be used to investigate binaural frequency comparisons but by sampling large frequency and level ranges it should be possible to gain a "big picture" of the important issues for collecting this data and give some idea as to whether an effect is indeed present. This would allow further more specific investigation at a later point. A point of note is the use of the term "ipsilateral" frequency response area. While this is a binaural response for the sake of brevity and ease this binaural response will be referred to as the ipsilateral response or iFRA.

2.2.2 Methods

2.2.2.1 Subjects

Subjects were tricolour pigmented guinea pigs (Cavia porecellus). Guinea pigs (GPs) were housed in gender specific home cages under standardised lighting conditions (12 hours of light/ 12 hours of dark). The temperature was maintained at 20 ± 2˚c and relative humidity was approximately 55 ± 10%.

2.2.2.2 Surgical Procedure

Neurophysiological recordings were conducted in a sound attenuating chamber (Whittingham acoustics Ltd). Animals were anesthetised initially with an intra-peritoneal injection of Urethane (0.9 – 1.3 $g.kg^{-1}$ in a 20% solution, Sigma). Dose of anaesthetic was gauged according to weight (approximately 4.5 ml/kg). Upon induction of anaesthesia 0.2 ml of Atropine Sulphate (600$\mu g.ml^{-1}$ Pheonix Pharma) was administered sub-cutaneously to suppress bronchial secretions. During the course of the experiment the forepaw withdrawal reflex was tested every 30 minutes. If a withdrawal reflex was evoked supplementary analgesia was administered. This consisted of an intramuscular injection of 0.2-0.3 ml of Hypnorm (Fentanyl citrate 0.315$mg.ml^{-1}$, Fluanisone 10$mg.ml^{-1}$, Janssen). Upon induction of surgical anaesthesia the animal was placed on a thermal blanket (Harvard Apparatus) and the throat, ears and head were shaved using animal

clippers. The tragi were removed bilaterally to enable ease of access to each ear canal, the ear canal was then inspected under microscope and any blockages or ear wax removed. The trachea was then cannulated using polythene tubing (Portex), this had an outside diameter of 2.4mm and an inside diameter of 2mm.

The animal was then placed on a specifically designed platform with accompanying stereotaxic head frame. The animals top jaw was placed around a bite bar and a nose plate was fitted over the top jaw to keep head position. Perspex specula were then placed at the beginning of the ear canal and aligned, using a microscope aimed through each speculum, such that there was a clear view of the tympanic membrane. The tracheal tube was then connected to a respiratory pump supplying 100% oxygen (BOC) mixed with air, breathing was controlled and monitored throughout the experiment. The end tidal $C0^2$ was maintained within the normal range of 28-38 mm Hg, respiratory rate and volume were controlled to ensure steady breathing. At this point a rectal thermometer was placed using petroleum jelly (Vaseline) as a lubricant. The temperature of the heating mat beneath the animal was monitored and maintained using a feedback system between the homeothermic blanket control unit (Harvard Apparatus) and rectal thermometer keeping the animal at a core temperature of 38˚c. An electrocardiogram was used to monitor cardiac function with electrodes inserted to the skin either side of the thorax.

To expose the skull an incision was made along the sagittal midline and the skin cleared from the skull using blunt dissection. The periostium and temporalis muscles were removed using a cortary tool and any tissue remaining on the skull cleared. Middle ear pressure cannot be equalised under anaesthetic. Therefore a small hole was made in the auditory bulla and a length of 0.5mm diameter polythene tube inserted into each, a seal was made between the tube and the skull using petroleum jelly (Vaseline). To increase the stability of the recordings efforts were made to reduce pressure of the cerebro-spinal fluid.

Connective tissue above the foramen magnum was cleared and a small hole made in the *dura mater*.

Primary auditory cortex (AI) is located caudal and ventral to the Pseudo Sylvian sulcus in the Guinea Pig. Before the craniotomy was performed an outline marking the boundaries of the underlying primary auditory cortex was pencilled on the skull (Wallace *et al.*, 2000). The overlying morphology of the skull is a reasonable indicator of the position of the underlying cortex. Thus, the rostral boundary of the craniotomy was marked along the coronal plane with a line 1mm rostral to the frontal edge of Bregma, this line was marked 1 mm in front of the Pseudo Sylvian sulcus. A parallel line was then marked 5mm caudally to the initial line this marked the caudal boundary of AI. The dorsal boundary was marked parallel to the midsaggital plane 10 mm ventral from the sagittal midline. The ventral boundary was marked parallel to the midsaggital plane a further 6 mm ventral of the previous line. A craniotomy was then performed using a dental drill (Minitech Dentimex) and the *dura mater* subsequently removed. As mentioned primary cortex lies caudal-ventral to Pseudo Sylvian sulcus with the lateral suture overlying the middle of AI. These markers were used to place the electrode. In addition to anatomical markers physiological markers such as response latency and responses to noise and tones were used to identify primary auditory cortex. The tonotopy of AI is arranged such that frequency ascends caudally; so if sampling of higher frequencies was required the electrode could be moved caudally. The tonotopy reverses at the border between AI and DC: if any such reversal in the tonotopy was found then the electrode was moved rostrally to ensure it was in AI. Once the electrode was positioned the exposed cortex was covered with agar (1.5 % agar in 0.9% saline).

2.2.2.3 Stimulus generation and delivery

Fig 2.2.1 is a schematic diagram of the input and output hardware used for physiological recordings. The stimulus generation program was written in RPvdsEx graphical programming interface.

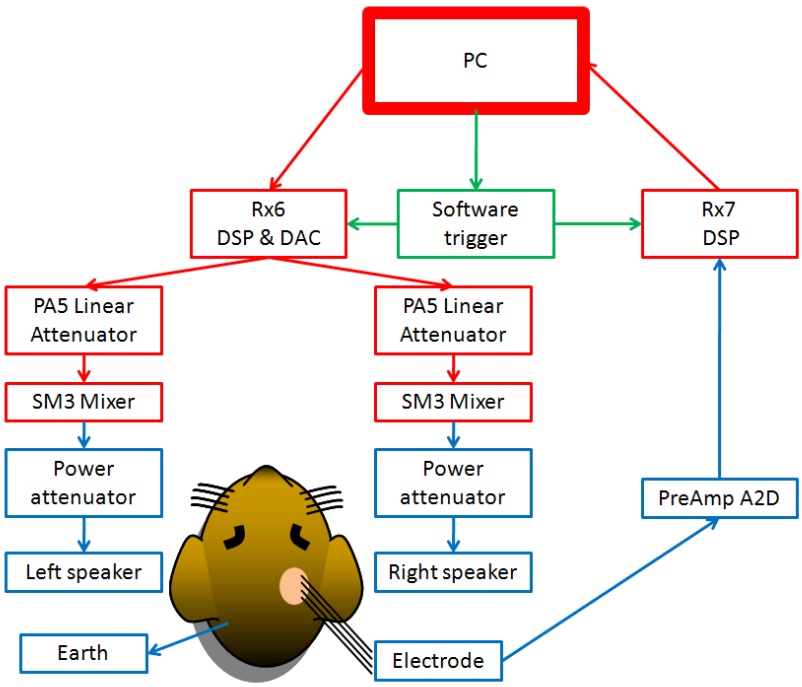

Fig 2.2.1. Schematic diagram of the auditory stimulation and neurophysiological recording equipment (see text for detail).

A Viglen PC (Pentium III, 600MHz CPU and 256 MB RAM) loaded this program into the RX6 digital signal processor (Tucker-Davis Technologies) via the recording program (BrainWare). Under the control of BrainWare the RX6 was used to generate auditory stimuli (noise and tones). The signal was then passed through a 24 bit sigma delta digital-to-analogue (D2A) converter before being attenuated with the programmable attenuator version 5 (PA5). The attenuated signal then passed through SM3 mixers. The power attenuator shown in the schematic could be programmed before the beginning of a recording session, the level of attenuation was determined from the sound calibration. Stimuli were then presented binaurally through custom-modified Radioshack 40-1377

tweeters (M. Ravicz, Eaton Peabody Laboratory, Boston, MA) these were held by the hollow specula leading to the respective tympanic membranes.

2.2.2.4 Sound system calibration

The calibration files generated were used to adjust the attenuation of the signal to ensure that equal dB SPL signal levels were presented for tones of differing frequencies. Stimulus level was controlled online by varying the attenuation from the PA5 attenuators. To calibrate the system full level signals, no attenuation from PA5 and the power attenuators were presented giving a level of 120dB SPL. The signals were white noise bursts presented individually to each ear. The signals were recorded using a condenser microphone (Bruel and Kjaer 4134) with an attached probe tube. The microphone was inserted through a small hole in the specula so that it was between the tympanic membrane and the speaker approximately 3mm from the tympanic membrane. The noise bursts was presented 20 times and then the averaged Fourier transform was recorded. The microphone and probe tube characteristics were saved. A Matlab program was written to take the sound presentation characteristic and adjust levels so that the output signal was always at the required dB SPL, this program automatically selected the correct signal levels for each frequency and created the BrainWare file to be run for each unit.

2.2.2.5 Neurophysiological recordings

Neurophysiological recordings were conducted using in-house manufactured multi-electrodes arrays (Bullock *et al.*, 1988). Glass insulated tungsten tips were manufactured specifically to have small exposed tips (8-12 µm) to ensure good isolation. These electrodes were then mounted onto a printed circuit board (usually 4-6 electrodes to each multi-electrode array) and held using epoxy resin. Each electrode was then individually connected to an individual channel of the circuit board and a conductive seal was made between the circuit board

and the electrode using electro-conductive paint. The printed circuit board had 16 channels available, as less than 8 channels were needed for electrodes, the remaining 8 channels were wired together and used as an earth.

In addition to the sound delivery system Fig 2.2.1 also demonstrates the recording system for neurophysiological recordings. The signals from the multi-electrode array were band-pass filtered (0.16 -6000 Hz) using a high impedance head-stage (TDT RA16AC) and digitised using a pre-amplifier (TDT RA16PA). This signal was then passed to a digital signal processor (RX7) where the signal was again filtered under the control of BrainWare (400-3000 Hz). Recordings were monitored online using BrainWare. Only activity which exceeded a certain threshold, in amplitude, were recorded. This recording was temporally windowed around the spike and lasted 1105 µsec. Spike shapes could be monitored using BrainWare and if required they could be viewed on an oscilloscope or listened to via loudspeaker.

2.2.2.6 Search stimulus

The search stimulus was generally a broadband noise bursts (48 kHz wide, 100 ms duration, 2 ms $\cos^2$ on-off ramp, 70 dB SPL) this was presented either monaurally (either contralaterally or ipsilaterally) or binaurally, occasionally tones were used. Once some response was elicited a frequency response area (FRA) was gathered by presenting single tones of 100ms duration (6ms $\cos^2$ on-off ramp) at a rate of 1 per second at a number of frequencies and levels (level generally ranged between 20 and 80 dB SPL).

2.2.2.7 Spike Sorting

Though the impedance of the small tipped electrodes allowed for good isolation units were still sorted to ensure that the data collected were from single units and not the combined response of multiple units or just noise. Spike sorting

was conducted on the waveforms of the recorded data. Before this could be done the data was converted from the recorded BrainWare format into a .plx file using Matlab to be used in the Plexon offline spike sorting software. A number of different dimensions could be studied at a single site, depth and on a particular channel, these included the receptive field, rate level functions and potentially multiple variations of the relevant stimulus paradigm. Therefore to ensure that the same unit was observed across all of these different manipulations all of the waveforms were packaged together into a single file using Matlab before sorting.

Waveforms were aligned by their maximum peak before Principal Component Analysis (PCA) was performed. PCA is a statistical technique which aims to explain the data, in our case waveforms, in terms of a weighted function of the principal components. The principal components themselves are derived from the data set by finding a vector which explains the most variance in the data, this would be labelled as PC1. All of the vectors in PCA are selected as to be orthogonal to each other and so the PC2 is selected to be a vector orthogonal to PC1 which explains the second most amount of the variance, the same process is applied to give the desired amount of principal components. Plexon selects the first 3 principal components and then for each waveform attributes a weight value for each principal component so that $W_n = w_{n1}PC_1 + w_{n2}PC_2 + w_{n3}PC_3$. Where W_n is the new approximation of the n^{th} waveform in principal components, w_{ni} is the weight attributed to PC_i of the n^{th} waveform. An error function is used to find the weights that minimize the mean square error between the original waveform and the new waveform W_n. Once the waveforms are represented as weighted sums of PC vectors they can be plotted as clusters in two or more dimensions and sorted according to the grouping of these clusters (Fig 2.2.2).

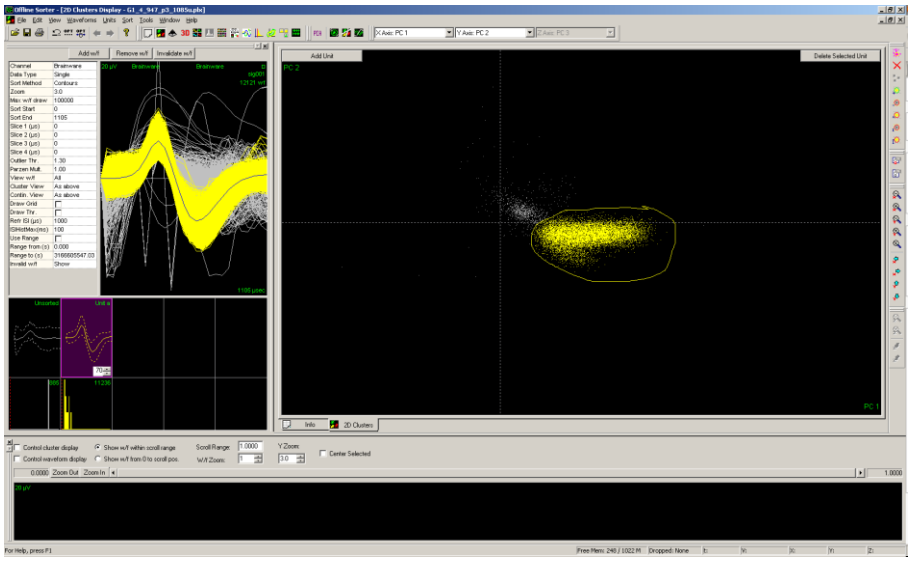

Fig 2.2.2. An example of spike sorting using Plexon. PCA has been run on the waveforms (top-left box) and then the PCs plotted in PCA space with the first two PCs (top-right). On this basis it is easy to separate out the real spike to the noise (highlighted yellow is the spike and grey is the noise).

2.2.2.8 Data collection

2.2.2.8.1 Collecting the monaural contralateral frequency response area

Once a single unit had been well isolated on at least one electrode data collection began. Initially a frequency response area was taken. Pure tones of 100 ms duration, with a cosine on/off ramp (6 ms) were presented monaurally to the contralateral ear to the recording site. As frequency tuning varies little within a cortical column previously recorded frequency response areas could often be used to inform the selection of frequencies for the current unit. If this selection proved to be inappropriate a second frequency response area would be taken to ensure as much as possible of, generally all of, the frequency response area was recorded. Generally two repeats was sufficient to characterise frequency tuning, however, between 3 and 5 repeats were collected.

As the tuning width varies from unit to unit it was necessary to change the resolution or frequency step size in order to gain an informed view of the frequency response area.

2.2.2.8.2 Collecting ipsilateral frequency response areas

Once the monaural contralateral frequency response area had been measured two or three contralateral tone frequencies were chosen. Attempts were made to restrict contralateral tones to frequency/level combinations within the frequency response area of the contralateral tuning. As 4 channel multi-electrode arrays were used to record data there were often 2 or more units present across the different electrodes, each unit with its own frequency tuning. This meant that a decision had to be made as to which unit the contralateral tone frequencies were catered for. This in turn meant that for the other units recorded the frequency/level combinations would be non-optimal as they were not specifically chosen for this unit. For the selected units the frequency/level combinations of the contralateral tones were selected such that the one frequency was presented at characteristic frequency, one above and below and these were matched (as far as could be) for rate by adjusting the level and frequency.

The frequency/level combinations used for the monaural contralateral frequency response area were also used for presentation to the ipsilateral ear. In this way for each selected contralateral tone condition a different ipsilateral frequency response area (iFRA) was measured. Contralateral and ipsilateral tones were gated on and off simultaneously both with 100 ms duration and a cosine on/off ramp of 6 ms. Binaural frequency/level combinations were presented in random order for as long as the unit would still respond.

2.2.2.9 Data analysis

2.2.2.9.1 Contralateral frequency response area

An algorithm was coded in Matlab to establish the threshold of cells at each frequency. The algorithm was similar to one used before in previous studies (Sutter and Schreiner, 1991). Before the frequency tuning curve (FTC) algorithm was applied the receptive fields were filtered using a 3 x 3 window applied across the entire receptive field. The window had a pyramidal weighting, i.e. the middle element had the highest value. This filtering was necessary because the frequency tuning was measured using very few repeats and so it was possible to encounter extreme values attributable to noise, these could affect the resultant FTC.

A frequency/level combination was considered to have evoked a response if the firing rate was larger than the spontaneous rate + criterion value * (maximal response – spontaneous rate). The spontaneous rate was the rate average of the lowest level used for each frequency (Sutter and Schreiner, 1991). The criterion value was set as 0.4. This algorithm essentially traces around the FRA at these rates yielding an estimate of tuning.

2.2.2.9.2 Unit and iFRA classification

Fig. 2.2.3 displays an example of the tuning properties measured under different conditions. The top plot shows the monaural FRA for the contralateral ear. The black line indicates the tuning threshold derived using the algorithm above. Markers (+,o,•) indicate the level and frequency of the contralateral tones presented with the ipsilateral Frequency Response Areas (iFRAs). As can be seen the tuning is well defined and can be easily distinguished. The bottom three plots are the iFRAs these were normalised by dividing the iFRA firing rates by the firing rate when the contralateral tone was presented on its own.

Therefore a firing rate >1 indicates facilitation and a firing rate <1 indicates inhibition of the contralateral alone firing rate. The obtained iFRAs were generally poorly defined; the left plot shows inhibition in the centre (between 12-16kHz) with facilitation at higher frequencies (>22kHz), the middle plot appears to fluctuate at random and the right most plot demonstrates inhibition in the centre (12-22 kHz) and excitation above and below this. An adapted version of the algorithm used to characterise contralateral tuning was applied but frequently produced poor estimates of tuning.

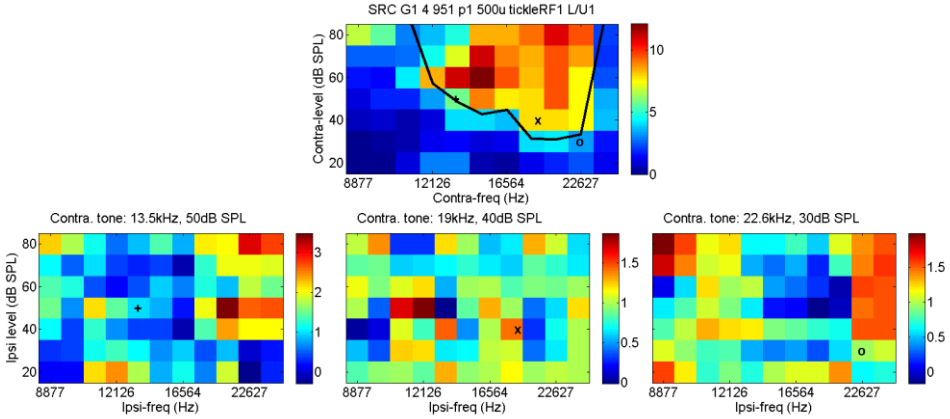

Fig 2.2.3. Example unit with "noisy" iFRAs. The top plot is the contralateral frequency tuning with the contralateral frequency response characterised as described in 2.2.2.9.1. Symbols indicate the chosen contralateral frequency/level combination to be used with the iFRAs. Bottom plots demonstrate the iFRAs stimulated using three different contralateral tone frequencies. Colours indicate the normalised firing rate relative to the contralateral tone (see text above).

Due to the difficulties in designing an adequate algorithm for classifying iFRAs a simple approach was taken. iFRAs were qualitatively assessed with the criteria that if response areas were very noisy (little order in spiking behaviour at neighbouring frequency/level combinations) no classification was made (Fig 2.2.3, bottom row, middle plot). If either a facilitatory or inhibitory region existed and the surrounding regions followed a similar trend (i.e. a decrease or increase in rate with increasing level, respectively) and the rate in this region changed by 20% or more, the iFRA was classified as EE or EI, respectively. If the surrounding region had a different interaction type (Fig 2.2.3, bottom row, far right plot) then the iFRA was classified as of mixed type. The classification for

the iFRA measured when the contralateral tone nearest to contralateral CF was used to make the overall unit classification, as this would correspond closest to the conventional unit classification.

2.2.2.9.3 Mean data

In order to analyse firing rates across the population, two approaches were taken. In the first subsets of the conditions in the iFRA were considered as ILD functions and a mean ILD for each CTF calculated. In the second instance normalised population receptive fields were created by resampling in the frequency dimension and averaging across CTF.

2.2.2.9.3.1 Population mean ILD functions

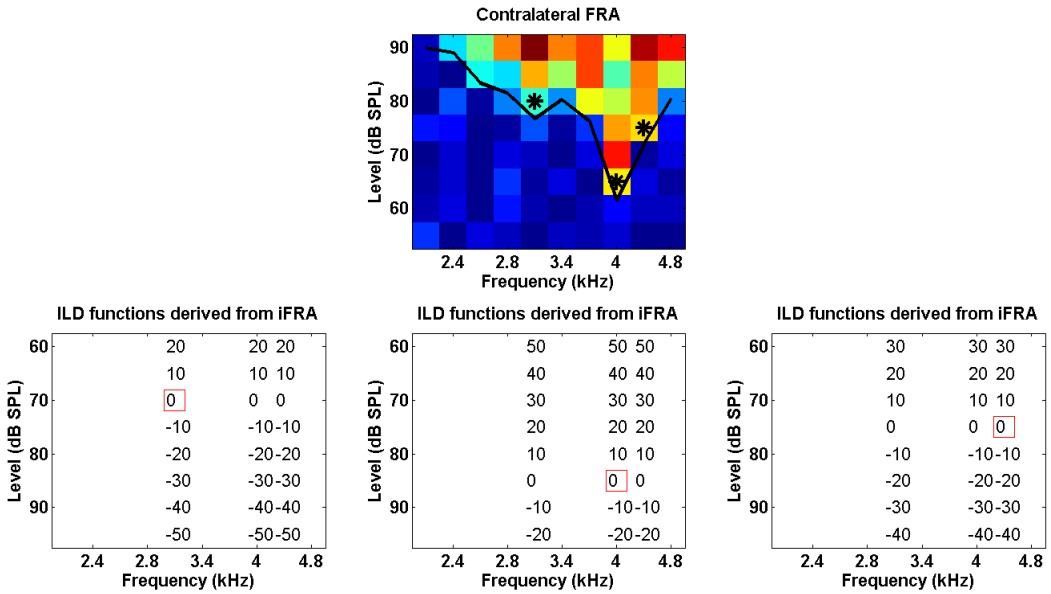

Fig. 2.2.4. Description of the calculation of ILD functions. Top – Contralateral FRA. Black line indicates the threshold criterion. Black stars indicate the contralateral tone level and frequencies presented when the iFRAs were collected. The frequencies and levels used for the iFRA were identical to those used for the contralateral FRA. Bottom – The ILD functions measured within each iFRA. Red squares indicate the level and frequency of the contralateral tone for each iFRA. Values indicate the size of the ILD at each frequency/level combination used.

Each of the iFRAs can be viewed as a set of ILD functions (see Fig 2.2.4). For each unit there were 3 iFRAs, one for each of the three contralateral tone frequencies presented (Fig 2.2.4, top). These same frequencies were also presented to the ipsilateral ear when collecting the iFRA (three for each contralateral frequency, Fig 2.2.4, bottom left to right). The main hypothesis would predict that the ipsilateral inhibition is strongest when frequencies are matched across the ears. Thus, we considered 3 ipsilateral frequencies matching the contralateral tones. This meant there were 9 possible contralateral/ ipsilateral frequency combinations (e.g. low contralateral vs. low ipsilateral, low contralateral vs. middle ipsilateral, low contralateral vs. high ipsilateral, etc.) The ILD functions were first normalised with respect to the overall minimum and maximum firing rate from all 9 ILD functions collected. The population mean, across units, for each of the 9 ILD functions was then calculated.

2.2.2.9.3.2 Mean iFRAs

Taking the mean iFRA posed two problems. The iFRAs for each unit sampled ipsilateral frequencies differently, and the contralateral tone frequencies used depended on the FRA tuning of each unit. Therefore each iFRA was up-sampled to yield the same minimum frequency step. Then for each iFRA the frequency axis was expressed in octaves from the contralateral tone frequency. Each iFRA was then normalised with respect to the contralateral tone alone firing rate and the mean iFRA calculated for each contralateral frequency.

2.2.3 Results

2.2.3.1 Main effects

Data was recorded from 96 units in total (30 SUs, 66 Mus). Each units binaural interaction type was classified (see section 2.2.2.9.2). 64/96 units could be classified as EI and 20/96 were classified as EE according to the iFRA recorded with a contralataral tone at CF (12 could not be readily classified). In addition to the basic binaural classification units could be classified according to the effect of contralateral tone frequency on iFRAs (80/96 units). Five main effects were observed, an example is provided of each.

2.2.3.1.1 Change in the binaural interaction type

19/80 units could be classified as demonstrating a change in binaural interaction type when varying the contralateral frequency/level combination (all of these were EI units). Fig 2.2.5 displays an example where the contralateral tone was presented below, at and above the contralateral characteristic frequency (cCF). When the contralateral tone was below contralateral CF (cCF) the iFRA demonstrated a facilitatory region corresponding roughly to the cFRA. When the contralateral tone was presented at cCF the iFRA was inhibitory centred on 16.6kHz, just below cCF. When the contralateral tone was presented above cCF the iFRA was again facilitatory. This facilitation is related to frequency/level combination for the contralateral tone being placed outside the cFRA and is discussed further later in this chapter.

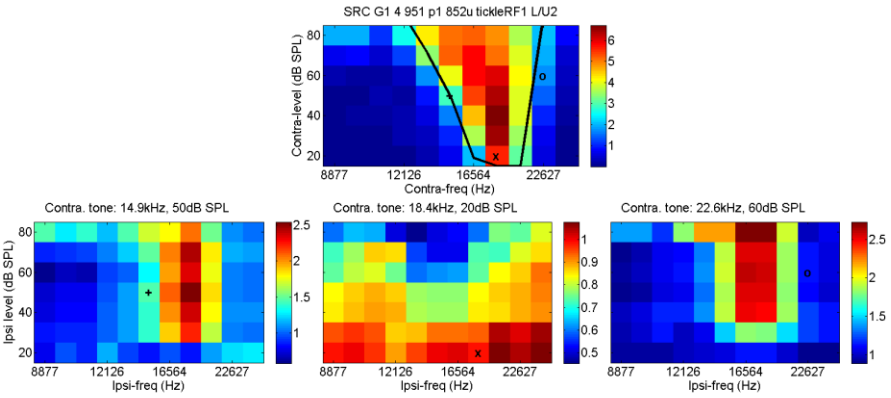

Fig 2.2.5. Example unit demonstrating a change in the binaural interaction type in the iFRA. Bottom left and right are classified as facilitatory responses whereas the middle response is inhibitory. Demonstrating a changing binaural interaction type with contralateral frequency/level combinations.

2.2.3.1.2 Change in the inhibitory best frequency (iBF)

25/80 units could be classified as demonstrating a change in the iBF with changes to the contralateral frequency (25/64 EI units, see Fig 2.2.6). For the middle contralateral tone condition an inhibitory response was observed with a deep and large inhibitory region being evoked just below the contralateral tone frequency. When the contralateral tone frequency was increased the deepest inhibitory region moved to 23 kHz, the frequency of the contralateral tone. Though it should be noted that a large inhibitory region was still present below 16 kHz, the frequency range previously maximally inhibited. For those units demonstrating a change the iBF changed with the contralateral tone frequency generally moving to neighbouring frequency regions. This was often, though not always, associated with a change in the ipsilateral characteristic frequency (iCF, see glossary of terms for distinction between BF and CF).

One other point to note is that at frequencies surrounding the inhibitory region there were slight facilitatory responses, this was true for almost all iFRAs measured regardless of the effect being seen. The lowest contralateral tone condition did not evoke a strong contralateral rate (Fig 2.2.6, top plot), potentially leading to the change in interaction type. Here there was a non-monotonic facilitatory region around 20.7kHz (peak activity at ~50dB SPL) with

one inhibitory sideband (>23kHz) and an area of very little change to rate (<16kHz).

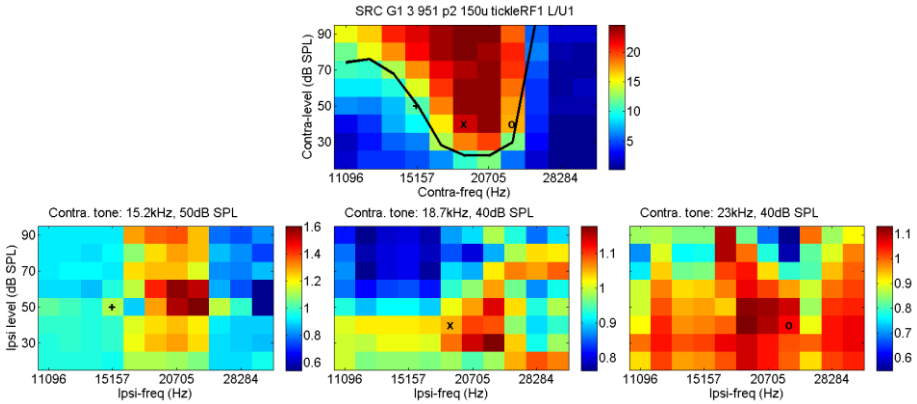

Fig 2.2.6. Example unit demonstrating a change in the iBF. The bottom middle panel demonstrates the iBF (largest inhibition) at ~12kHz at high ipsilateral levels (firing rate is <80% of the contralateral alone rate). In comparison the firing rate at high ipsilateral levels at ~23kHz is >90% of the contralateral alone rate. The bottom right panel demonstrates the iBF moved with contralateral frequency to 23kHz (where the firing rate is now <60% of the contralateral firing rate).

2.2.3.1.3 No change in the iCF

20/80 units did not demonstrate a change in the iCF, though there were often changes in the size of the iFRA (20/64 EI units). This sometimes corresponded with a change to the iBF. An example of this can be seen in Fig 2.2.7. For this unit one contralateral tone condition was selected at cCF with one above and one below. Despite large changes to the contralateral tone frequency presented the iCF for the iRFA remained at ~20 kHz for each condition. Despite no change in the iCF a change in the receptive field was often observed. For example the size of the iFRA was larger for the cCF tone when compared with the lower contralateral tone condition though the essential shape remained similar. The relative change between the cCF and higher contralateral tone condition was somewhat different with proportionally more inhibition being generated at the higher frequencies for the higher condition.

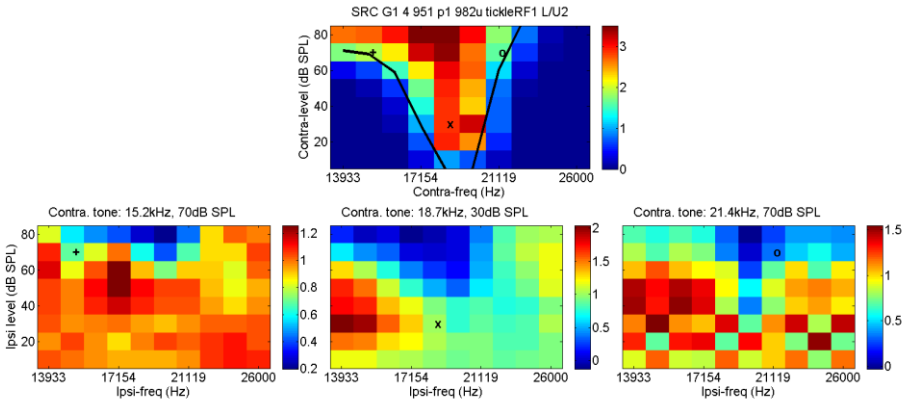

Fig 2.2.7. Example unit demonstrating no change in the iCF. For all bottom panels the iCF remained at ~20kHz. The size and shape of the iFRA varied but iCF remained stable.

2.2.3.1.4 No change to the facilitatory best frequency (fBF)

5/80 units demonstrated EE interactions in all contralateral conditions tested, none of these units demonstrated a change in fBF (5/20 EE units). An example of this can be seen in Fig 2.2.8.

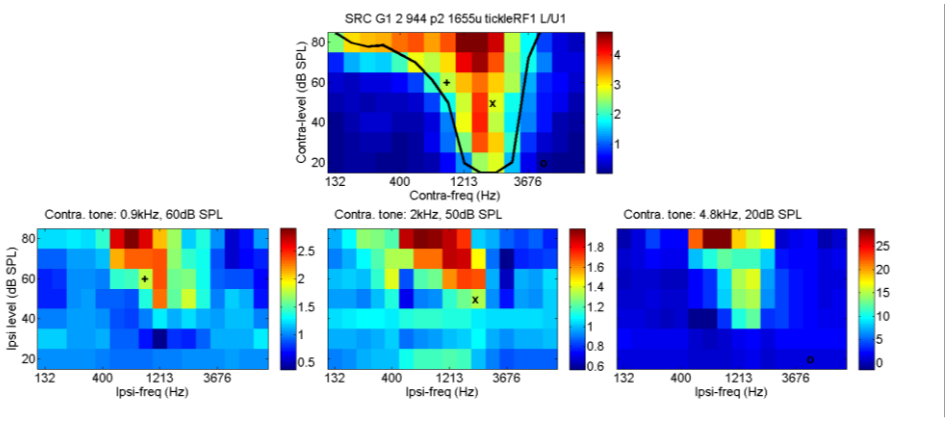

Fig 2.2.8. Example unit demonstrating no change in the fBF. For all bottom panels the iBF remained at ~696Hz. The size and shape of the iFRA varied but iCF remained stable.

For each contralateral tone condition the iFRAs measured were facilitatory in nature. Despite the contralateral tone frequency being varied by almost 2.5 octaves the fBF remained at ~700Hz, this was often reflected in the facilitatory CF also.

2.2.3.1.5 Mixed monotonic and non-monotonic FRAs

For a number of units (11/80) there was a mixture of rate/level responses across the two ears. Fig 2.2.9 is an example of this. The cFRA is regular with a monotonic increase in rate with increasing level for all frequencies the cell was sensitive to. However the iFRAs collected were all non-monotonic with respect to rate/level. For some units the contralateral ear produced a non-monotonic response with the iFRA being monotonic. The proportions in each classification are summarised in Table 2.2.1.

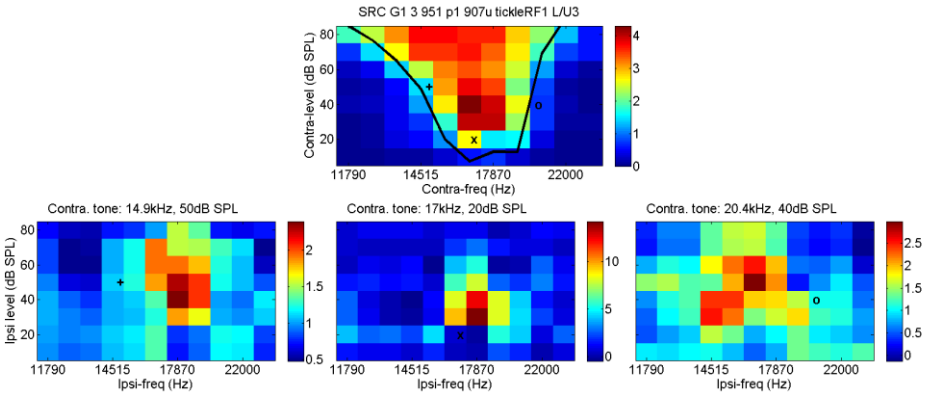

Fig 2.2.9. Example unit demonstrating mixed monotonic and non-monotonic FRA. Top panel demonstrates the cFRA was monotonic, bottom panels demonstrate non-monotonic iFRAs.

Binaural interaction classification	Number of units in class	Change in Interaction type	Change in iBF	No change in iCF	No change to fBF	Mixed monotonic-Non-monotonic	
EI	64	19	25	20	0	4	
EE	20	0	0	0	5	7	
no classification	12	0	0	0	0	0	

Table 2.2.1. Summary of the number of units within each classification.

2.2.3.2 Explaining changes in interaction type

As shown above some units would display inhibition for some contralateral tone frequencies and facilitation for others. This appeared to be related to the level of contralateral tone with respect to the contralateral frequency tuning of the

unit. More specifically, when the contralateral tone level was below threshold criterion (the black line on the cFRA plots, see 2.2.2.9.1) a facilitatory response could be evoked in an otherwise inhibitory unit. The iFRA classifications (see 2.2.2.9.1) were used to group data into inhibitory and facilitatory iFRA groups (96 and 106 iFRAs were in each group respectively). Data were then binned according to the level of the contralateral tone relative to the criterion threshold.

This grouping appeared to produce two distinct populations (Fig 2.2.10). The criterion threshold appeared to be related to the transition in iFRA classification from excitatory to inhibitory with increasing level. The majority of facilitatory iFRAs (88/106) were at or below 0 dB. The majority of inhibitory iFRAs (81/96) were at or above 0 dB. An independent-samples t-test was conducted to compare contralateral sound level relative to criterion threshold in the inhibitory and excitatory groups. A significant difference was found between the two groups ($p<0.05$).

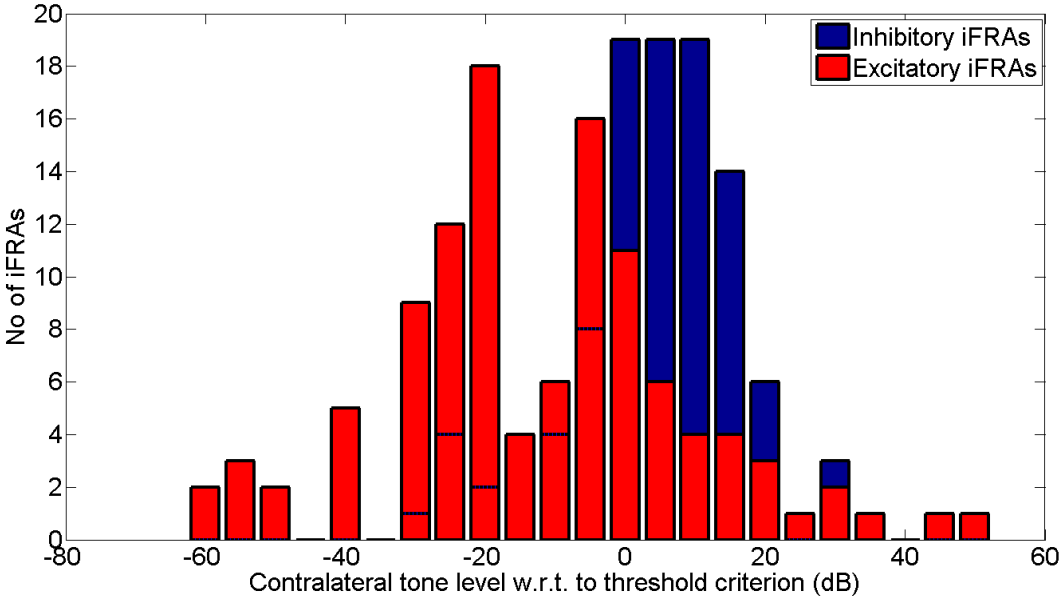

Fig 2.2.10. The signal level of the contralateral tone relative to the contralateral threshold at that frequency (dB). Bars indicate the number of iFRAs at each signal level for each iFRA classification. The legend indicates the classification of the iFRA. Horizontal blue dotted lines indicate the height of underlying inhibitory IFRA bars.

2.2.3.3 Mean ILD functions

To test whether CTF had an effect on the shape of the ILD functions across the population of units, the ILD functions from all inhibitory iFRAs were averaged based on the ILD value for each CTF/ITF combination (see 2.2.2.9.3.1 for details). For the low CTF condition the low ITF provided the greatest degree of inhibition (Fig 2.2.11). The middle ITF provided a similar, though slightly smaller reduction in rate. By contrast the high ITF yielded markedly less inhibition at the highest ILDs tested. For each CTF the matched frequency condition proved to be the most effective inhibitor of the contralateral alone response. In the low CTF condition the high ITF did not inhibit as much as the matched (low) ITF at all ILDs. By contrast in the high CTF condition is was only at large ILDs that the low ITF began to distinguish itself from the matched (high) condition. In addition to this the low ITF produced the least amount of inhibition for the middle CTF. Taken together this suggests an asymmetry in performance where low frequency ITF produce less inhibition compared to the high ITF.

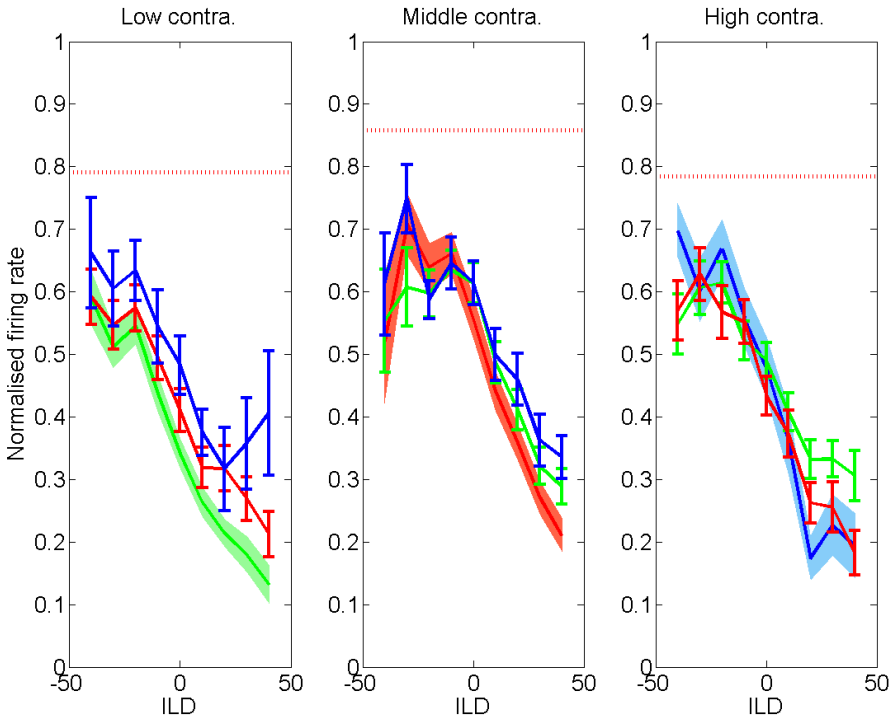

Fig 2.2.11. Mean ILD functions for different ipsilateral/contralateral frequency combinations. Panels indicate the contralateral tone frequency. Line colours indicate the ipsilateral tone frequency for low, middle and high contralateral tone frequencies (green, red and blue, respectively). Error bars are the standard error of means.

2.2.3.4 Mean iFRA

Mean iFRAs were calculated from the normalised unit iFRAs (see section 2.2.2.9.3.2). For the low CTF condition the most inhibited frequency, when averaged across level, was at 0.4 octaves below the middle CTF. For the middle CTF condition the most inhibited frequency was at middle CTF. For the high CTF condition the most inhibited frequency was at 0.3 octaves above the middle CTF. While the most inhibited frequency was modulated by the contralateral tone frequency there were a number of points worthy of note. The middle CTF condition demonstrated the deepest inhibition (i.e. lowest minima in terms of normalised rate) and the largest inhibitory region, possibly reflecting the importance of the contralateral CF on binaural interactions. Furthermore the most inhibited region (i.e. the frequency/level combination at which the deepest inhibition was found) was below the middle CTF, in terms of frequency. For the low and high CTF conditions the most inhibited region was roughly matched with the contralateral frequency. Also worthy of note is that for each CTF condition a small amount of facilitation was found at lower ipsilateral levels. The strength of this facilitation was usually greater at frequencies away from the most inhibited frequency.

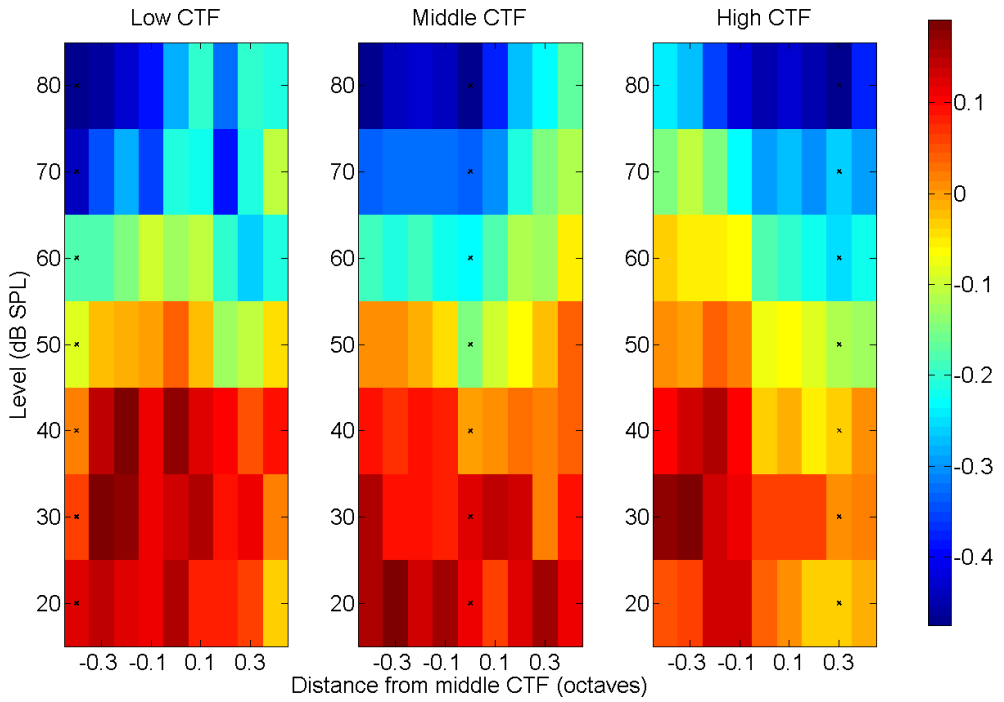

Fig 2.2.12. Mean iFRA for each CTF. Colours indicate the mean normalised rate. The dotted line indicates the iCF of the iFRA.

2.2.4 Discussion

Contralateral tone frequency can have a profound effect on the inhibitory frequency tuning of cortical cells. This effect could be observed in a number of different ways. Changing the CTF changes the iBF in 25/64 of EI cells, and this is often accompanied by a change in iCF. For those EI units (20/64) where no change in iCF was observed, changes in the iFRA were nevertheless evident. For the remaining EI cells (19/64) changes in the CTF produced changes in the interaction type above and below the cCF leading to a predominantly facilitatory interaction. The majority of these facilitatory iFRAs (79%) occurred when presenting the contralateral tone on the edge of or outside of the cFRA.

The result of these individual changes can be summarised by viewing the mean ILD functions and inhibitory iFRAs. For each binaural frequency combination, matching the two frequencies produced the largest modulations in spike rate. Increases in the binaural frequency difference decreased the degree of

modulation in spike rate, demonstrating the frequency specific nature of the inhibition observed. For the biggest binaural frequency differences (high CTF vs. low ITF and low CTF vs. high ITF) the high CTF condition produced the largest modulation in spike rate. The mean iFRAs also demonstrate that the size of the inhibitory region is large for the middle CTF. cCF was most frequently at or close to the middle CTF suggesting that binaural interactions close to CF would account for proportionally more of the broadband response than frequencies away from CF. This idea is difficult to validate as changes in frequency have been shown to change the optimal ABL/ILD combination for cortical cells (Kitzes, 2008).

Overall there is good evidence to suggest that binaural interactions are frequency specific. This experiment served the purpose of demonstrating the proof of concept and highlighting the key issues in collecting these data. One issue with the current study was the inherent noise found in the iFRAs this made it difficult to design an algorithm that could effectively define and classify iFRAs. Due to the large size of frequency/level combinations required to collect an iFRA few repeats were possible, generally a maximum of 10 were collected. By reducing the number of frequency/level combinations the number of repeats could be increased potentially reducing the noise encountered. Complexity in the iFRAs was also a barrier to a more thorough quantitative assessment. As seen in Fig 2.2.6 (bottom right panel) multiple inhibitory regions bounded by excitatory regions could exist, complicating automated classification. Another issue was that the CTFs used were individually catered to each unit. This meant that quantitative assessment of the effect of binaural frequency difference on the rate modulation of cells was not possible. These issues were addressed in the next study.

2.3 ILD functions and interaural frequency differences

2.3.1 Introduction

The previous experiment successfully demonstrated that binaural processing in cortical neurons depends on the particular combination of frequencies in the left and right ears, and that comparisons can be frequency specific. The design used in the previous chapter elucidated key issues in the collection of these data. Categorising the results was problematic due to the poorly defined iFRAs (see binaural classification in the previous chapter). As well as being noisy, iFRAs could take complex forms involving more than one inhibitory region (Change in the inhibitory best frequency (iBF). Another issue was that the iFRAs were specifically prescribed to fit the tuning of each cell which made quantitative assessment of frequency changes difficult to investigate across the entire population. This in turn did not allow for quantitative comparison of these data with other literature.

Several behavioural studies have measured the effect of varying the interaural frequency difference (IFD) of carriers on the sensitivity to envelope ITDs in humans (Nuetzel and Hafter, 1981; Saberi, 1998; Blanks *et al.*, 2007). In all of these studies, increasing the carrier IFD decreased the sensitivity to envelope ITDs of AM tones. Blanks et. al (2007) used sinusoidally amplitude modulated (SAM) pure tones (3 kHz) where modulator frequencies were either 128 or 129 Hz. This creates a binaural beat of 1 Hz where the stimuli are perceived as moving from one side of the head to the other. Listeners were asked whether the perceived location moved from left to right or vice versa for sounds with different IFDs. The listeners used were all highly trained in a similar task and performance was therefore deemed to reflect close-to-optimal performance (Blanks *et al.*, 2007). For all listeners, performance remained close to 100% correct with IFDs as large as 1 octave, beyond this performance declined. The

exact point at which performance began to fall varied with individual, but all listeners performed at above 70% correct at IFDs of 1.5 octaves. ITD envelope sensitivity appears to be relatively robust to changes in IFD, however, as performance was close to ceiling (100% correct) at smaller IFDs, detecting changes in performance over this range was not possible.

Francart and Wooters (2007) measured JNDs in ILDs for listeners using 1/3 octave noise bands while varying the interaural centre frequency difference of the noise bands. Listeners were presented sounds over headphones and were asked to indicate which direction the sound was perceived (left or right). Listeners were first trained at this task, though very little improvement was observed. As noise bands were used, it is more useful to discuss the amount of overlap between noises, rather than the interaural centre frequency difference. Noise stimuli were 1/3 octave wide, consequently the interaural centre frequency difference had to be greater than 1/6 octave to ensure no spectral overlap occurred between noise bands. For the highest frequencies used (4kHz), JNDs remained relatively stable while the spectrum overlapped, but a relatively large JND increase (from ~ 2dB to 4dB) was observed when the IFD became 1 octave (with a 2/3 octaves gap between overlapping spectrum). A small but significant increase in JND occurred at most centre frequencies with a 1/3 octave shift (almost no overlap), with very little change between 1/6 and 1/3 octaves. In addition to increases in the JNDs, the variability of these data also increased with IFD. Two points are particularly worthy of note here: Firstly, large changes in JND were observed once the interaural centre frequency was 1 octave (2/3 octaves in terms of spectral overlap). Second, while ILD JNDs only increased by between 1-3 dB, this still equates to large changes in angular position in open-field listening environments. For instance at 4 kHz, the highest frequency tested in this study, a 2 dB increase can be equivalent to approximately 10˚.

Unfortunately, the differences between these studies are too great to allow for quantitative comparison. Differences existed in the stimuli presented (noise bands vs. tones), procedure (method of constant stimuli vs. Levitt tracker) and proficiency of listeners (highly trained vs. relatively untrained). However, both these studies demonstrate that for IFDs of 1 octave, large changes in performance are observed in both envelope ITD and ILD across-frequency binaural comparisons.

Blanks et. al (2007) studied the effect of mismatching frequency on the synchronization to binaural beats in the IC in rabbits. SAM tones were utilised; envelopes had a 1 Hz difference between the two ears. The stimulus level and the envelope frequency were optimised to evoke the greatest synchronicity. Two conditions were tested: First, the contralateral carrier frequency was held constant while the ipsilateral carrier frequency varied and vice versa. The stable carrier frequency was selected to be either at BF or CF of a given cell while the frequency at the other ear was systematically varied in 0.25 octave steps. Data were collected in high frequency cells; CFs were between ~ 2 and 30 kHz. Electrophysiological recordings measured synchronisation (using the Synchronization Index, (Goldberg and Brown, 1969) of neuronal firing rate to the sound envelope. The authors found that increasing the IFD reduced the synchronisation of cells to sound stimuli. On average, the synchronisation rate fell rapidly at small IFDs (less than 0.5 octaves). Despite a large decrease in synchronisation at small IFDs significant synchronisation to the stimuli was found with frequency mismatches as large as 2-3 octaves. Synchronisation was greater at equivalent IFDs when the ipsilateral frequency was lower than the contralateral frequency. This demonstrated a frequency-related asymmetry in the ability to encode envelope ITD information when mismatching frequency.

This present study aimed to address the issues found in the previous section to allow measurement of the effective changes in IFD that lead to reductions in the ability to faithfully encode ILD information. By reducing the number of stimulus

conditions, it was possible to increase the number of repeats. As opposed to measuring entire iFRAs, ILD functions were measured using different frequency combinations (in the same manner as the mean data plot in the previous chapter).

2.3.2 Methods

2.3.2.1 Data collection

Data was collected from the auditory cortex of anaesthetised guinea pigs using the methods outlined in section 2.2.2.5. ILD functions were collected in the presence of a contralateral tone. Three contralateral tone frequencies were used: one at, one below and one above cCF. The frequency step size between each contralateral tone was equal on a logarithmic scale, this resulted in a frequency range of twice the frequency step size for each cell. Three frequency ranges were tested: 0.25, 0.5 and 1 octave steps. It was hoped that these changes would give a more robust picture of across-frequency integration of binaural cues, and facilitate a more quantitative approach to data analysis. In turn this data could be compared to the studies mentioned above to examine the similarity between the encoding of ILD and envelope ITD information.

In the previous chapter, the position of the contralateral tone within the frequency response area was shown to influence the type of binaural interaction found. Even for binaurally inhibitory cells, excitatory iFRAs were observed when the contralateral frequency and level were outside of the cFRA. In order to prevent this, an extra stage was added to the data collection phase to ensure contralateral tones levels evoked a strong response and to try and match the evoked responses at each contralateral frequency. Once three tone frequencies had been selected, contralateral rate-level functions were recorded at each frequency. Level values were selected to be at threshold from a visually

determined threshold for each respective frequency value. Again, stimuli were pure tones (100ms duration) with a cosine on/off ramp (6ms) presented monaurally to the contralateral ear. Frequency-level combinations were presented in a randomly interspersed order and 10 repetitions of each frequency-level combination were recorded. The rate-level responses generated were then used to inform the selection of the level of the contralateral tones for ILD functions. Levels chosen were selected to be within 10-20 dB of threshold and the levels at different frequencies were matched for rate.

ILD functions were then collected: Contralateral levels and frequencies had been selected based on rate-level plots, while ipsilateral levels were generated in Matlab using the acoustic calibration for that particular experiment. As before three contralateral frequencies were selected. At, below and above cCF, for the CF condition only one ILD function was measured (with 0 IFD) and above and below CF three ILD functions were measured at different IFDs. This meant in total seven ILD functions were recorded. Ipsilateral levels were set so that the ILD would be between + 20 dB and – 20 dB (at 5 dB steps). To further investigate the nature of the binaural interaction it was necessary to quantify the monaural contribution. A subset of the binaural stimuli was used, consisting of: three contralateral tones (one for each contralateral tone condition) and seven ipsilateral rate/level functions (one for each of the seven ILD functions). The ipsilateral rate/level functions only used three sound levels the same sound levels as those used to produce ILD of +20, 0 and -20 dB ILD.

2.3.2.2 Data analysis

Data was collected using BrainWare (developed by J. Schnupp, University of Oxford). Frequency-response areas, rate-level functions and ILD data were collected in separate recording runs (as outlined above) and hence a unique file was generated for each paradigm. These files were combined in Matlab for all

recordings on a single electrode at a single position and converted into one '.plx' file (for spike sorting using the Off-line Sorter v. 2.8.8; Plexon, Texas, USA). Spike sorting was then performed on this amalgamated data for each electrode. Merging data files assisted in ensuring all data for a given set of files was collected from the same single-unit. Data were then converted back to the original separate format (FRA, rate-level and ILD files) and exported to Matlab. Subsequently, data were converted into dB SPL from the stored attenuation values based on information contained in sound level calibration files recorded during the experiment.

2.3.2.2.1 Calculating Spike rates by temporally windowing over the stimulus duration

Only spikes within a selected time window were considered for further processing. For each individual unit, all spikes recorded were collated across conditions yielding a PSTH of all the spikes evoked regardless of condition. The PSTH onset response was manually selected as the start of the time window. The length of the time window was always 100ms.

2.3.2.2.2 Significance test

In order to ascertain whether one condition was statistically different from another a bootstrap statistical test was carried out on the spike count distributions. The spike counts for each presentation were compiled into a list of spike counts for each condition. To compare two conditions the two lists of spike counts were used. The two spike lists were sampled with replacement to form two new spike lists. From these, the difference between the means was calculated. This process was repeated 500 times, each time yielding the difference between the means. This gave the probability distribution function (PDF) for differences between the two means. From this, the probability of the

actual difference between the means occurring by chance, if the two spike lists came from a single distribution, was extracted. If the probability of the two spike lists coming from the same underlying distributions was less than 0.05, then they were considered to be statistically different from one another.

This test was applied to the ILD functions for each frequency condition; there was a total of nine ILDs (between -20 dB and +20dB) for each ILD function. Modulation of the spike rate of the contralateral-alone tone (monaural) when the ipsilateral tone was added (binaural) indicates a binaural interaction. A statistical difference between the means only needed to be present at one of 9 ILDs to demonstrate that a modulation of the contralateral-alone firing rate had occurred. Bonferroni correction was made to test the *P* value.

2.3.2.2.3 Monaural classification

Classifications of the monaural rate-level functions for each ear were conducted for each neuron recorded. Rate-level functions were normalised using the maximum and minimum of these data. Data were sub-divided into three categories: increasing (+), decreasing (-), and unresponsive (O). In order to determine whether the functions were increasing or decreasing, a straight line was fitted using a least squares approach. After examining the distribution of the gradients, a gradient of 0.125 was selected to classify the rate-level function. If $m \geq 0.125$, the unit was classified as (+), if $m \leq -0.125$, the classification was (-), and when $-0.125 < m < 0.125$, the classification was (O).

2.3.2.2.4 Binaural classification

For this data set I attempted to use a rigorous method for classifying the binaural interactions encountered. Individual ILD functions were used to make binaural classifications. The first step was to exclude ILD functions which

showed no significant modulation in firing rate, as described above. The next step involved normalising the ILD functions , using the equation:

$$\hat{fr} = \frac{(\frac{(\vec{fr}-C)}{max(\vec{fr}-C)})}{2} + 1 \qquad \textbf{eq. 2.3.1}$$

Where $\hat{fr}$ is the normalised firing rates of the ILD function, $\vec{fr}$ is the firing rates of the ILD function and C is the firing rate for the contralateral tone alone.

Principal component analysis (PCA) was applied to reduce the dimensionality of the ILD functions using the Matlab, function '*princomp*'. *Princomp* takes in an *n* by *p* matrix, with *n* observations and *p* variables, and returns a matrix of *p* by *p* PC coefficients (describing each PC in descending order of varance) and *n* by *p* matrix of scores (describing the weight of each PC describing each observation). When the amount of variance explained by *p* PCs was >90%, the number of PCs was deemed adequate to describe the data, hence only *p* PCs were used to describe the data in subsequent data analyses. Agglomerative clustering analysis (see next section) was then carried out on the selected PC weights to form two clusters: those that were binaurally inhibitory and those that were binaurally facilitatory. For each unit ILD functions were measured below, above and at CF of the cell. In addition above and below CF additional ILD functions were measured where an IFD was introduced. Each ILD function carried an classification as defined by the PCA and subsequent clustering. Only the ILD functions where contralateral and ipsilateral frequency were equal were used as part of the unit classification (one below, above and at CF). Where the majority of the ILD functions were classified as inhibitory the unit classification was given as binaurally inhibitory and if the majority of ILD functions were classified as facilitatory the unit was classified as binaurally facilitatory.

2.3.2.2.5 Clustering

Two types of cluster analysis were applied to the data to determine the number of clusters which would adequately define the types of binaural interactions occurring.

2.3.2.2.5.1 K-means clustering

Firstly, k-means clustering was used. This method aims to partition *m* observations into *k* clusters (where *k*<m). The algorithm is an iterative process which begins with *k* initial clusters taken by grouping data points at random. From this, a centroid value is calculated, i.e., the mean. A minimisation algorithm is employed in order to increase similarity of data points, based on the PC weights within each cluster. This can be expressed mathematically as follows:

$$argmin \sum_{i=1}^{k} \sum_{x_j \epsilon S_i} \| x_j - \mu_i \| \qquad \textbf{eq. 2.3.2}$$

Where $S = \{S_1, \ldots, S_k\}$ are the *k* sets of observations (k clusters) and μ_i is the mean of points in S_i.

2.3.2.2.5.2 Agglomerative clustering

The second cluster analysis employed was an agglomerative clustering algorithm. Before the ILD functions could be clustered it was necessary to create a dissimilarity matrix; a dissimilarity matrix describes a pairwise distinction between *m* objects. This measure is a one-dimensional value that indicates the strength of similarity/dissimilarity. The ILD functions were already reduced in dimensionality from 9 ILD points to *n* PCs (see description of princomp in 2.3.2.2.4). This meant each of the *m* observations had *n* PC scores

which described the ILD function. These scores could be used to describe an observation in *n* dimensional Euclidean space. The distance between two points expressed as the square root of the sum of squares describes the dissimilarity of two ILD functions:

$$\| M_{ij} \| = \sqrt{\sum_{k=1}^{n} (PC_{ik} - PC_{jk})^2} \qquad \textbf{eq. 2.3.3}$$

Where PC is an *m* by *n* matrix with the PC scores for *m* ILD functions described by *n* principal components and where *i* and *j* = 1 to m. k-means clustering always starts with k clusters and reduces the distances of the cluster members from the centroid iteratively. Agglomerative clustering begins with *m* clusters where the observation itself is the centroid of the cluster. A linkage function is used to measure the distance between two groups of objects, in this way the two closest clusters can then be clustered together. This incremental decrease in the number of clusters is continued until a specified number of clusters have been found. Ward's linkage was used to specify the distance between two clusters. Ward's linkage computes the increase in the error sum of squares (ESS) after fusing two clusters into a single cluster. The method seeks to minimise the increase in ESS as successive clusters are added. This was implemented in Matlab using the linkage function and the ward option. This function minimises the increase in ESS finding the pairing clusters with the smallest distance between their centroids. The centroid of cluster *r* is found as follows:

$$\bar{x}_r = \frac{1}{n_r} \sum_{i=1}^{n_r} x_{ri} \qquad \textbf{eq. 2.3.4}$$

Where n_r is the number of objects in cluster r. The between-cluster sum of squares is defined as the sum of the squares of the distances between the cluster centroids and the all cluster members. The equivalent distance is:

$$d(r,s) = \sqrt{n_r n_s \frac{(\| \bar{x}_r - \bar{x}_s \|_2)^2}{n_r + n_s}}$$ **eq. 2.3.5**

Where $\| \bar{x}_r - \bar{x}_s \|_2$ is the Euclidean distance, $\bar{x}_r$ and $\bar{x}_s$ are the centroids of the clusters *r* and *s*.

2.3.2.2.6 Normalising ILD functions when finding the mean ILD function

Before the population mean ILD function was calculated they were normalised relative to the firing rate when the contralateral tone was presented on its own (contralateral alone):

$$\hat{fr} = \frac{(\sum_{i=1}^{n} fr_i - C)}{fr_{max}}$$ **eq. 2.3.7**

Where $\hat{fr}$ is the normalised ILD function, *fr* is the original firing rate (and *i* is the ILD number and n the total number of ILDs tested), *C* is the contralateral tone alone firing rate and fr_{max} the maximum firing rate evoked across all *i*.

2.3.2.2.7 Frequency specificity index

The hypothesis to be tested is that the cell firing rate should change most when the ipsilateral and contralateral stimuli are matched in frequency and hence are frequency specific. For an inhibitory cell this would mean a zero IFD condition would result in more inhibition (a lower firing rate) than a >0 IFD condition. In order to assess the degree to which binaural interactions were frequency specific across the population, a contrast sensitivity metric was applied:

$$SI(f_i) = \frac{m(f_i) - n(f_i)}{m(f_i) + n(f_i)}$$ **eq. 2.3.8**

Where $m(f_i)$ and $n(f_i)$ are the total firing rate for each ILD functions in the matched (IFD = 0) and mismatched (IFD>0) conditions, respectively, and *i* is the contralateral frequency index. When the mismatched tone condition evokes a greater response than the matched condition, the SI value is negative, demonstrating greater inhibition across the ILD function.

2.3.3 Results

2.3.3.1 Individual unit examples

Fig 2.3.1 demonstrates a representative example of a single cortical cell. Three contralateral tone frequencies were selected within the FRA (see Fig 2.3.1, top row). For high and low contralateral frequencies, combinations of the three selected frequencies were presented to create frequency-matched and mismatched ILD functions (see Fig 2.3.1, bottom left and right panels). Most units that were binaurally inhibitory for both high and low contralateral frequencies exhibited the greatest amount of inhibition when the contralateral and ipsilateral frequencies were matched (thick lines in the bottom row of Fig 2.3.1 and 2.3.2). The condition with the largest separation between contralateral and ipsilateral frequencies (e.g. low contralateral frequency/high ipsilateral frequency and vice versa) produced the least inhibition of the firing rate, while smaller frequency differences (e.g. low contralateral frequency/middle ipsilateral frequency) were more similar to the matched-frequency condition. For some units, the firing rate at negative ILDs (low ipsilateral level) was comparable to or below the contralateral-alone rate (Fig 2.3.1), whereas for other units facilitation was consistently observed (Fig 2.3.2).

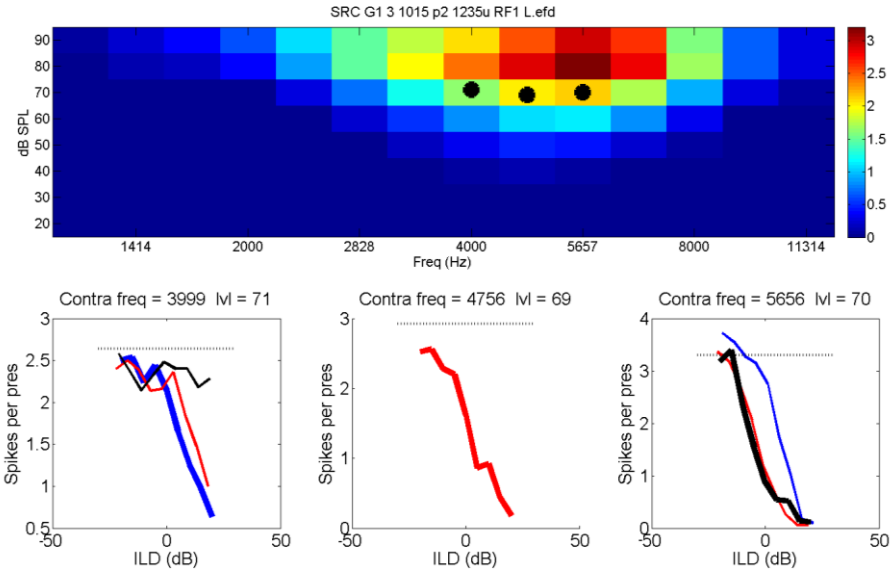

Fig 2.3.1. Example binaurally inhibitory unit. Top: FRA for contralateral tones. Black dots indicate contralateral tone frequency/level. Bottom: Each panel represents one contralateral frequency/level combination. Coloured lines indicate ipsilateral tone frequency: lowest, intermediate and highest are represented by blue, red and black respectively. Dashed black lines indicate contralateral alone rate.

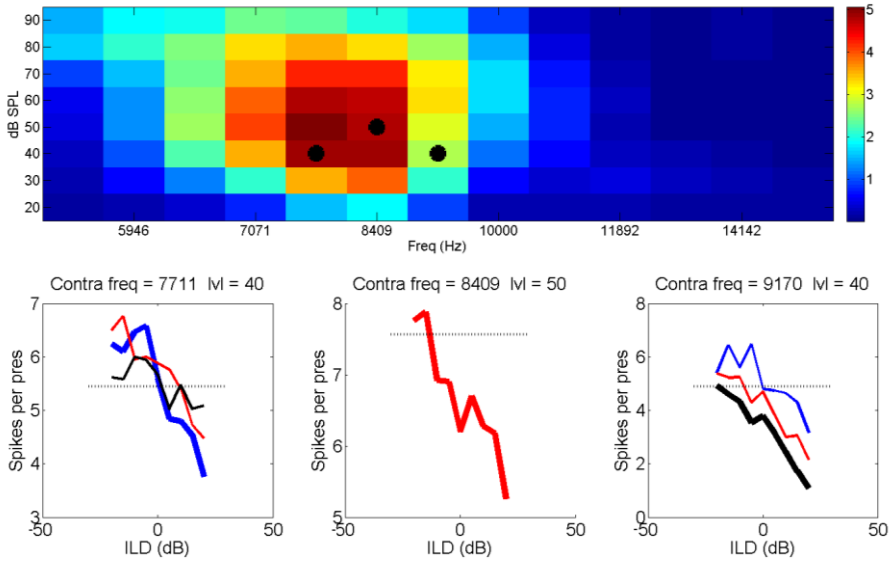

Fig 2.3.2. Example binaurally inhibitory unit. For explanation of the plot see Fig 2.3.1.

As with the previous experiment, some units demonstrated a change in binaural interaction type (Fig 2.3.3). This unit displays a binaurally inhibitory interaction for the matched cCF tone condition (bottom row, middle panel) but binaurally facilitatory interactions for most ipsilateral frequencies for the above-cCF and

below-cCF conditions (bottom row, left and right panels). As shown previously, these changes were most probably related to low firing rate caused by presenting the the CTF outside of the cFRA.

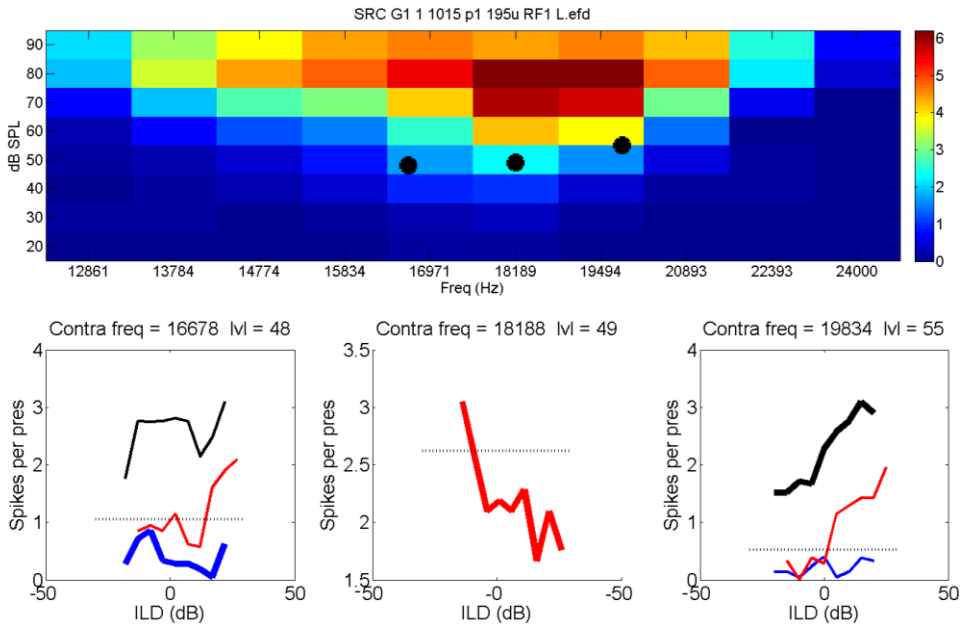

Fig 2.3.3. Example mixed unit. For explanation of the plot see Fig 2.3.1.

Some facilitatory units were also recorded. While inhibitory units seemed to display changes to the ILD function based on IFD changing the IFD did not appear to affect binaural classification of the units. For example Fig. 2.3.4 demonstrates no discernable change to the ILD functions with changes in IFD.

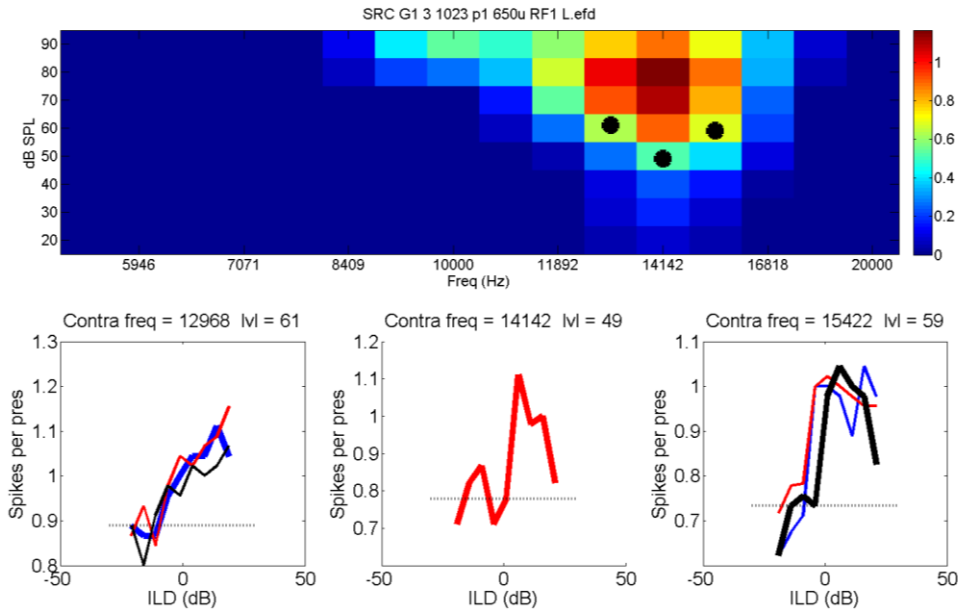

Fig 2.3.4. Example EE unit. For explanation of the plot see Fig. 2.3.1.

2.3.3.2 Classification of response types

Trends in the individual data pointed to frequency specificity in inhibitory cells and no frequency specificity in facilitatory cells. To quantify the degree of frequency specificity, binaural and monaural contributions were first classified and these classifications used to test frequency specificity. The classification scheme used in this thesis is different from more traditional classification schemes in a number of ways: Firstly, as the data collected aimed to look at across-frequency binaural interactions, a number of binaural classifications were made for each unit (with different frequency combinations), whereas in other schemes binaural interactions are only considered at the best frequency for each unit. Secondly, arbitrary criteria are often used to define the binaural interaction. In these data, a statistical test was applied to find the significant binaural interactions and then cluster analyses were applied. The classification of monaural and binaural responses has been studied many times in auditory cortex. Thus increasing the stimulus battery to comprehensively study this interaction was not a priority, however some efforts were made to record this data for comparison with existing studies.

2.3.3.2.1 Classification of binaural responses based on the ILD function

Before the classification began there were 223 candidate units (122 multi units and 101 single units). For each unit 7 ILD functions had been recorded therefore there were 1561 ILD functions in total to consider in the classification scheme. In order to classify units based on the ILD functions a principal component analysis (PCA) was carried out. When ILD data were plotted in PC space using the first two PCs, it was apparent that no clear clusters were present, though it appeared as if two poorly defined clusters may exist separated between -0.5>PC(1)<0 (Fig 2.3.5).

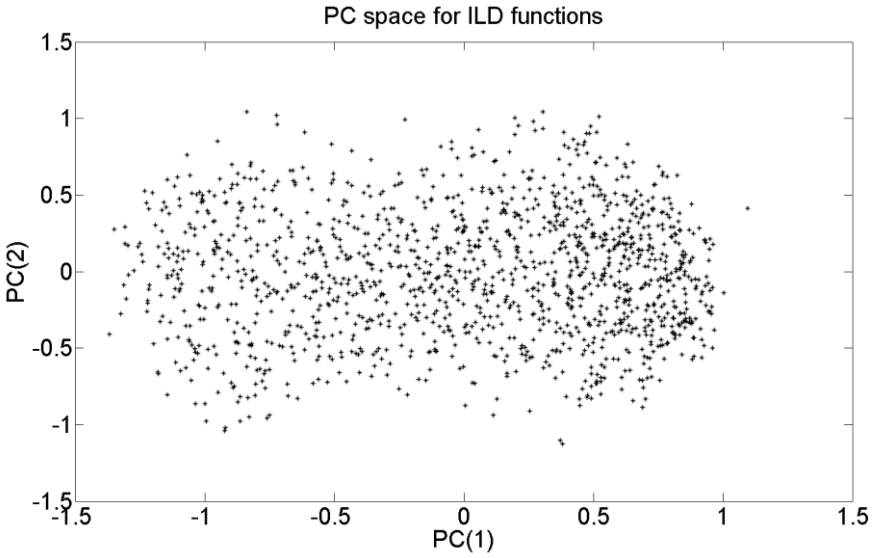

Fig 2.3.5. The first two components of PC space of ILD functions. No clear clusters were evident in the data.

If this represented two distinct clusters one might expect the distribution of PC(1) scores to be bimodal. Inspection of the PC(1) scores distribution reveals that while the distribution is skewed it is not bimodal but forms a skewed continuum (Fig 2.3.6).

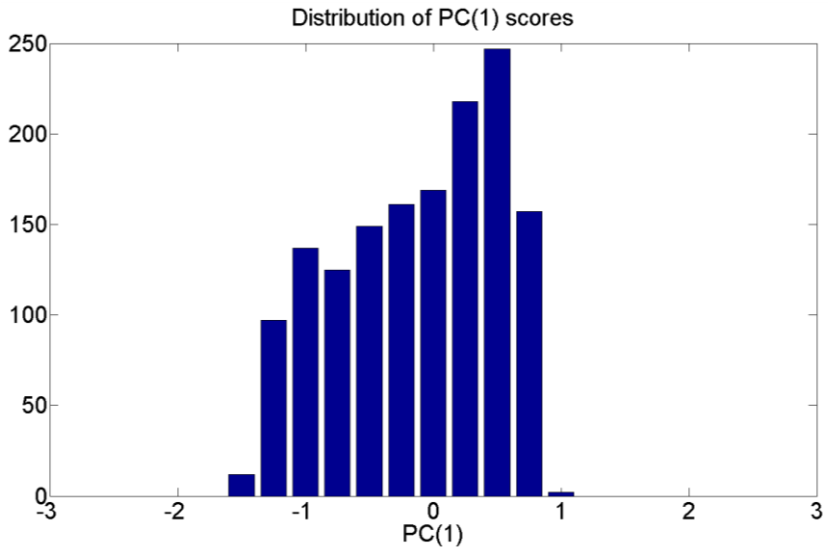

Fig 2.3.6. The number of ILD functions falling within each PC(1) score bin. This distribution is unimodal and skewed toward a PC(1) score of 0.5. A bimodal bin might indicate discrete clusters in PC space.

ILD functions were tested for significant binaural interactions by analysing the modulations in spike rate (see section 2.3.2.2.6). Only binaural interactions that were significant were included in further analysis (522/1461 ILD functions). Principal component analysis was then carried out only on these significant ILD functions. The first PC explained 79% of the variance in the ILD functions. The first 3 PCs explained over 90% of the variance (Fig 2.3.7).

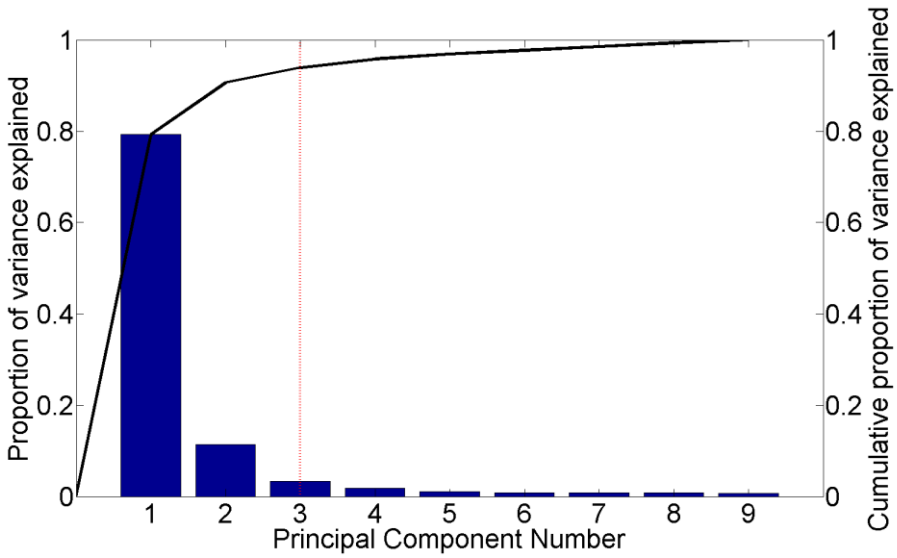

Fig 2.3.7. Proportion of variance explained by PCA of ILD functions with significant binaural interactions. Bars display the proportion of variance explained by each PC, thick black line indicates the cumulative proportion of variance explained by increasing PCs, dashed black line indicates the number of PCs required to explain >90% of the variance.

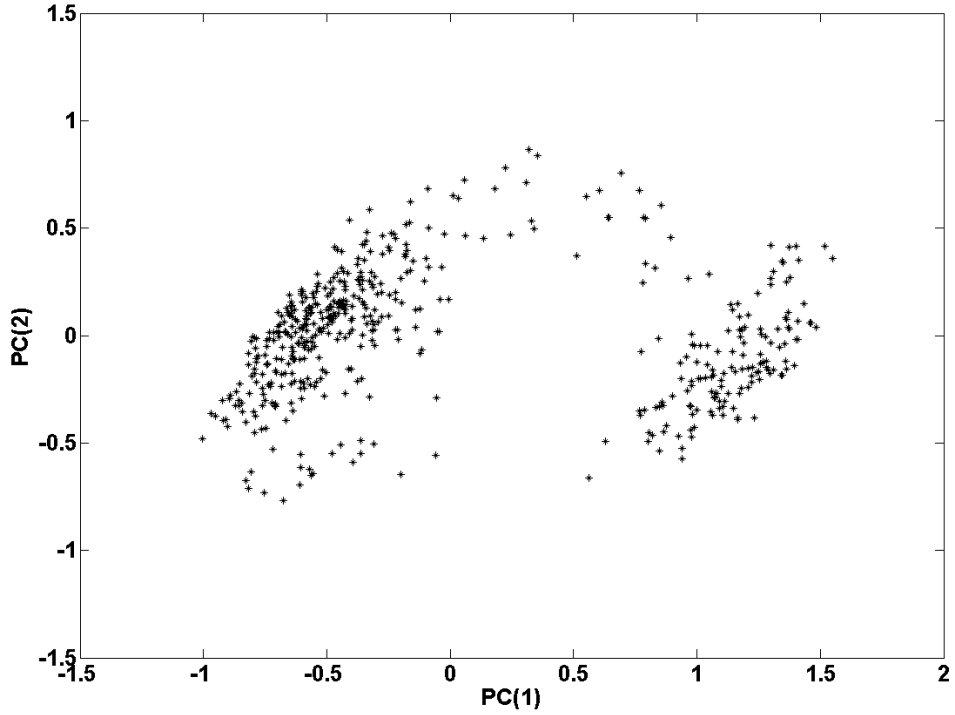

Fig 2.3.8. Scatter diagram of the first two PCs of the ILD functions after non-significant binaural interactions are removed. Two distinct clusters appear to be present.

This was selected as the threshold for the dimensionality of the data and all further processing was carried out on these 3 components. Fig 2.3.8 displays the data points of the first two principal components. In these two dimensions two fairly distinct clusters were evident in these data.

In order to objectively decide how many clusters these data should be partitioned into, two different clustering approaches were applied: Agglomerative and k-means clustering analyses (see section 2.3.2.2.5). The benefit of increasing the number of clusters can be assessed by comparing the total distance from centroids of each cluster, the smaller this distance the more similar the individual points within a cluster. As k-means clustering is an iterative process starting with randomly selected cluster centroids, the resultant cluster centroids were not always the same. Therefore the k-means clustering was carried out 50 times and the mean and standard deviation for each number of clusters displayed (Fig. 2.3.9).

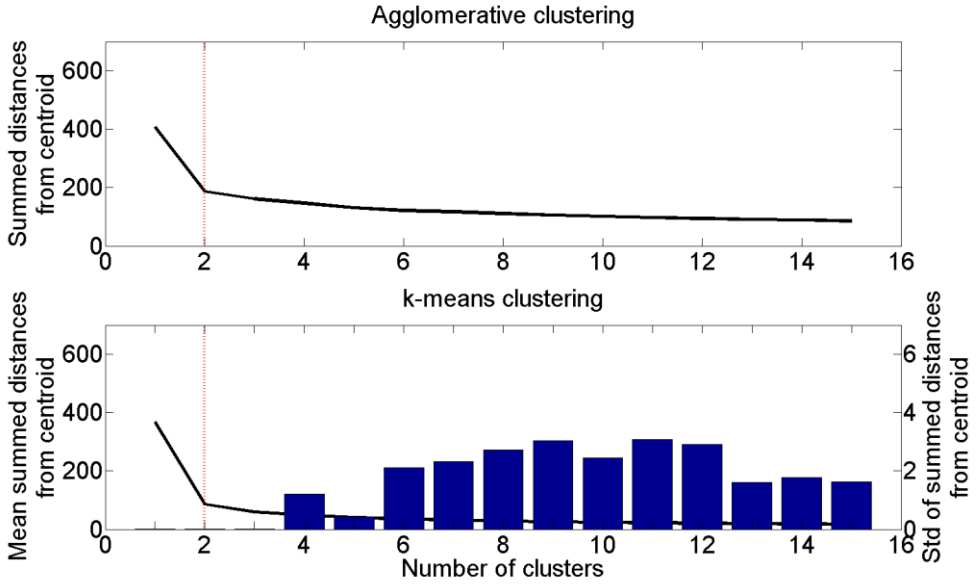

Fig 2.3.9. The effect of number of clusters on the summed distances from the centroid (see text). Top row, agglomerative clustering. Bottom row, k-means clustering. K-means clustering lead to a number of potential outcomes , therefore the mean summed distances from centroid was used (black line). Blue bars indicate the standard deviation for each number of clusters after 50 repeats.

As the number of clusters was increased from one to two, both techniques displayed a dramatic drop in the total distance from centroids. Subsequent increases in the number of clusters increased the distance from the centroid very little. As k-means clustering allowed for variation in the measured clusters it offered an opportunity to measure the stability or reproducibility of the clusters. The stability of the clusters, measured by the standard deviation, was robust at ≤3 clusters. After this, the clusters changed markedly with each repeat. As increasing the number of clusters above two had little effect in better-defining the data, this was deemed sufficient for partitioning the data.

Fig 2.3.10 displays the 2 selected clusters when using k-means clustering with the mean normalised ILD functions and standard errors for each of the clusters. Cluster 1 (black stars and line) inhibits the contralateral alone rate with increasing ipsilateral level. Whereas cluster 2 (red stars and line) facilitates the contralateral alone rate with increasing ipsilateral level. While two clusters

seemed adequate to characterise the ILD functions in this dataset, it is noteworthy that there were a number of previously documented subtypes.

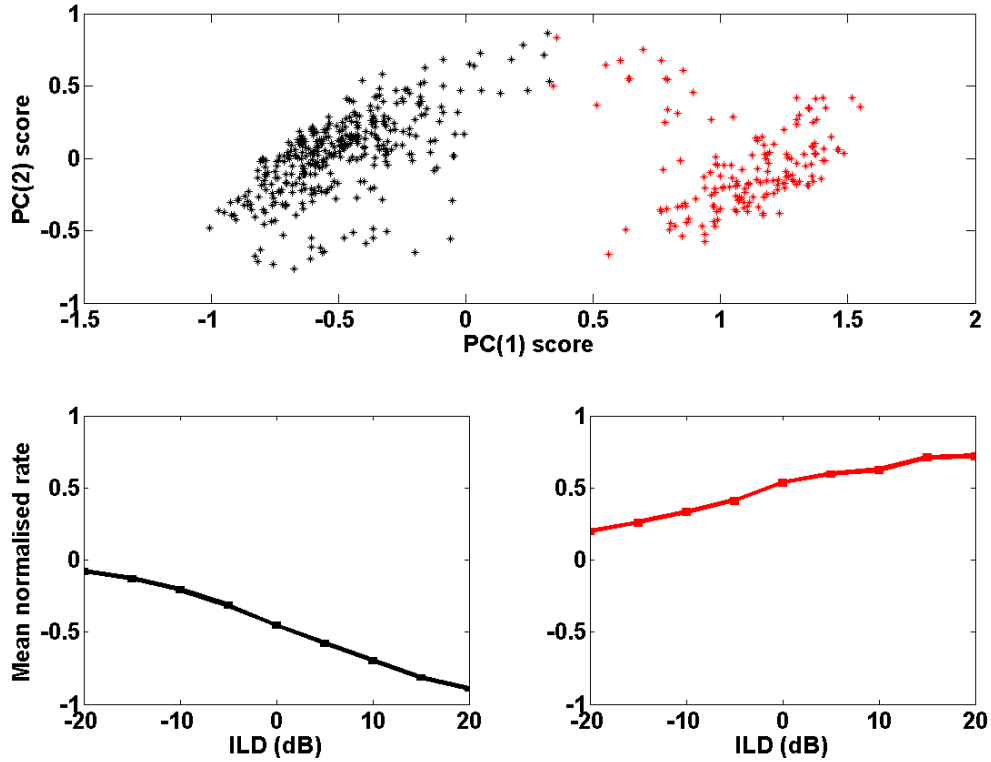

Fig 2.3.10. Two distinct clusters obtained using k-means clustering. Top panel, ILD functions plotted using the first two PCs. Colours distinguish the two clusters. Bottom panel, mean ILD functions for two clusters.

When the data was split into three clusters (using agglomerative clustering, Fig 2.3.11), the additional cluster (cluster 3) resembled the mixed facilitatory/inhibitory ILD functions described in previous studies (Semple and Kitzes, 1993; Nakamoto *et al.*, 2004). For successive increases in the number of clusters an increasing number of subtypes of the facilitatory and inhibitory clusters were found. For instance, in Fig 2.3.13 there were 5 clusters defined. Clusters 1, 3 and 5 were sub-classified from the previously described facilitatory cluster (Fig 2.3.10) and clusters 1 and 3 demonstrated facilitation beginning at different ILDs, one was instantly modulated by increasing ILD (cluster 5) the other demonstrated modulation of rate at higher ILDs. Similarly clusters 2 and 3

were subtypes of the inhibitory class and demonstrated changes at different ILDs.

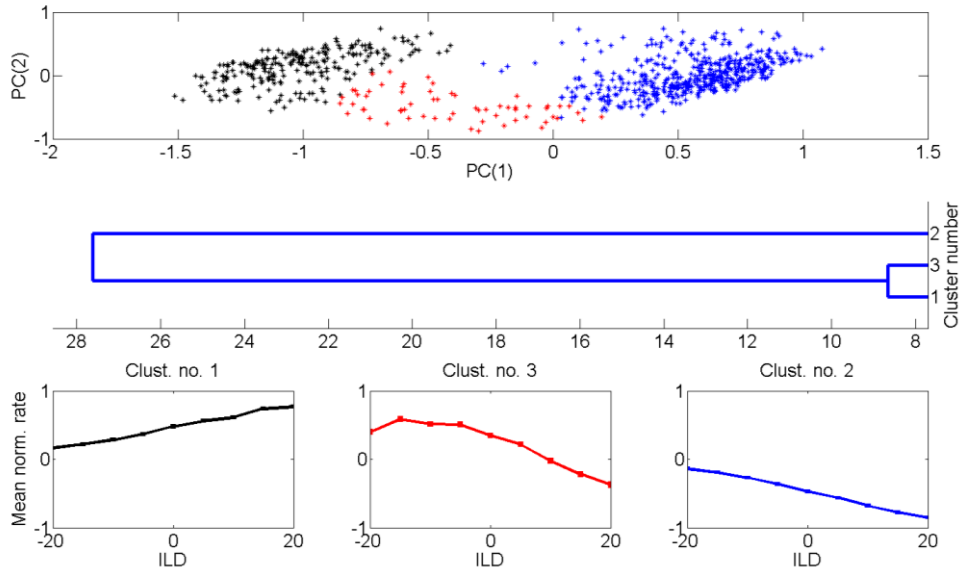

Fig 2.3.11. Three distinct clusters obtained using agglomerative clustering. Top panel, ILD functions plotted using the first two PCs, colours define each separate cluster. Middle panel, dendrogam demonstrating the relatedness of each cluster. Bottom panels, mean ILD function for each cluster.

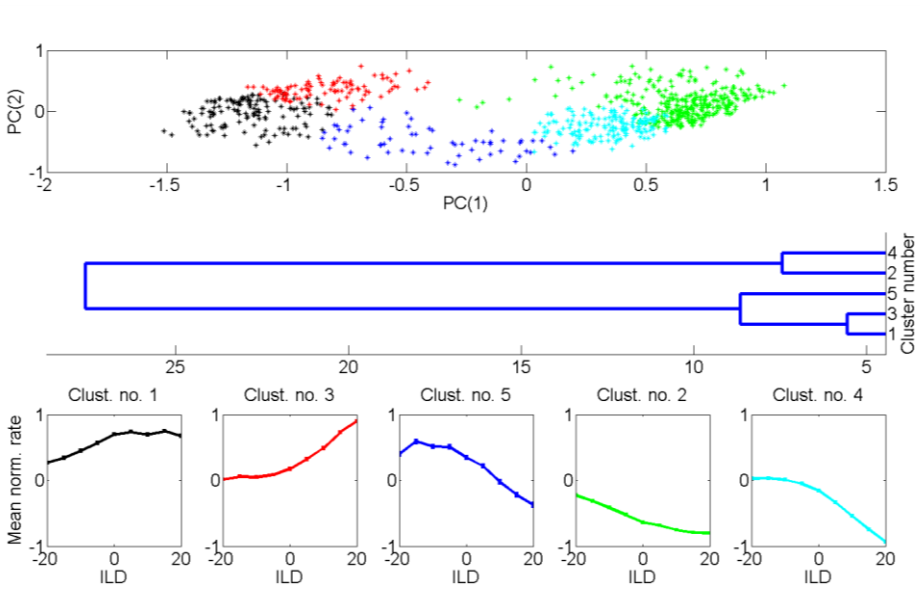

Fig 2.3.12. Five distinct clusters obtained using agglomerative clustering.

2.3.3.2.2 Monaural classification

Monaural rate level functions were also recorded, for each ILD function a subset of the corresponding monaural inputs were used (for more details see section 2.3.2.1). Using these a monaural classification was made for each ILD function (see section 2.3.2.2.3). Cortical cells exhibit a low spontaneous firing rate in auditory cortex. Due to the low number of spikes that can potentially be inhibited, recording monaural inhibition is generally not possible. For this reason, monaural classifications were restricted to being either excitatory (E) or unresponsive (O) to stimulation. As it was necessary to characterise cFRA this meant that only units with an excitatory contralateral response were recorded from (an ihbitory cFRA could not be characterised when little or no inhibition could be detected or if the cell was unresponsive). Ipsilateral rate-level functions could increase (monotonic, tuned to high SPLs), decrease (monotonic, tuned to low SPLs), a combination of an increase and a decrease (non-monotonic, tuned to middle SPLs), or exhibit little change in response. A straight line was fitted to each rate-level function. If the line gradient fell below a given threshold value, it was considered to be decreasing (E^-), if above it was increasing (E^+), and all other responses were considered unresponsive (O). As only one contralateral and three ipsilateral responses could be classified this meant that there were three possible monaural classifications (EE^-, EE^+ or EO).

A total of 522 ILD functions recorded using a mixture of binaural frequency combinations were found to show significant modulations of the ILD function ($p < 0.05$); 168/522 were classified as binaurally facilitatory, while 354/522 were classified as being binaurally inhibitory. Within the binaurally-facilitatory class 23, 125 and 20 ILD functions could be classified monaurally as EE^-, EE^+ and EO, respectively (Table 2.3.1). Within the binaurally-inhibitory class 60, 190 and 104 ILD functions can be classified monaurally as EE^-, EE^+, and EO, respectively (Table 2.3.1).

Binaural Classification	Monaural Classification			Total
	EE^-	EE^+	EO	
Facilitatory (F)	23	125	20	168
Inhibitory (I)	60	190	104	354

Table 2.3.1. Number of ILD functions within each binaural/monaural classification scheme

2.3.3.2.3 Classification using the monaural and binaural responses

It has previously been suggested that monaural classifications are not predictive of binaural classifications (Zhang *et al.*, 2004). To test this monaural and binaural classifications were considered together. As before each point in PC space reflects one ILD function, which could be recorded at, above or below cCF (Fig 2.3.13).

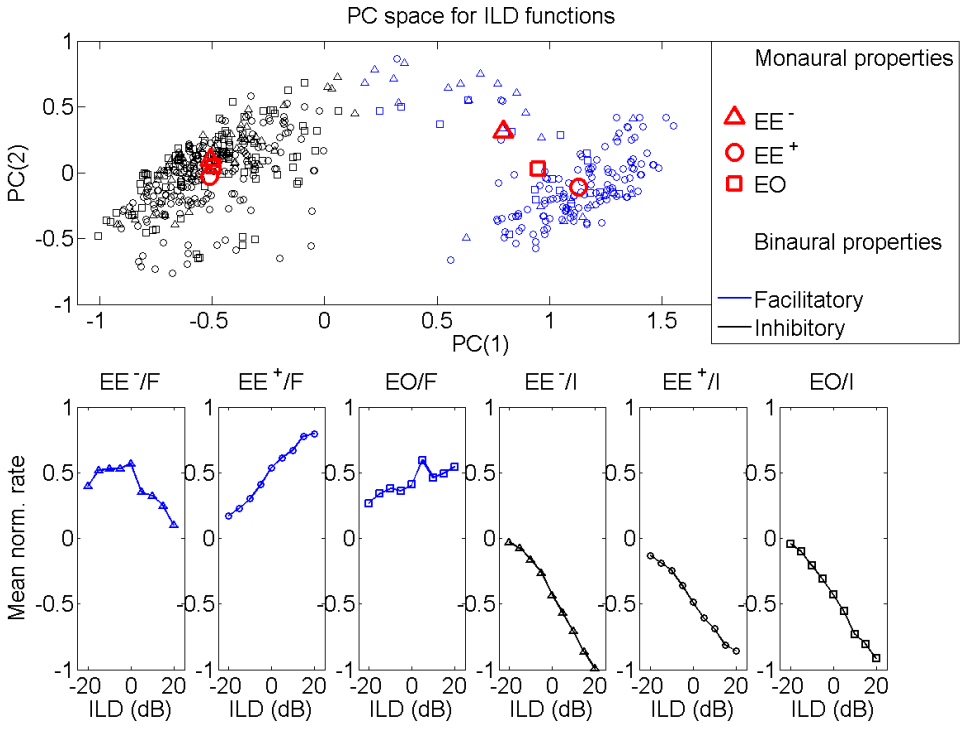

Fig 2.3.13. Monaural vs. binaural classifications. Top panel, PC space for all ILD functions. Symbols indicate the monaural classifications, colours indicated the binaural properties (see legend). Bottom panels, mean ILD functions for each monaural/binaural classification type. Titles refer to the classification type.

Monaural classifications were taken for both ears and for each corresponding ILD function. Of the two separate clusters defined (as above), one was binaurally-facilitatory (blue points in top panel), the other inhibitory (black points in top panel). Within each cluster subdivisions based on monaural classifications were defined. For conditions leading to a binaurally-facilitatory classification (/F) the characteristic shape of the ILD function differed depending on the monaural classification on (Fig 2.3.13, bottom leftmost panels): For those classified as EE^-/F the rate was initially facilitated above the contralateral-alone firing rate (mean normalised rate of 0) and increasing the ipsilateral level (or ILD) produced a reduction in rate. Increasing the ipsilateral level also decreased the monaural ipsilateral rate and hence there is agreement between the binaural and monaural response, although a number of ILD functions included in this class were essentially mixed facilitatory-inhibitory responses (see cluster 3 in Fig 2.3.11). This meant that for this class the decreases in rate at high ILDs were mainly due to this subpopulation existing within the class. The difference in ILD function shape is also evident by the differences observed in the centroids for each monaural class (Fig 2.3.13, top panel red symbols amongst the facilitatory cluster).

For the remaining classes, individual inspection of responses revealed that changes were due to average changes in the entire population and were not due to the grouping of responses with different forms. The mean ILD function of the EE^+/F class was slightly above the contralateral-alone rate at low ipsilateral levels and demonstrated relatively large increases in rate with increases in ipsilateral level. This is consistent with the ipsilateral monaural properties, increasing facilitation in conditions where the ipsilateral rate is high. The EO/F class demonstrated proportionally less facilitation at low ipsilateral levels than the EE^+/F class but greater facilitation at low ipsilateral levels.

For conditions leading to inhibitory binaural classifications (/I) mean ILD functions did not change as noticeably as for the facilitatory units. This is reflected by both the proximity of the centroids in PC space (Fig 2.3.15, top panel) and the similarity of the mean ILD functions (Fig 2.3.13, bottom panel). Despite ILD functions being more similar the monaural classifications did appear to sensibly influence ILD function. EE^- produced the deepest inhibition, EE^+ the shallowest inhibition and EO was in between the two.

Units were classified using the ILD classifications (see 2.3.2.2.4 for full details) at matched frequencies (one for each CTF). This meant for each unit there were 3 ILD classifications. A unit was classified as binaurally inhibitory (BI) if the majority of the ILD classifications were inhibitory (/I in Fig 2.3.13) and binaurally facilitatory (BF) if the majority of ILD classifications were facilitatory (/F in Fig 2.3.13). From the total population of cells 121 could be classified. Of these 88/121 were classified as being BI and 33/121 as binaurally BF. Of the 88 BI units 38 were single units (SU) and 50 were multi units (MU). Of the 33 BF units 17 were SUs and 16 were MUs (see Table 2.3.2).

Binaural classification	Single units	Multi-units
BI	38	50
BF	17	16

Table 2.3.2. Number of cells within each binaural classification scheme

2.3.3.3 The effect of IFD on ILD functions

ILD functions were collected while varying IFD in order to assess the frequency specificity of ILD integration. Fig 2.3.14 demonstrates the effect of IFD on the mean ILD functions for BI cells. The top row of Fig 2.3.15 displays the small step size condition. Three IFDs were presented in this condition 0, 0.125 and 0.25 octaves (top row; black, red and blue lines respectively). Very little difference was observed between the 0 and 0.125 octaves IFD conditions for either the low

(0.125 octaves below CF) or high (0.125 octaves above CF) CTF conditions. The 0.25 octave IFD condition deviated from the 0 IFD condition for the low CTF but not for the high CTF (top left and top right panels, respectively).

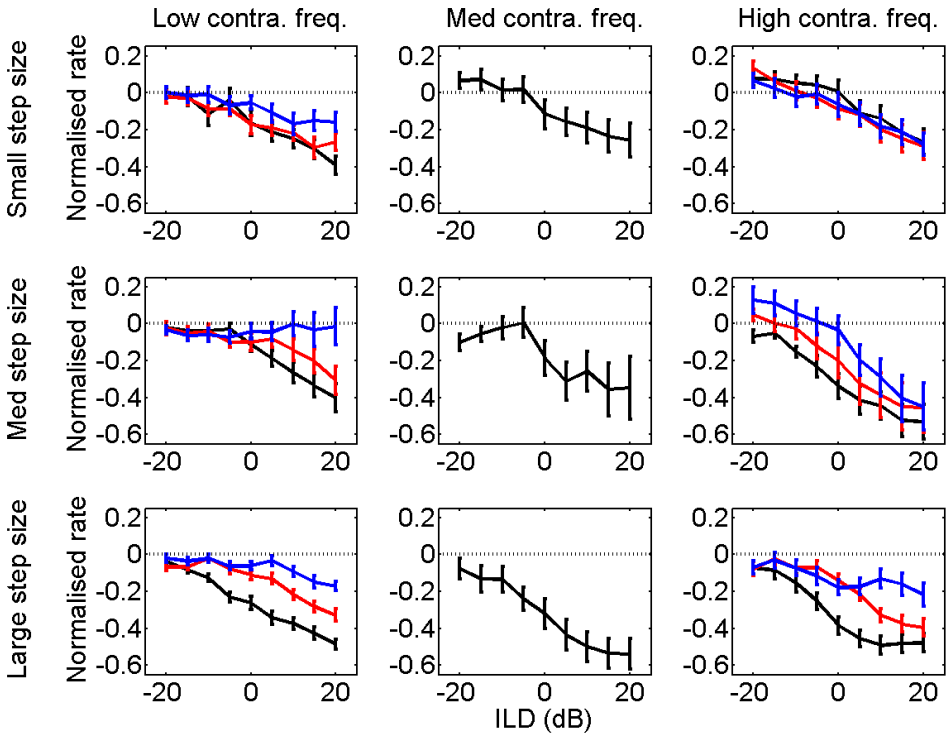

Fig 2.3.14. Mean normalised firing rate for three frequency steps sizes. Small, medium and large step sizes were used 0.125, 0.25 and 0.5 octaves from top to bottom row, respectively. From left to right panels indicated low, medium and high contralateral frequency, respectively. Line colours indicates the IFD where black, red and blue lines indicate no IFD, the smallest IFD (equal to the step size) and the largest IFD (twice the step size), respectively. When the contralateral tone was low (left panels) Error bars are the standard error of means. Dashed lines indicate the contralateral alone rate.

The middle row of Fig 2.3.14 displays the medium step size condition. Three IFDs were presented in this condition 0, 0. 25 and 0.5 octaves (black, red and blue lines respectively). Small differences were observed between the 0 and 0.25 octaves IFD conditions for both the low (0.25 octaves below CF) and high (0.25 octaves above CF) CTF conditions. These differences were only present at positive ILDs for the low CTF condition. The 0.5 octave IFD condition at the low CTF remained close to the contralateral alone rate at all ILDs whereas positive ILDs displayed inhibition of the rate below the contralateral alone rate (<0 normalised rate).

At large step sizes (bottom row of Fig 2.3.14) three IFDs were presented: 0, 0.5 and 1 octaves (black, red and blue lines respectively). Differences were observed between the 0 IFD and all other IFD conditions for both high and low CTF conditions. The smaller IFD (0.5 octaves) produced less inhibition of rate than the larger IFD condition (1 octave) at most ILDs. For the 1 octave condition and at low CTF it was not until the ILD was greater than 5dB that the function began to noticeably deviate from the contralateral alone rate.

For the low CTF conditions less inhibition of rate is observed at large ILDs (>0dB ILD) when an IFD is introduced (blue and red lines versus black lines). The smaller IFD condition (red lines) demonstrates the more inhibition than the larger IFD condition (blue lines) for each step size tested. This trend is also present in the high CTF condition, however, an asymmetry appears present when comparison is made across CTF conditions at equal step sizes. Overall for both high and low CTF conditions the matched frequency (0 IFD) condition generally produced the greatest inhibition of contralateral firing rate and increasing IFD progressively reduced the amount of inhibition.

2.3.3.4 Specificity Index

To demonstrate trends in the population on a unit-by-unit basis a frequency specificity index (SI) was used (see section 2.3.2.2.7). The aim of the metric was to compare the amount of inhibition of rate when an IFD was present versus when no IFD was present (Eqn. 2.3.8). A negative SI value indicates the firing rate, across ILD, was lower for the matched condition (0 IFD) compared with the mismatched condition (0>IFD) and vice versa. Two SI values were measured for each cell; one at the low CTF (SI_1) and one at the high CTF (SI_2). Fig 2.3.15 displays the result of this analysis. Symbols falling below the diagonal line indicate ILD functions were, on average, inhibited more for the matched condition than the mismatched condition and hence could be thought of as

demonstrating frequency specific inhibition ($SI_1+SI_2<0$). A significantly larger proportion (74/88) of BI units demonstrated frequency specific inhibition (sign test, $p< 0.01$). Of these 34 were single units and 40 were multi units, for both unit types a significant proportion displayed frequency specific inhibition (sign test, $p<0.01$ and $p<0.01$ respectively).

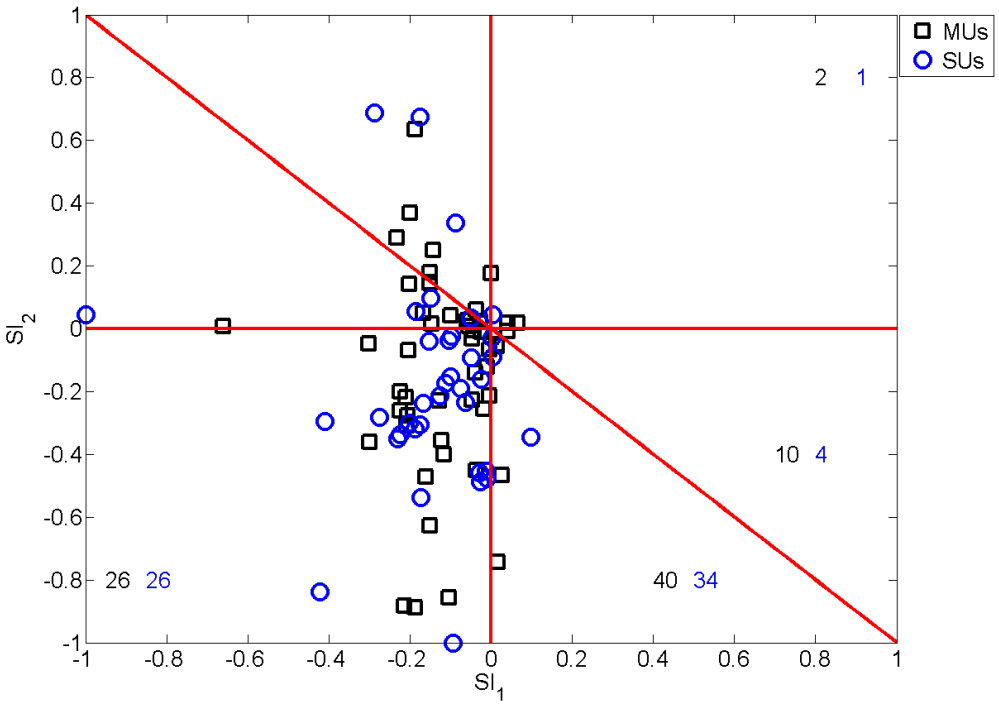

Fig 2.3.15. Specificity Index (SI) for BI units. Symbols indicate the SI value for each CTF condition: SI_1 and SI_2 for low and high, respectively. Blue circles indicate single units and black squares indicate multi units. Values indicate the number of units that fell below the diagonal (lower bottom right), above the diagonal (upper bottom right), in the bottom left quadrant (bottom left) and in the top right quadrant (top right).

Symbols located in the bottom left quadrant (Fig 2.3.15) indicate greater inhibition for the matched condition for both CTFs. Symbols located in the top right indicate stronger inhibition for the mismatched condition at both CTFs. A majority of BI units (52/90) had negative SI values at both CTFs (bottom left quadrant), whereas, very few units (3/90) demonstrated positive SI values at both CTFs (top right quadrant). This demonstrates that BI units were far more likely to demonstrate lower firing rates for both CTFs in the matched versus the mismatched condition. The mean ILD functions (Fig 2.3.14) suggests a slight

asymmetry in the data. Of the BI cells 77/88 gave negative SI_1 values whereas 60/90 gave negative SI_2 values. This suggested a frequency asymmetry, a paired two-sided sign test carried out on the SIs revealed this was not found a significant effect ($p<0.05$).

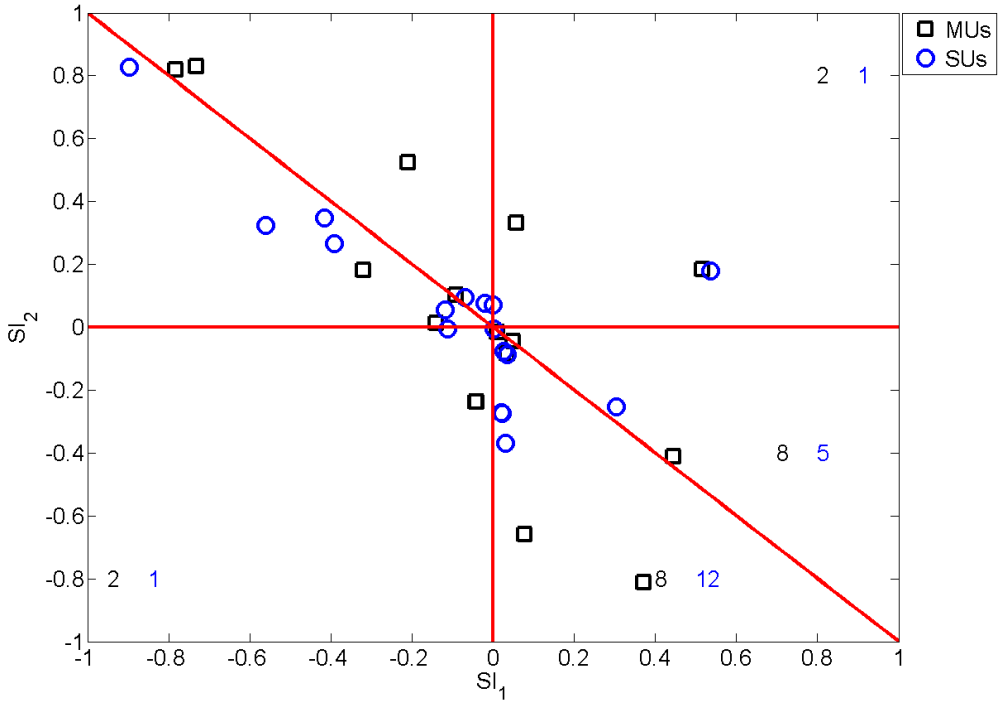

Fig 2.3.16. Specificity Index (SI) for BF units. Symbols indicate the SI value for each CTF condition: SI_1 and SI_2 for low and high, respectively. Blue circles indicate single units and black squares indicate multi units. Values indicate the number of units that fell below the diagonal (lower bottom right), above the diagonal (upper bottom right), in the bottom left quadrant (bottom left) and in the top right quadrant (top right).

BF units were also analysed using the SI index (Fig 2.3.16). The slight majority (20/33) fell beneath the diagonal line, though this was not significant (sign test, $p>0.05$). An equal amount of BF cells exhibited positive and negative SI values at both CTFs (3/33 for both). Overall there was no evidence that BF cells differentially integrated inputs based on frequency. BF units were also grouped according to IFD size no significant effect was found for any of the three step size (sign test, $p>0.05$ for all three step sizes).For this reason no further analysis was carried out on the BF population.

To test the effect of IFD on BI cells three step sizes were used: 0.25, 0.5 and 1 octaves. Frequency specific inhibition (points beneath the diagonal line) was found for just 15/23 when applying small IFDs (0.25 octaves), this was not significant (sign test, $p<0.05$).

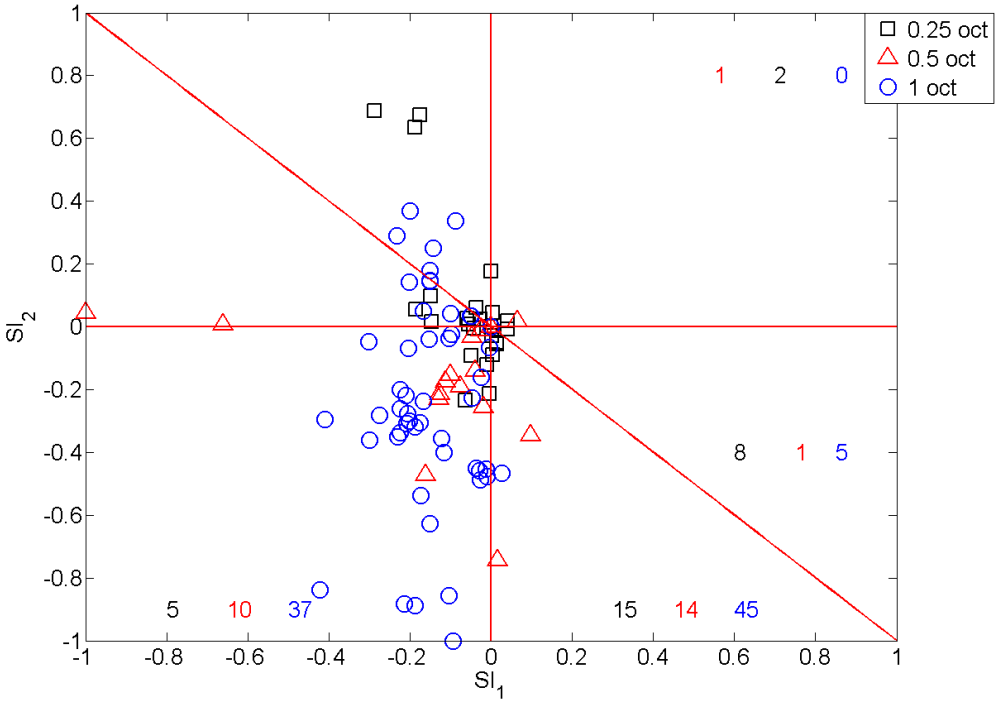

Fig 2.3.17. Specificity Index (SI) at different IFD sizes for EI cells. Symbols indicate the SI value for each CTF condition: SI_1 and SI_2 for low and high, respectively. IFD size is given by the figure legend. Values indicate the number of units that fell below the diagonal (lower bottom right), above the diagonal (upper bottom right), in the bottom left quadrant (bottom left) and in the top right quadrant (top right).

At larger IFDs (0.5 and 1 octaves) the significant majority of cells exhibited frequency specific inhibition (14/15 and 45/50, respectively). The proportion of cells showing frequency specific inhibition became progressively more significant with increasing IFD (sign test, $p > 0.2$, $p < 9^{-4}$ and $p < 10^{-9}$ for 0.25, 0.5 and 1 octave IFDs, respectively). At the smallest frequency step (0.25 octaves) only a small proportion of cells (5/23) demonstrated negative SI values at both CTFs (see Fig 4.3.17). At larger frequency separations the majority of cells, 10/15 and 37/50, demonstrated negative SI values at both CTFs, for steps of 0.5 and 1 octaves respectively. Again asymmetry was tested using a paired two-sided sign test carried out on the SIs at each IFD, this revealed no significant difference between the SI_1 and SI_2 conditions ($p>0.05$ for all IFDs).

2.3.3.5 Significance vs. frequency difference

The relative changes in proportion of units yielding significant inhibition aids quantification of the importance of the change in frequency difference. For instance if no units demonstrated significant inhibition at a given frequency difference it would be tempting to conclude that behaviourally the animal would not be able to lateralise the stimulus. Presenting the data in this format also allows for comparison of studies in the ITD literature (see discussion). BI units were grouped according to IFD and the proportion of units in each group showing significant inhibition calculated (Fig. 4.3.20).

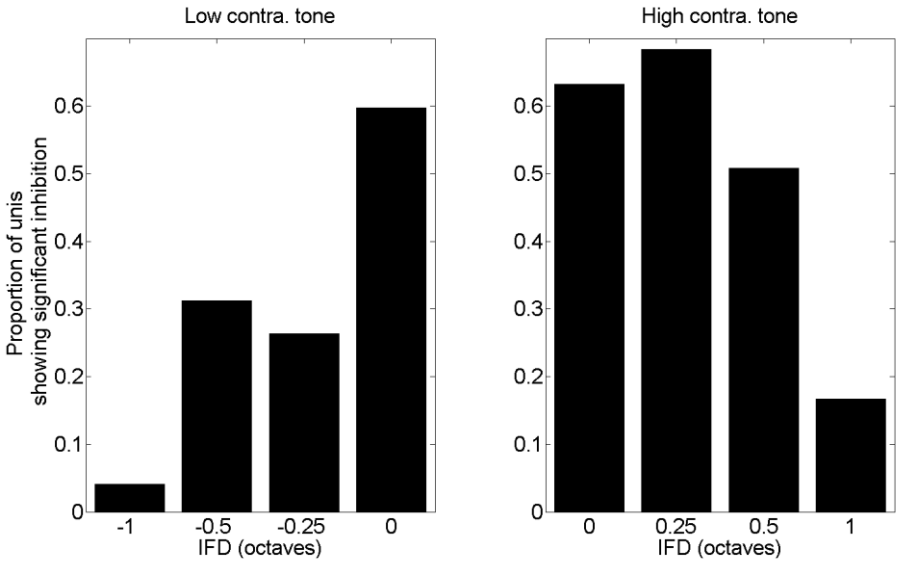

Fig 4.3.20. Proportion of units demonstrating significant inhibition within the ILD function at each frequency step.

For both high and low frequency conditions the proportion of units demonstrating a significant interaction with no frequency difference was almost equal. For the low contralateral tone condition even 0.25 octave steps significantly decreased ($\chi^2 = 13.578$, df = 4, $p<0.002$) the proportion of units evoking a significant interaction, when compared to the matched condition.

There was relatively little difference in the proportions calculated at -0.5 and -0.25 octaves. The difference between the 0 and 1 octave IFD groups was large, approximately 0.55, and highly significant ($\chi^2 = 42.12$, df = 4, $p<0.0000001$).

For the high contralateral condition no statistical difference was found between the 0 IFD group and the 0.25 or the 0.5 octave IFD groups ($\chi^2 = 0.06$, df = 4, $p>0.002$ and $\chi^2 = 2.067$, df = 4, $p>0.002$, respectively). The difference between the 0 and 1 octave IFD groups was, however, statistically significant ($\chi^2 = 28.58$, df = 4, $p<0.002$). Visually the proportions began to drop at 0.5 octaves, however, there was no significant difference between the 0.25 and 0.5 groups ($\chi^2 = 1.88$, df = 4, $p<0.002$). In summary, smaller IFDs were required to produce significant reductions in proportions for the low CTF condition. For the high CTF condition a 1 octave IFD was required to produce a significant reduction in proportions.

2.3.4 Discussion

The major findings of this experiment are as follows. When only significant modulations in firing rate were considered two distinct binaural classes were identified: facilitatory (BF) and inhibitory (BI). These two populations responded in different ways to IFDs. For those defined as BF there was no effect of introducing an IFD found. For those units defined as binaurally inhibitory a clear trend existed whereby sufficiently increasing the IFD decreased the depth of inhibition measured when compared to the 0 IFD condition. The size of the IFD seemed to be related to the amount of inhibition (Fig 2.3.14, the greater the IFD the greater the difference from the 0 IFD condition). In addition the IFD needed to be larger the 0.25 octaves to observe frequency specific inhibition. The low CTF condition appeared to produce greater inhibition for identical IFDs when compared to the high CTF condition. These differences were observed in the mean iFRA and ILD functions in the previous study (Figs 2.2.11-12, respectively) and also in the mean ILD functions

here (Fig 2.3.14). This difference was not statistically significant when tested using the SI.

2.3.4.1 Classification of the response types

The first point for discussion is the classification of response types. Response types can be classified based on monaural or binaural properties, or a combination of the two. Studies have often used both the monaural and binaural responses to classify cells (Irvine *et al.*, 1996; Rutkowski *et al.*, 2000; Zhang *et al.*, 2004). A plethora of categorisations have been proposed, the usefulness of which has been brought into question. This is for at least two reasons; the first is that binaural responses exist on a continuum and do not form well defined groups, so categorisation just refers to an arbitrary grouping and not a unique population (Campbell *et al.*, 2006). Secondly, it has been suggested that monaural responses do not directly predict binaural responses (Zhang *et al.*, 2004).

As PCA was used to define ILD functions it was possible to test whether the binaural responses formed well defined groups or were on a continuum. There are two main methods for collecting ILD functions; the EML-constant method and the ABL-constant method. ABL/ILD functions form a continuous distribution of response properties. When reduced into principal component space no clearly separable clusters are found (Campbell *et al.*, 2006). The study here collected ILD functions using the EML-constant method. In this reduced dimensionality a continuum of response type was also found. The two clusters found (inhibitory and excitatory) were imposed to allow the separation of binaural classification type by insisting on a statistically significant binaural interaction. Binaural response type existed on a continuum skewed toward an increased probability of encountering binaurally inhibitory responses in high frequency cells.

While the study here did not aim to rigorously quantify monaural responses, data were collected that allowed the categorisation of cells based on monaural and binaural responses. Particular monaural response combinations are more commonly associated with a specific binaural response than others, for instance cells with a EO monaural classification are more likely to be binaurally inhibitory than binaurally facilitatory (Reale and Kettner, 1986; Irvine *et al.*, 1996; Rutkowski *et al.*, 2000; Zhang *et al.*, 2004). The power of these predictions, however, is weak. For example one might expect that EE monaural responses should most frequently predict a facilitatory binaural response but studies in AI have demonstrated that this is not the case (Rutkowski *et al.*, 2000; Zhang *et al.*, 2004). In this study 88% of facilitatory units were classified as EE (either EE^+ or EE^-) and 71% of inhibitory units were classified as EE, therefore 12% and 29% were EO respectively. This demonstrates again that the monaural classification only indicates an increased likelihood of a particular binaural interaction type but is a poor predictor.

Combining subcategories used here to form just four groups EE/F, EE/I, EO/F and EE/I comparison can be made with a previous study in the guinea pig AI. This study found that 68% of EE units were inhibitory and 31% facilitatory (Rutkowski *et al.*, 2000). Similarly in this study 63% (250/398) of EE classifications were inhibitory and 37% (148/398) were facilitatory. Rutkowski et al. (2000) also found that 83% of EO units were inhibitory and 17% facilitatory. Again a similar percentage was found here where 84% (104/124) of the EO class were inhibitory and 16% (20/124) facilitatory. Overall there is reasonable agreement in percentages of each binaural class given particular monaural responses despite different data collection and classification methods. As in previous studies the monaural responses classified here were not directly predictive of the binaural response classification. Despite this there was evidence that the monaural responses were predictive of the amount of inhibition or facilitation evoked. For instance responses defined as EE^+/F would have greater facilitation than those defined as EO/F. This suggests that the

monaural responses do contribute toward the binaural response but the binaural response cannot be elucidated using monaural stimuli alone.

2.3.4.2 Across frequency binaural interactions

The frequency specificity of ILD computations was investigated. A contrast sensitivity index was applied (the SI) in order to measure the effect of introducing an IFD on the ILD function. The SI directly compared the evoked firing rates when the IFD was zero (the two frequencies were matched) and when an IFD was introduced (either 0.25, 0.5 or 1 octave IFD). This comparison was carried out twice for each cell, once when the CTF was below CF and once when it was above. A significant proportion of BI cells demonstrated frequency specific inhibition ($SI_1+SI_2<0$) whereas no significant effect was found for BF cells. One might expect that BF cells would demonstrate frequency specific facilitation ($SI_1+SI_2>0$) and it is not clear why integration of binaural cues was frequency specific in BI cells but not in BF cells.

Envelope ITD and ILD information appear to be represented in the same population of cells, i.e. high frequency neurons in auditory areas at and above the LSO (Tsuchitani and Boudreau, 1966; Creutzfeldt *et al.*, 1980; Batra *et al.*, 1993; Joris and Yin, 1995; Kuwada *et al.*, 1997; Palmer and Kuwada, 2005). Therefore one might expect to see a similar ability in computing across frequency comparisons in the coding of envelope ITD and ILD information. Blanks et. al (2007) measured the bandwidth over which significant synchronisation could be made on a neuron by neuron basis in IC. The median value for the ipsilateral ear was 1.25 octaves. A number of differences exist between this study and the one conducted here: different species were used, different measures were used, different auditory areas were tested and the testing methods varied. Though direct comparison cannot be made it is worthy of discussion. With decreasing ipsilateral frequency (for the high CTF) the proportion of units that produce significant inhibition falls drastically (and

significantly) at 1 octave. Thus for the population the majority of units yield significant inhibition over a 1 octave range. Interestingly for signals at high frequencies (1kHz and above) large difference in envelope ITD and ILD lateralisation performance are not observed until a 1 octave IFD has been introduced (Blanks *et al.*, 2007; Francart and Wouters, 2007).

2.3.4.3 Functional role of frequency specificity

The frequency specificity of BI cells and lack of it for BF cells potentially underlies a dichotomy in functional role. It is possible there is a functional gain to be had to effectively processing ILD information in separate frequency channels. One potential reason is that without frequency specificity the ILDs would be averaged across frequency first. The proportional change in ILD when varying spatial position is not equal at every frequency, some frequency ranges are more sensitive than others (Feddersen et al., 1957). If ILDs not frequency specific the auditory system could lose information over these more sensitive frequency ranges. Another potential reason could be that one population of cells is used for localising high frequency sounds (BI) and one could be used for detecting high frequency cells (BF). At present the reason for this dichotomy is unclear.

3 Bibliography

Abel, S. M., Giguère, C., Consoli, A., and Papsin, B. C. (**2000**). "The effect of aging on horizontal plane sound localization," The Journal of the Acoustical Society of America **108**, 743.

Abel, S. M., and Hay, V. H. (**1996**). "Sound Localization the Interaction of Aging, Hearing Loss and Hearing Protection," Scandinavian Audiology **25**, 3-12.

Aitkin, L. M., and Martin, R. L. (**1987**). "The representation of stimulus azimuth by high best-frequency azimuth-selective neurons in the central nucleus of the inferior colliculus of the cat," Journal of neurophysiology **57**, 1185.

Altshuler, M. W., and Comalli, P. E. (**1975**). "Effect of stimulus intensity and frequency on median horizontal plane sound localization," Journal of Auditory Research.

Anderson, H., and Wedenberg, E. (**1965**). "A new method for hearing tests in the guinea pig," Acta Oto-laryngologica **60**, 375-393.

Batra, R., Kuwada, S., and Stanford, T. R. (**1993**). "High-frequency neurons in the inferior colliculus that are sensitive to interaural delays of amplitude-modulated tones: evidence for dual binaural influences," Journal of neurophysiology **70**, 64.

Beyerl, B. D. (**1978**). "Afferent projections to the central nucleus of the inferior colliculus in the rat," Brain Research **145**, 209-223.

Blanks, D. A., Roberts, J. M., Buss, E., Hall, J. W., and Fitzpatrick, D. C. (**2007**). "Neural and behavioral sensitivity to interaural time differences using amplitude modulated tones with mismatched carrier frequencies," JARO-Journal of the Association for Research in Otolaryngology **8**, 393-408.

Blauert, J. (**1969**). "Sound localization in the median plane," Acustica **22**, 205–213.

Blodgett, H. C., Jeffress, L. A., and Taylor, R. W. (**1958**). "Relation of masked threshold to signal-duration for various interaural phase-combinations," The American Journal of Psychology **71**, 283-290.

Bolia, R. S., D'Angelo, W. R., Mishler, P. J., and Morris, L. J. (**2001**). "Effects of hearing protectors on auditory localization in azimuth and elevation," Human Factors: The Journal of the Human Factors and Ergonomics Society **43**, 122.

Boudreau, J. C., and Tsuchitani, C. (**1968**). "Binaural interaction in the cat superior olive S segment," Journal of neurophysiology **31**, 442.

Boudreau, J. C., and Tsuchitani, C. (**1970**). "Cat superior olive S-segment cell discharge to tonal stimulation," Contributions to sensory physiology **4**, 143.

Bourbon, W. T., Evans, T. R., and Deatherage, B. H. (**1968**). "Effects of Intensity on “Critical Bands” for Tonal Stimuli as Determined by Band Limiting," The Journal of the Acoustical Society of America **43**, 56.

Bronkhorst, A. W. (**1995**). "Localization of real and virtual sound sources," Journal of the Acoustical Society of America.

Brosch, M., Selezneva, E., and Scheich, H. (**2005**). "Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys," The Journal of neuroscience **25**, 6797.

Brown, C., Beecher, M., Moody, D., and Stebbins, W. (**1980**). "Localization of noise bands by Old World monkeys," The Journal of the Acoustical Society of America **68**, 127.

Brugge, J. F., Reale, R. A., and Hind, J. E. (**1996**). "The structure of spatial receptive fields of neurons in primary auditory cortex of the cat," The Journal of neuroscience **16**, 4420.

Brunso Bechtold, J., Thompson, G., and Masterton, R. (**1981**). "HRP study of the organization of auditory afferents ascending to central nucleus of inferior colliculus in cat," The Journal of Comparative Neurology **197**, 705-722.

Bullock, D., Palmer, A., and Rees, A. (**1988**). "Compact and easy-to-use tungsten-in-glass microelectrode manufacturing workstation," Medical and Biological Engineering and Computing **26**, 669-672.

Burdick, C. K. (**1979**). "The effect of behavioral paradigm on auditory discrimination learning: A literature review," Journal of Auditory Research.

Butler, R. A. (**1986**). "The bandwidth effect on monaural and binaural localization," Hearing research **21**, 67-73.

Butler, R. A., and Musicant, A. D. (**1993**). "Binaural localization: influence of stimulus frequency and the linkage to covert peak areas," Hearing research **67**, 220-229.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (**1986**). "Detection of simple and complex tones in fixed and random conditions," The Journal of the Acoustical Society of America **79**, S48.

Caird, D., and Klinke, R. (**1983**). "Processing of binaural stimuli by cat superior olivary complex neurons," Experimental Brain Research **52**, 385-399.

Calford, M. B., and Aitkin, L. M. (**1983**). "Ascending projections to the medial geniculate body of the cat: evidence for multiple, parallel auditory pathways through thalamus," The Journal of neuroscience **3**, 2365.

Campbell, R. A. A., King, A. J., Nodal, F. R., Schnupp, J. W. H., Carlile, S., and Doubell, T. P. (**2008**). "Virtual adult ears reveal the roles of acoustical factors and experience in auditory space map development," The Journal of neuroscience **28**, 11557.

Campbell, R. A. A., Schnupp, J. W. H., Shial, A., and King, A. J. (**2006**). "Binaural-level functions in ferret auditory cortex: evidence for a continuous distribution of response properties," Journal of neurophysiology **95**, 3742.

Cant, N., and Morest, D. (**1977**). "Small cells of the anterior subdivision of the anteroventral cochlear nucleus (AVCN) of the cat," Anat. Rec **187**, 544.

Carlile, S. (**1990**). "The auditory periphery of the ferret. I: Directional response properties and the pattern of interaural level differences," The Journal of the Acoustical Society of America **88**, 2180.

Clarey, J. C., Barone, P., Irons, W. A., Samson, F. K., and Imig, T. J. (**1995**). "Comparison of noise and tone azimuth tuning of neurons in cat primary

auditory cortex and medical geniculate body," Journal of neurophysiology **74**, 961.

Comalli, P., and Altshuler, M. W. (**1976**). "Effect of stimulus intensity, frequency, and unilateral hearing loss on sound localization," J Auditory Res **16**, 275-279.

Comalli, P., and Altshuler, M. W. (**1976**). "Effect of stimulus intensity, frequency, and unilateral hearing loss on sound localization," J Auditory Res **16**, 275-279.

Covey, E., Kauer, J. A., and Casseday, J. H. (**1996**). "Whole-cell patch-clamp recording reveals subthreshold sound-evoked postsynaptic currents in the inferior colliculus of awake bats," The Journal of neuroscience **16**, 3009-3018.

Creutzfeldt, O., Hellweg, F. C., and Schreiner, C. (**1980**). "Thalamocortical transformation of responses to complex auditory stimuli," Experimental Brain Research **39**, 87-104.

Dai, H., Scharf, B., and Buus, S. (**1991**). "Effective attenuation of signals in noise under focused attention," The Journal of the Acoustical Society of America **89**, 2837.

Djelani, T., Porschmann, C., Sahrhage, J., and Blauert, J. (**2000**). "An interactive virtual-environment generator for psychoacoustic research II: Collection of head-related impulse responses and evaluation of auditory localization," Acta Acustica united with Acustica **86**, 1046-1053.

Druga, R., and Syka, J. (**1984**). "Projections from auditory structures to the superior colliculus in the rat," Neuroscience letters **45**, 247-252.

Durlach, N. I., and Colburn, H. S. (**1978**). "Binaural phenomena," Handbook of perception **4**, 365-466.

Ebert Jr, C. S., Blanks, D. A., Patel, M. R., Coffey, C. S., Marshall, A. F., and Fitzpatrick, D. C. (**2008**). "Behavioral sensitivity to interaural time differences in the rabbit," Hearing research **235**, 134-142.

Ehret, G., and Schreiner, C. (**1997**). "Frequency resolution and spectral integration (critical band analysis) in single units of the cat primary

auditory cortex," Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology **181**, 635-650.

Elverland, H. (**1978**). "Ascending and intrinsic projections of the superior olivary complex in the cat," Experimental Brain Research **32**, 117-134.

Euston, D. R., and Takahashi, T. T. (**2002**). "From spectrum to space: the contribution of level difference cues to spatial receptive fields in the barn owl inferior colliculus," The Journal of neuroscience **22**, 284.

Evans, E. (**1972**). "The frequency response and other properties of single fibres in the guinea-pig cochlear nerve," The Journal of physiology **226**, 263.

Evans, E., Pratt, S., Spenner, H., and Cooper, N. (**1992**). "Comparisons of physiological and behavioural properties: Auditory frequency selectivity," Auditory physiology and perception, 159–169.

Faingold, C. L., Gehlbach, G., and Caspary, D. M. (**1989**). "On the role of GABA as an inhibitory neurotransmitter in inferior colliculus neurons: iontophoretic studies," Brain research **500**, 302-312.

Fay, R. R. (**1988**). *Hearing in vertebrates: A psychophysics databook* (Hill-Fay Associates).

Feddersen, W., Sandel, T., Teas, D., and Jeffress, L. (**1957**). "Localization of High Frequency Tones," The Journal of the Acoustical Society of America **29**, 988.

Fletcher, H. (**1940**). "Auditory patterns," Reviews of Modern Physics **12**, 47.

Fletcher, H., and Munson, W. A. (**1937**). "Relation between loudness and masking," Journal of the Acoustical Society of America.

Francart, T., and Wouters, J. (**2007**). "Perception of across-frequency interaural level differences," The Journal of the Acoustical Society of America **122**, 2826.

Frens, M., and Opstal, A. J. (**1995**). "A quantitative study of auditory-evoked saccadic eye movements in two dimensions," Experimental Brain Research **107**, 103-117.

Friauf, E., and Ostwald, J. (**1988**). "Divergent projections of physiologically characterized rat ventral cochlear nucleus neurons as shown by intra-

axonal injection of horseradish peroxidase," Experimental Brain Research **73**, 263-284.

Fuzessery, Z., and Pollak, G. (**1985**). "Determinants of sound location selectivity in bat inferior colliculus: a combined dichotic and free-field stimulation study," Journal of neurophysiology **54**, 757.

Gäsler, G. (**1954**). "Über die Hörschwelle für Schallereignisse mit verscgieden breitem Frequenzspektrum," Acustica, 408-414.

Gescheider, G. A. (**1997**). *Psychophysics: the fundamentals* (Lawrence Erlbaum).

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hearing research **47**, 103-138.

Glendenning, K., Baker, B., Hutson, K., and Masterton, R. (**1992**). "Acoustic chiasm V: inhibition and excitation in the ipsilateral and contralateral projections of LSO," The Journal of Comparative Neurology **319**, 100-122.

Glendenning, K. K., and Baker, B. N. (**1988**). "Neuroanatomical distribution of receptors for three potential inhibitory neurotransmitters in the brainstem auditory nuclei of the cat," The Journal of Comparative Neurology **275**, 288-308.

Godfrey, D. A., Kiang, N., and Norris, B. E. (**1975**). "Single unit activity in the dorsal cochlear nucleus of the cat," The Journal of Comparative Neurology **162**, 269-284.

Goldberg, J. M., and Brown, P. B. (**1968**). "Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study," Journal of neurophysiology **31**, 639.

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

Good, M. D., and Gilkey, R. H. (**1996**). "Sound localization in noise: The effect of signal to noise ratio," The Journal of the Acoustical Society of America **99**, 1108.

Good, M. D., Gilkey, R. H., and Ball, J. M. (**1997**). "The relation between detection in noise and localization in noise in the free field," Binaural and spatial hearing in real and virtual environments, 349-376.

Green, D. M. (**1988**). *Profile analysis: Auditory intensity discrimination* (Oxford University Press, USA).

Greenberg, G. Z., and Larkin, W. D. (**1968**). "Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: The probe-signal method," Journal of the Acoustical Society of America.

Guinan, J. J., Norris, B. E., and Guinan, S. S. (**1972**). "Single auditory units in the superior olivary complex: II: locations of unit categories and tonotopic organization," International Journal of Neuroscience **4**, 147-166.

Hafter, E. R., and Saberi, K. (**2001**). "A level of stimulus representation model for auditory detection and attention," The Journal of the Acoustical Society of America **110**, 1489.

Hafter, E. R., Schlauch, R. S., and Tang, J. (**1993**). "Attending to auditory filters that were not stimulated directly," Journal of the Acoustical Society of America.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectro temporal pattern analysis," The Journal of the Acoustical Society of America **76**, 50.

Harrison, J., and Howe, M. (**1974**). "Anatomy of the afferent auditory nervous system of mammals," Handbook of sensory physiology **5**, 284-336.

Harrison, J., and Warr, W. (**1962**). "A study of the cochlear nuclei and ascending auditory pathways of the medulla," The Journal of Comparative Neurology **119**, 341-379.

Harrison, R., and Evans, E. (**1979**). "Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig," European Archives of Oto-Rhino-Laryngology **224**, 71-78.

Hartmann, W. M., and Constan, Z. A. (**2002**). "Interaural level differences and the level-meter model," The Journal of the Acoustical Society of America **112**, 1037.

Hartmann, W. M., and Rakerd, B. (**1989**). "On the minimum audible angle: A decision theory approach," Journal of the Acoustical Society of America.

Hartmann, W. M., Rakerd, B., and Gaalaas, J. B. (**1998**). "On the source-identification method," The Journal of the Acoustical Society of America **104**, 3546.

Hawkins Jr, J., and Stevens, S. (**1950**). "The masking of pure tones and of speech by white noise," Journal of the Acoustical Society of America.

Heffner, H. (**1978**). "Effect of auditory cortex ablation on localization and discrimination of brief sounds," Journal of neurophysiology **41**, 963.

Heffner, H. E. (**1998**). "Auditory awareness," Applied Animal Behaviour Science **57**, 259-268.

Heffner, H. E., and Heffner, R. S. (**1984**). "Sound localization in large mammals: Localization of complex sounds by horses," Behavioral neuroscience **98**, 541.

Heffner, H. E., and Heffner, R. S. (**2008**). "High-frequency hearing," The Senses: A Comprehensive Reference. Audition **3**, 55-60.

Heffner, H. E., Heffner, R. S., Tollin, D., Populin, L., Moore, J., Ruhland, J., and Yin, T. (**2005**). "The sound-localization ability of cats," Journal of neurophysiology **94**, 3653.

Heffner, H. E., Ravizza, R. J., and Masterton, B. (**1969a**). "Hearing in primitive mammals, III: Tree shrew (Tupaia glis)," Journal of Auditory Research **9**, 12-18.

Heffner, H. E., Ravizza, R. J., and Masterton, B. (**1969b**). "Hearing in primitive mammals: IV, bushbaby (Galago senegalensis)," Journal of Auditory Research **9**, 19-23.

Heffner, R., and Heffner, H. (**1980**). "Hearing in the elephant (Elephas maximus)," Science **208**, 518.

Heffner, R. S., and Heffner, H. E. (**1982**). "Hearing in the elephant (Elephas maximus): Absolute sensitivity, frequency discrimination, and sound localization," Journal of Comparative and Physiological Psychology **96**, 926.

Heffner, R. S., and Heffner, H. E. (**1987**). "Localization of noise, use of binaural cues, and a description of the superior olivary complex in the smallest carnivore, the least weasel (< xh: i> Mustela nivalis</xh: i>)," Behavioral neuroscience **101**, 701.

Heffner, R. S., and Heffner, H. E. (**1988a**). "Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure," Hearing research **36**, 221-232.

Heffner, R. S., and Heffner, H. E. (**1988b**). "Sound localization and use of binaural cues by the gerbil (Meriones unguiculatus)," Behavioral neuroscience **102**, 422.

Heffner, R. S., and Heffner, H. E. (**1988c**). "Sound localization in a predatory rodent, the northern grasshopper mouse (Onychomys leucogaster)," Journal of Comparative Psychology **102**, 66.

Heffner, R. S., and Masterton, R. B. (**1990**). "Sound localization in mammals: brain-stem mechanisms," Comparative perception **1**, 285-314.

Henning, G. B. (**1974**). "Detectability of interaural delay in high frequency complex waveforms," The Journal of the Acoustical Society of America **55**, 84.

Hill, K., Stange, G., and Mo, J. (**1989**). "Temporal synchronization in the primary auditory response in the pigeon," Hearing research **39**, 63-73.

Hine, J. E., Martin, R. L., and Moore, D. R. (**1994**). "Free-field binaural unmasking in ferrets," Behavioral neuroscience **108**, 196.

Hirsch, J.A., and Martinez, L.M. (**2006**). "Circuits that build visual cortical receptive fields" Trends in Neurosciences **29**, 30-39.

Hirsh, I. J. (**1948**). "The influence of interaural phase on interaural summation and inhibition," The Journal of the Acoustical Society of America **20**, 536.

Houben, D., and Gourevitch, G. (**1979**). "Auditory lateralization in monkeys: an examination of two cues serving directional hearing," The Journal of the Acoustical Society of America **66**, 1057.

Houtsma, A. J. M. (**2004**). "Hawkins and Stevens revisited with insert earphones," The Journal of the Acoustical Society of America **115**, 967.

Huang, A. Y., and May, B. J. (**1996**). "Sound orientation behavior in cats. II. Mid frequency spectral cues for sound localization," The Journal of the Acoustical Society of America **100**, 1070.

Huffman, R. F., and Henson Jr, O. (**1990**). "The descending auditory pathway and acousticomotor systems: connections with the inferior colliculus," Brain research reviews **15**, 295-323.

Irvine, D. (**1986**). "The auditory brainstem," Progress in sensory physiology **7**, 1-279.

Irvine, D. (**1987**). "A comparison of two methods for the measurement of neural sensitivity to interaural intensity differences," Hearing research **30**, 169-179.

Irvine, D., and Gago, G. (**1990**). "Binaural interaction in high-frequency neurons in inferior colliculus of the cat: effects of variations in sound pressure level on sensitivity to interaural intensity differences," Journal of neurophysiology **63**, 570.

Irvine, D., Park, V., and McCormick, L. (**2001**). "Mechanisms underlying the sensitivity of neurons in the lateral superior olive to interaural intensity differences," Journal of neurophysiology **86**, 2647.

Irvine, D., Rajan, R., and Aitkin, L. (**1996**). "Sensitivity to interaural intensity differences of neurons in primary auditory cortex of the cat. I. Types of sensitivity and effects of variations in sound pressure level," Journal of neurophysiology **75**, 75.

Irving, R., and Harrison, J. (**1967**). "The superior olivary complex and audition: a comparative study," The Journal of Comparative Neurology **130**, 77-86.

Jacobsen, T. (**1976**). "Localization in noise," Acoustics Laboratory, Technical University, Lynby, Denmark.

Jeffress, L. A. (**1948**). "A place theory of sound localization," Journal of Comparative and Physiological Psychology **41**, 35.

Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (**1962**). "Masking and interaural phase: II. 167 cycles," Journal of the Acoustical Society of America.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory nerve fibers to single tones," The Journal of the Acoustical Society of America **68**, 1115.

Joris, P., and Yin, T. C. T. (**2007**). "A matter of time: internal delays in binaural processing," TRENDS in Neurosciences **30**, 70-78.

Joris, P. X., Smith, P. H., and Yin, T. C. T. (**1998**). "Coincidence detection minireview in the auditory system: 50 years after Jeffress," Neuron **21**, 1235–1238.

Joris, P. X., and Yin, T. (**1995**). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences," Journal of neurophysiology **73**, 1043.

Kacelnik, O., Nodal, F. R., Parsons, C. H., and King, A. J. (**2006**). "Training-induced plasticity of auditory localization in adult mammals," PLoS biology **4**, e71.

Kavanagh, G. L., and Kelly, J. B. (**1987**). "Contribution of auditory cortex to sound localization by the ferret (Mustela putorius)," Journal of neurophysiology **57**, 1746.

Kelly, J. B. (**1980**). "Effects of auditory cortical lesions on sound localization by the rat," Journal of neurophysiology **44**, 1161.

Kelly, J. B., Kavanagh, G. L., and Dalton, J. C. H. (**1986**). "Hearing in the ferret (Mustela putorius): thresholds for pure tone detection," Hearing research **24**, 269-275.

Kelly, J. B., and Sally, S. L. (**1988**). "Organization of auditory cortex in the albino rat: binaural response properties," Journal of neurophysiology **59**, 1756.

Kiang, N. Y. (**1965**). "Discharge Patterns of Single Fibers in the Cat's Auditory Nerve," (DTIC Document).

King, A. J., Nelken, I. (**2009**). "Unravelling the principles of auditory cortical processing: can we learn from the visual system?," Nature neuroscience **12**, 698-701.

King, A. J., Schnupp, J. W. H., and Doubell, T. P. (**2001**). "The shape of ears to come: dynamic coding of auditory space," Trends in cognitive sciences **5**, 261-270.

Kitzes, L. (**2008**). "Binaural interactions shape binaural response structures and frequency response functions in primary auditory cortex," Hearing research **238**, 68-76.

Kitzes, L. M., Wrege, K. S., and Cassady, J. M. (**1980**). "Patterns of responses of cortical cells to binaural stimulation," The Journal of Comparative Neurology **192**, 455-472.

Klein, S. A. (**2001**). "Measuring, estimating, and understanding the psychometric function: A commentary," Attention, Perception, & Psychophysics **63**, 1421-1455.

Klink, K. B., Bendig, G., and Klump, G. M. (**2006**). "Operant methods for mouse psychoacoustics," Behavior research methods **38**, 1-7.

Klug, A., Park, T. J., and Pollak, G. D. (**1995**). "Glycine and GABA influence binaural processing in the inferior colliculus of the mustache bat," Journal of neurophysiology **74**, 1701.

Knudsen, E. I., Blasdel, G. G., and Konishi, M. (**1979**). "Sound localization by the barn owl (Tyto alba) measured with the search coil technique," Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology **133**, 1-11.

Kohlrausch, A., and Sander, A. (**1995**). "Phase effects in masking related to dispersion in the inner ear: II. Masking period patterns of short targets," Journal of the Acoustical Society of America.

Kollmeier, B., Gilkey, R. H., and Sieben, U. K. (**1988**). "Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model," The Journal of the Acoustical Society of America **83**, 1852.

Köppl, C. (**1997**). "Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba," The Journal of neuroscience **17**, 3312.

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," The Journal of the Acoustical Society of America **62**, 157.

Kuwada, S., Batra, R., and Fitzpatrick, D. C. (**1997**). "Neural processing of binaural temporal cues," Binaural and spatial hearing in real and virtual environments, 399-425.

Langford, T. L. (**1984**). "Responses elicited from medial superior olivary neurons by stimuli associated with binaural masking and unmasking," Hearing research **15**, 39-50.

Leek, M. R. (**2001**). "Adaptive procedures in psychophysical research," Attention, Perception, & Psychophysics **63**, 1279-1292.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," Journal of the Acoustical Society of America **49**, 467–477.

Lewald, J., Dörrscheidt, G. J., and Ehrenstein, W. H. (**2000**). "Sound localization with eccentric head position," Behavioural brain research **108**, 105-125.

Liberman, M. C., and Kiang, N. Y. (**1978**). "Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity," Acta Oto-laryngologica.

Licklider, J. (**1948**). "The influence of interaural phase relations upon the masking of speech by white noise," Journal of the Acoustical Society of America.

Linschoten, M. R., Harvey, L. O., Eller, P. M., and Jafek, B. W. (**2001**). "Fast and accurate measurement of taste and smell thresholds using a maximum-likelihood adaptive staircase procedure," Attention, Perception, & Psychophysics **63**, 1330-1347.

Lorenzi, C., Gatehouse, S., and Lever, C. (**1999**). "Sound localization in noise in normal-hearing listeners," The Journal of the Acoustical Society of America **105**, 1810.

Macmillan, N. A., and Creelman, C. D. (**2005**). *Detection theory: A user's guide* (Lawrence Erlbaum).

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited," The Journal of the Acoustical Society of America **111**, 2219.

Makous, J. C., and Middlebrooks, J. C. (**1990**). "Two dimensional sound localization by human listeners," The Journal of the Acoustical Society of America **87**, 2188.

Martin, R. L., and Webster, W. R. (**1987**). "The auditory spatial acuity of the domestic cat in the inter aural horizontal and median vertical planes," Hearing research **30**, 239-252.

Masterton, B., Heffner, H., and Ravizza, R. (**1969**). "The evolution of human hearing," The Journal of the Acoustical Society of America **45**, 966.

May, B. J., and Huang, A. Y. (**1996**). "Sound orientation behavior in cats. I. Localization of broadband noise," The Journal of the Acoustical Society of America **100**, 1059.

May, B. J., Kimar, S., and Prosen, C. A. (**2006**). "Auditory filter shapes of CBA/CaJ mice: Behavioral assessments," The Journal of the Acoustical Society of America **120**, 321.

Middlebrooks, J. (**1992**). "Narrowband sound localization related to acoustical cues," (IEEE), pp. 0_53-50_54.

Middlebrooks, J. C., and Green, D. M. (**1990**). "Directional dependence of interaural envelope delays," The Journal of the Acoustical Society of America **87**, 2149.

Middlebrooks, J. C., and Green, D. M. (**1991**). "Sound localization by human listeners," Annual Review of Psychology **42**, 135-159.

Middlebrooks, J. C., and Pettigrew, J. D. (**1981**). "Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location," The Journal of neuroscience **1**, 107.

Middlebrooks, J. C., Xu, L., Eddins, A. C., and Green, D. M. (**1998**). "Codes for sound-source location in nontonotopic auditory cortex," Journal of neurophysiology **80**, 863.

Mills, A. W. (**1958**). "On the minimum audible angle," Journal of the Acoustical Society of America.

Mills, A. W. (**1960**). "Lateralization of high-frequency tones," Journal of the Acoustical Society of America.

Molino, J. A. (**1974**). "Psychophysical verification of predicted interaural differences in localizing distant sound sources," Journal of the Acoustical Society of America.

Moore, B. C. J. (**2003**). *An introduction to the psychology of hearing* (Emerald Group Pub Ltd).

Moore, B. C. J., and Glasberg, B. R. (**1987**). "Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns," Hearing research **28**, 209-225.

Moore, J. M., Tollin, D. J., and Yin, T. C. T. (**2008**). "Can measures of sound localization acuity be related to the precision of absolute location estimates?," Hearing research **238**, 94-109.

Morest, D. K., and Oliver, D. L. (**1984**). "The neuronal architecture of the inferior colliculus in the cat: defining the functional anatomy of the auditory midbrain," The Journal of Comparative Neurology **222**, 209-236.

Moushegian, G., Rupert, A., and Gidda, J. (**1975**). "Functional characteristics of superior olivary neurons to binaural stimuli," Journal of neurophysiology **38**, 1037.

Mrsic-Flogel, T. D., King, A. J., and Schnupp, J. W. H. (**2005**). "Encoding of virtual acoustic space stimuli by neurons in ferret primary auditory cortex," Journal of neurophysiology **93**, 3489.

Musicant, A. D., and Butler, R. A. (**1984**). "The influence of pinnae based spectral cues on sound localization," The Journal of the Acoustical Society of America **75**, 1195.

Nakamoto, K. T., Zhang, J., and Kitzes, L. M. (**2004**). "Response patterns along an isofrequency contour in cat primary auditory cortex (AI) to stimuli varying in average and interaural levels," Journal of neurophysiology **91**, 118.

Nelken, I., Chechik, G., Mrsic-Flogel, T. D., King, A. J., and Schnupp, J. W. H. (**2005**). "Encoding stimulus information by spike numbers and mean response time in primary auditory cortex," Journal of Computational Neuroscience **19**, 199-221.

Nelson, D. R. (**1967**). "Hearing thresholds, frequency discrimination, and acoustic orientation in the lemon shark, Negaprion brevirostris (Poey)," Bulletin of Marine Science **17**, 741-768.

Niemiec, A. J., Yost, W. A., and Shofner, W. P. (**1992**). "Behavioral measures of frequency selectivity in the chinchilla," Journal of the Acoustical Society of America.

Nodal, F., Bajo, V., Parsons, C., Schnupp, J., and King, A. (**2008**). "Sound localization behavior in ferrets: comparison of acoustic orientation and approach-to-target responses," Neuroscience **154**, 397-408.

Nuetzel, J. M., and Hafter, E. R. (**1981**). "Discrimination of interaural delays in complex waveforms: Spectral effects," The Journal of the Acoustical Society of America **69**, 1112.

Ocklenburg, S., Hirnstein, M., Hausmann, M., and Lewald, J. (**2010**). "Auditory space perception in left-and right-handers," Brain and cognition **72**, 210-217.

Oliver, D. L. (**1984**). "Dorsal cochlear nucleus projections to the inferior colliculus in the cat: a light and electron microscopic study," The Journal of Comparative Neurology **224**, 155-172.

Oliver, D. L. (**1987**). "Projections to the inferior colliculus from the anteroventral cochlear nucleus in the cat: possible substrates for binaural interaction," The Journal of Comparative Neurology **264**, 24-46.

Osen, K. K. (**1969**). "Cytoarchitecture of the cochlear nuclei in the cat," The Journal of Comparative Neurology **136**, 453-483.

Palmer, A. (**1987**). "Physiology of the cochlear nerve and cochlear nucleus," British medical bulletin **43**, 838.

Palmer, A. (**2007**). "Anantomy and physiology of the auditory brainstem," Auditory evoked potentials, 200-208.

Palmer, A., and Kuwada, S. (**2005**). "Binaural and spatial coding in the inferior colliculus," The inferior colliculus, 377–410.

Palmer, A., and Russell, I. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hearing research **24**, 1-15.

Palmer, A. R., and Shackleton, T. M. (**2009**). "Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency," JARO-Journal of the Association for Research in Otolaryngology **10**, 233-250.

Park, T. J., Grothe, B., Pollak, G. D., Schuller, G., and Koch, U. (**1996**). "Neural delays shape selectivity to interaural intensity differences in the lateral superior olive," The Journal of neuroscience **16**, 6554.

Park, T. J., and Pollak, G. D. (**1993**). "GABA shapes sensitivity to interaural intensity disparities in the mustache bat's inferior colliculus: implications for encoding sound location," The Journal of neuroscience **13**, 2050-2067.

Patterson, R., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1988**). "An efficient auditory filterbank based on the gammatone function," APU report **2341**.

Patterson, R. D. (**1976**). "Auditory filter shapes derived with noise stimuli," The Journal of the Acoustical Society of America **59**, 640.

Patterson, R. D., Nimmo Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," The Journal of the Acoustical Society of America **72**, 1788.

Patterson, R. D. a. M., B.C. (ed). (**1986**). *Auditory filters and excitation patterns of frequency resolution* (Academic Press, London).

Patterson, T. I. R., and Irino, T. (**1997**). "A time-domain, level-dependent auditory filter: The gammachirp," J. Acoust. Soc. Am **101**, 412-419.

Perrott, D. R., Ambarsoom, H., and Tucker, J. (**1987**). "Changes in head position as a measure of auditory localization performance: auditory psychomotor coordination under monaural and binaural listening conditions," The Journal of the Acoustical Society of America **82**, 1637.

Pfeiffer, R. R. (**1966**). "Classification of response patterns of spike discharges for units in the cochlear nucleus: tone-burst stimulation," Experimental Brain Research **1**, 220-235.

Phillips, D., and Irvine, D. (**1981**). "Responses of single neurons in physiologically defined area AI of cat cerebral cortex: sensitivity to interaural intensity differences," Hearing research **4**, 299-307.

Phillips, D., Judge, P., and Kelly, J. (**1988**). "Primary auditory cortex in the ferret (Mustela putorius): neural response properties and topographic organization," Brain Research **443**, 281-294.

Phillips, D. P. (**1990**). "Neural representation of sound amplitude in the auditory cortex: effects of noise masking," Behavioural brain research **37**, 197-214.

Pickles, J. (**1975**). "Normal critical bands in the cat," Acta Oto-laryngologica **80**, 245-254.

Pickles, J. (**1979**). "Psychophysical frequency resolution in the cat as determined by simultaneous masking and its relation to auditory nerve resolution," The Journal of the Acoustical Society of America **66**, 1725.

Pickles, J. O. (**1982**). *An introduction to the physiology of hearing* (academic Press London).

Pierce, A. H. (**1901**). *Studies in auditory and visual space perception* (Longmans, Green, and Co.).

Radziwon, K. E., June, K. M., Stolzberg, D. J., Xu-Friedman, M. A., Salvi, R. J., and Dent, M. L. (**2009**). "Behaviorally measured audiograms and gap detection thresholds in CBA/CaJ mice," Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology **195**, 961-969.

Rajan, R., Aitkin, L., Irvine, D., and McKay, J. (**1990**). "Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters," Journal of neurophysiology **64**, 872.

Ramon, Y., and Cajal, S. (**1909**). "Histologie du systeme nerveux de l'homme et des vertebres," Maloine, Paris, 774–838.

Ravizza, R. J., Heffner, H. E., and Masterton, B. (**1969**). "Hearing in primitive mammals. I. Opossum (Didelphis virginianus)," Journal of Auditory Research **9**, 1-7.

Rayleigh, L. (**1907**). "On our perception of sound direction," Philos. Mag **13**, 214–232.

Reale, R. A., and Kettner, R. E. (**1986**). "Topography of binaural organization in primary auditory cortex of the cat: effects of changing interaural intensity," Journal of neurophysiology **56**, 663.

Recanzone, G. H., Guard, D. C., and Phan, M. L. (**2000**). "Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey," Journal of neurophysiology **83**, 2315.

Recanzone, G. H., Makhamra, S. D. D. R., and Guard, D. C. (**1998**). "Comparison of relative and absolute sound localization ability in humans," The Journal of the Acoustical Society of America **103**, 1085.

Reed, C. M., and Bilger, R. C. (**1973**). "A comparative study of S/N and E/N," The Journal of the Acoustical Society of America **53**, 1039.

Renaud, D. L., and Popper, A. N. (**1975**). "Sound localization by the bottlenose porpoise Tursiops truncatus," Journal of Experimental Biology **63**, 569.

Rhode, W., Oertel, D., and Smith, P. (**1983**). "Physiological response properties of cells labeled intracellularly with horseradish peroxidase in cat ventral cochlear nucleus," The Journal of Comparative Neurology **213**, 448-463.

Rhode, W. S., and Smith, P. H. (**1986**). "Physiological studies on neurons in the dorsal cochlear nucleus of cat," Journal of neurophysiology **56**, 287.

Richardson, B. W. (**1879**). "Some researches with Professor Hughes' new instrument for the measurement of hearing; the audiometer," Proceedings of the Royal Society of London **29**, 65.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," Journal of neurophysiology **30**, 769.

Ruggero, M. A., Santi, P. A., and Rich, N. C. (**1982**). "Type II cochlear ganglion cells in the chinchilla," Hearing research **8**, 339-356.

Rutkowski, R. G., Wallace, M. N., Shackleton, T. M., and Palmer, A. R. (**2000**). "Organisation of binaural interactions in the primary and dorsocaudal fields of the guinea pig auditory cortex," Hearing research **145**, 177-189.

Ryan, A. (**1976**). "Hearing sensitivity of the mongolian gerbil, Meriones unguiculatis," The Journal of the Acoustical Society of America **59**, 1222.

Saberi, K. (**1998**). "Modeling interaural-delay sensitivity to frequency modulation at high frequencies," The Journal of the Acoustical Society of America **103**, 2551.

Saberi, K., Dostal, L., Sadralodabai, T., and Bull, V. (**1991**). "Free-field release from masking," Journal of the Acoustical Society of America.

Sabin, A. T., Macpherson, E. A., and Middlebrooks, J. C. (**2005**). "Human sound localization at near-threshold levels," Hearing research **199**, 124-134.

Sachs, M. B., and Young, E. D. (**1980**). "Effects of nonlinearities on speech encoding in the auditory nerve," The Journal of the Acoustical Society of America **68**, 858.

Sachs, M. B., Young, E. D., and Lewis, R. H. (**1974**). "Discharge patterns of single fibers in the pigeon auditory nerve," Brain Research **70**, 431-447.

Saint Marie, R. L., Stanforth, D., and Jubelier, E. (**1997**). "Substrate for rapid feedforward inhibition of the auditory forebrain," Brain Research **765**, 173-176.

Sally, S. L., and Kelly, J. B. (**1992**). "Effects of superior olivary comples lesions on binaural responses in rat inferior colliculus," Brain Research **572**, 5-18.

Samson, F. K., Barone, P., Irons, W. A., Clarey, J. C., Poirier, P., and Imig, T. J. (**2000**). "Directionality derived from differential sensitivity to monaural and binaural cues in the cat's medial geniculate body," Journal of neurophysiology **84**, 1330.

Sandel, T., Teas, D., Feddersen, W., and Jeffress, L. (**1955**). "Localization of sound from single and paired sources," Journal of the Acoustical Society of America.

Sanes, D. H. (**1990**). "An in vitro analysis of sound localization mechanisms in the gerbil lateral superior olive," The Journal of neuroscience **10**, 3494.

Scharf, B. (**1970**). "Critical Bands," in *Foundations in Modern Auditory Theory*, edited by J. V. Tobias (Academic Press, New York).

Schmidt, S., Türke, B., and Vogler, B. (**1983**). "Behavioural audiogram from the bat, Megaderma lyra," Myotis **21**, 62–66.

Schnupp, J. W. H., Mrsic-Flogel, T. D., and King, A. J. (**2001**). "Linear processing of spatial cues in primary auditory cortex," Nature **414**, 200-204.

Scholes, C., Palmer, A. R., and Sumner, C. J. (**2010**). "Forward suppression in the auditory cortex is frequency specific," European Journal of Neuroscience.

Schreiner, C. E., and Mendelson, J. R. (**1990**). "Functional topography of cat primary auditory cortex: distribution of integrated excitation," Journal of neurophysiology **64**, 1442-1459.

Seeber, B. (**2002**). "A new method for localization studies," Acta Acustica united with Acustica **88**, 446-450.

Semple, M., and Aitkin, L. (**1979**). "Representation of sound frequency and laterality by units in central nucleus of cat inferior colliculus," Journal of neurophysiology **42**, 1626.

Semple, M., and Kitzes, L. (**1987**). "Binaural processing of sound pressure level in the inferior colliculus," Journal of neurophysiology **57**, 1130.

Semple, M., and Kitzes, L. (**1993**). "Binaural processing of sound pressure level in cat primary auditory cortex: evidence for a representation based on absolute levels rather than interaural level differences," Journal of neurophysiology **69**, 449.

Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1992**). "Across frequency integration in a model of lateralization," J. Acoust. Soc. Am **91**, 2276-2279.

Shaw, E. (**1974**). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," The Journal of the Acoustical Society of America **56**, 1848.

Shelton, B., and Searle, C. (**1978**). "Two determinants of localization acuity in the horizontal plane," The Journal of the Acoustical Society of America **64**, 689.

Smith, P. H., Joris, P. X., Carney, L. H., and Yin, T. C. T. (**1991**). "Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat," The Journal of Comparative Neurology **304**, 387-407.

Smith, P. H., Joris, P. X., and 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," The Journal of Comparative Neurology **331**, 245-260.

Smith, P. H., Joris, P. X., and Yin, T. C. T. (**1998**). "Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat," Journal of neurophysiology **79**, 3127.

Spangler, K. M., Warr, W. B., and Henkel, C. K. (**1985**). "The projections of principal cells of the medial nucleus of the trapezoid body in the cat," The Journal of Comparative Neurology **238**, 249-262.

Spiegel, M. F. (**1981**). "Thresholds for tones in maskers of various bandwidths and for signals of various bandwidths as a function of signal frequency," The Journal of the Acoustical Society of America **69**, 791.

Spoendlin, H. (**1978**). "The afferent innervation of the cochlea," in *Evoked Electrical Activity in the Auditory Nervous System*, edited by R. F. N. a. C. Fernandez (Academic Press, New York), pp. 21-41.

Stecker, G. C., and Middlebrooks, J. C. (**2003**). "Distributed coding of sound locations in the auditory cortex," Biological cybernetics **89**, 341-349.

Sterbing, S. J., Hartung, K., and Hoffmann, K. P. (**2003**). "Spatial tuning to virtual sounds in the inferior colliculus of the guinea pig," Journal of neurophysiology **90**, 2648.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988**). "Lateralization of complex binaural stimuli: A weighted image model," The Journal of the Acoustical Society of America **84**, 156.

Stevens, S. S., and Newman, E. B. (**1936**). "The localization of actual sources of sound," The American Journal of Psychology, 297-306.

Stotler, W. (**1953**). "An experimental study of the cells and connections of the superior olivary complex of the cat," The Journal of Comparative Neurology **98**, 401-431.

Su, T. I. K., and Recanzone, G. H. (**2001**). "Differential effect of near-threshold stimulus intensities on sound localization performance in azimuth and elevation in normal human subjects," JARO-Journal of the Association for Research in Otolaryngology **2**, 246-256.

Sumner, C. J. a. P., A. (**2010**). "Responses of the ferret auditory nerve to tones," in *Midwinter Meeting - Association for Research in Otolaryngology* (Anaheim).

Sutter, M. L., and Schreiner, C. E. (**1991**). "Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex," Journal of neurophysiology **65**, 1207.

Swets, J. A. (**1986**). "Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance," Psychological Bulletin **99**, 181.

Syka, J., Popelar, J., Kvasnak, E., and Astl, J. (**2000**). "Response properties of neurons in the central nucleus and external and dorsal cortices of the inferior colliculus in guinea pig," Experimental Brain Research **133**, 254-266.

Tanner, W. P. (**1956**). "Some general properties of the hearing mechanism."

Terhune, J. (**1974**). "Sound localization abilities of untrained humans using complex and sinusoidal sounds," Scandinavian Audiology **3**, 115-120.

Thurlow, W. R., and Runge, P. S. (**1967**). "Effect of induced head movements on localization of direction of sounds," Journal of the Acoustical Society of America.

Tollin, D. J. (**2003**). "The lateral superior olive: a functional role in sound source localization," The neuroscientist **9**, 127.

Tollin, D. J., Populin, L. C., Moore, J. M., Ruhland, J. L., and Yin, T. C. T. (**2005**). "Sound-localization performance in the cat: the effect of restraining the head," Journal of neurophysiology **93**, 1223.

Tollin, D. J., and Yin, T. C. T. (**2001**). "Comparison of the response properties of the principal cells of the lateral superior olive and the medial nucleus of the trapezoid body," Assoc Res Otolaryngol **24**, 57.

Trahiotis, C., and Bernstein, L. R. (**1986**). "Lateralization of bands of noise and sinusoidally amplitude modulated tones: Effects of spectral locus and bandwidth," The Journal of the Acoustical Society of America **79**, 1950.

Tsuchitani, C. (**1977**). "Functional organization of lateral cell groups of cat superior olivary complex," Journal of neurophysiology **40**, 296.

Tsuchitani, C. (**1988**). "The inhibition of cat lateral superior olive unit excitatory responses to binaural tone bursts. I. The transient chopper response," Journal of neurophysiology **59**, 164.

Tsuchitani, C. (**1994**). "The brain stem evoked response and medial nucleus of the trapezoid body," Otolaryngology--head and neck surgery: official journal of American Academy of Otolaryngology-Head and Neck Surgery **110**, 84.

Tsuchitani, C. (**1997**). "Input from the medial nucleus of trapezoid body to an interaural level detector," Hearing research **105**, 211-224.

Tsuchitani, C., and Boudreau, J. (**1966**). "Single unit analysis of cat superior olive S segment with tonal stimuli," Journal of neurophysiology **29**, 684.

Ulanovsky, N., Las, L., and Nelken, I. (**2003**). "Processing of low-probability sounds by cortical neurons," nature neuroscience **6**, 391-398.

Vater, M., Habbicht, H., Kössl, M., and Grothe, B. (**1992**). "The functional role of GABA and glycine in monaural and binaural processing in the inferior colliculus of horseshoe bats," Journal of Comparative Physiology A **171**, 541-553.

Von Békésy, G. (**1960**). "Experiments in hearing."

Wakeford, O. S., and Robinson, D. E. (**1974**). "Lateralization of tonal stimuli by the cat," Journal of the Acoustical Society of America.

Walker, K. M. M., Schnupp, J. W. H., Hart-Schnupp, S. M. B., King, A. J., and Bizley, J. K. (**2009**). "Pitch discrimination by ferrets for simple and complex sounds," The Journal of the Acoustical Society of America **126**, 1321.

Wallace, M. N., Rutkowski, R. G., and Palmer, A. R. (**2000**). "Identification and localisation of auditory areas in guinea pig cortex," Experimental Brain Research **132**, 445-456.

Warr, B. W. (**1972**). "Fiber degeneration following lesions in the multipolar and globular cell areas in the ventral cochlear nucleus of the cat," Brain Research **40**, 247-270.

Warr, W. B. (**1966**). "Fiber degeneration following lesions in the anterior ventral cochlear nucleus of the cat," Experimental neurology **14**, 453-474.

Watson, C. S. (**1987**). "Uncertainty, informational masking, and the capacity of immediate auditory memory."

Weber, D. L. (**1977**). "Growth of masking and the auditory filter," The Journal of the Acoustical Society of America **62**, 424.

Wenstrup, J. J. (**1984**). "Auditory sensitivity in the fish-catching bat, Noctilio leporinus," Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology **155**, 91-101.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). "Localization using nonindividualized head-related transfer functions," Journal-acoustical Society of America **94**, 111-111.

Wightman, F. L., and Kistler, D. J. (**1989**). "Headphone simulation of free-field listening. I: Stimulus synthesis," J. Acoust. Soc. Am **85**, 858-867.

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," Journal of the Acoustical Society of America.

Winer, J. A., Larue, D. T., Diehl, J. J., and Hefti, B. J. (**1998**). "Auditory cortical projections to the cat inferior colliculus," The Journal of Comparative Neurology **400**, 147-174.

Winer, J. A., and Schreiner, C. (**2005**). *The inferior colliculus* (Springer Verlag).

Wu, S. H., and Kelly, J. B. (**1992**). "NMDA, non-NMDA and glycine receptors mediate binaural interaction in the lateral superior olive: physiological evidence from mouse brain slice," Neuroscience letters **134**, 257-260.

Yao, L., and Peck, C. (**1997**). "Saccadic eye movements to visual and auditory targets," Experimental Brain Research **115**, 25-34.

Yin, T., and Chan, J. (**1990**). "Interaural time sensitivity in medial superior olive of cat," Journal of neurophysiology **64**, 465.

Yost, W. A. (**1971**). "Internal Delay Discrimination," The Journal of the Acoustical Society of America **50**, 88.

Yost, W. A. (**1974**). "Discriminations of interaural phase differences," Journal of the Acoustical Society of America.

Yost, W. A., and Dye Jr, R. H. (**1988**). "Discrimination of interaural differences of level as a function of frequency," The Journal of the Acoustical Society of America **83**, 1846.

Yost, W. A., and Shofner, W. P. (**2009**). "Critical bands and critical ratios in animal psychoacoustics: An example using chinchilla data," The Journal of the Acoustical Society of America **125**, 315.

Zhang, J., Nakamoto, K. T., and Kitzes, L. M. (**2004**). "Binaural interaction revisited in the cat primary auditory cortex," Journal of neurophysiology **91**, 101.

Zwicker, E. (**1965**). "Temporal Effects in Simultaneous Masking by White Noise Bursts," The Journal of the Acoustical Society of America **37**, 653.