Frequency selectivity measured both psychophysically and physiologically

Tobias Trevor Wells

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

July 2014

Abstract

The ability to resolve the individual frequency components of a complex sound is known as frequency selectivity. The auditory system seems to act as a series of overlapping band-pass filters or auditory filters (AF), the width of which describe frequency selectivity. It is a fundamental and extensively studied property of the auditory system, yet its neural basis is not fully understood. This is due, in part, to the fact that two distinct approaches are taken to explore it; psychophysics and physiology, the results of which are difficult to reconcile. AF are measured in quite different ways in the two sub-fields.

Psychophysics measures the ability of the system as a whole by using masking paradigms to measure the bandwidth of AF from behavioural data. A preferred method in human psychophysics is notched noise (NN) masking using forward masking with fixed signal level, since it does not suffer from confounds like suppression and off-frequency listening. In physiology bandwidth can be measured for single cells, or a population of cells, at various levels of the auditory system, and is traditionally done by observing the number of action potentials elicited in response to pure tone stimuli. The auditory system is known to be highly non-linear and so comparing the results of such vastly different approaches is problematic. Also, psychophysics measures the detection threshold of 'signal' sounds; how this compares to mean spike rates used in physiology is not clear.

Attempts have been made in the past to apply the same method in animals to measure both psychophysical and physiological bandwidths, with varying degree of success. No successful attempt has been made to use an up-to-date method used in human psychophysics. In this thesis I take a step towards comparing psychophysical and physiological results by 1) developing a novel method that allows forward masked NN bandwidths to be measured behaviourally in the ferret, and 2) applying the same

psychophysical paradigm to measuring bandwidth in guinea pig inferior colliculus (IC) and primary auditory cortex (A1) neurons. In addition a signal detection theory (SDT) approach is used on the physiological data to make results more comparable to psychophysical ones.

Results from the behavioural method show that it can be used to successfully measure both forward and simultaneously masked NN bandwidths in the same animal, and that these measurements are in close agreement with one another and with bandwidths measured using previous methods. Results from the guinea pig physiology study show that bandwidths measured from IC neurons using the psychophysical NN paradigm are narrower than pure tone estimates of bandwidth, in the same neurons. However, the NN estimates are in close agreement with auditory peripheral and perceptual bandwidths, a finding which differs substantially from previous studies. Unexpectedly, however, bandwidths estimated from A1 neurons using masking show much finer tuning at high frequencies than seen further down the auditory system. This tuning is not only narrower than pure tone tuning in these neurons, but also finer than psychophysically measured estimates, which represent the auditory system as a whole. However, this may be related to the greater non-linearity of cortical neurons compared with those in the midbrain and lower.

This work demonstrates that it is possible to reconcile different measurements of tuning in the auditory system by using appropriate methods. It also highlights the complex nature of auditory neurons and how care must be taken when measuring frequency selectivity using different approaches. In addition it provides a method for measuring auditory bandwidths psychophysically and physiologically in the same animal, allowing a direct comparison between the two; a vital step in investigating the neural basis of perceptual frequency selectivity.

Acknowledgements

I'd like to thank Chris for all the help and guidance he's given me throughout the entire PhD process. To Trevor, Mark Wallace and Alan for all the help, support and advice they gave me at a moment's notice; many a sudden panic suddenly disappeared upon their arrival. To Rachel, Sammy and Zoe for basically doing most of my ferret behavioural work for me, I could literally not have done this PhD without that. Also for never laughing at my ridiculous breathing apparatus (at least not to my face ☺). To Victor for spending the best part of 2 years wrestling sitafc with me and plugging all the holes that appeared every time a change was made. I hope the next generation of software gives you far fewer headaches.

A big thanks has to go to Liz for supplying me with cookies & cream Kitkats® and a much needed dose of calming conversation and sanity on a daily basis. If you ever leave I'm certain productivity at the IHR would plummet. Also to Ben, Joe, Pete, Rob, Steadmuffin, Joel and Fletch, not only for the help you all gave throughout the PhD, but for just being awesome and making the whole process fun.

A more general life thank you to Mum and Dad for all the support of every possible kind that you've given throughout not only the PhD but across the board. I'm taking my first belated steps into the adult world now so hopefully the burden will lessen a bit. Maybe. Also to Julie who's always been there for me, and who's support singlehandedly kept me sane throughout the writing up process. Thanks for putting up with all my whining and stress; I'd like to say it's over now but that would be a lie.

So in summary, it's been bloody hard, it's been great fun, and you all have to call me doctor now. Booyah!

Abbreviations

Area under receiver operating characteristic curve (AUROC)

Auditory brainstem response (ABR)

Auditory filter (AF)

Auditory nerve (AN)

Auditory nerve fibre (ANF)

Basilar membrane (BM)

Best frequency (BF)

Characteristic frequency (CF)

Cochlear nucleus (CN)

Confidence interval (CI)

Critical band(width) (CB)

Critical ratio (CR)

Dorsal cochlear nucleus (DCN)

Equivalent rectangular bandwidth (ERB)

Excitatory frequency tuning curve (eFTC)

Forward masked notched noise masking (Fwd)

Frequency response area (FRA)

Frequency tuning curve (FTC)

Inferior colliculus (IC)

Inhibitory frequency tuning curve (iFTC)

Input/output (I/O)

Intra-muscular (i.m)

Intra-peritoneal (i.p)

Just noticeable difference (JND)

Lateral lemniscus (LL)

Local field potential (LFP)

Maximum length sequence (MLS)

Medial geniculate body (MGB)

Medial olivocochlear (MOC)

Multi-unit (MU)

Notched noise (NN)

One-interval two-alternative forced choice (1I2AFC)

Otoacoustic emission (OAE)

Outer hair cell (OHC)

Post-/Peri-stimulus time histogram (PSTH)

Primary auditory cortex (A1)

Rate against level function (RLF)

Receiver operating characteristic (ROC)

Rippled noise (RN)

Root mean square (RMS)

Signal detection theory (SDT)

Signal frequency (SF)

Signal to noise ratio (SNR)

Simultaneously masked notched noise masking (Sim)

Single-unit (SU)

Spectro-temporal response field (STRF)

Stimulus-frequency otoacoustic emissions (SFOAE)

Superior olivary complex (SOC)

Tucker-Davis Technologies (TDT)

Two-interval two-alternative forced choice (2I2AFC)

Ventral cochlear nucleus (VCN)

ΔF = Distance in Hz between the centre frequency of the signal and the closest edges of the bands of the notched noise masker.

dB SPL = Level in decibels referenced to 20 μPa (Sound Pressure Level).

mmHg = millimetre of mercury, equivalent to 133.32 Pa.

Contents

CHAPTER 1: General introduction

The aim of auditory neuroscience is to understand how sound in the real world is detected, manipulated, analysed, and represented by the brain (and sensory nervous system) to allow the person or animal to gain information about their environment. This understanding can be sought at different levels. One can identify and investigate the way in which different aspects of sound are perceived, such as pitch, loudness, spatial location, and object segregation, as well as the system's limits and fidelity. It is possible to infer details about the auditory system using such psychoacoustic information, and this type of analysis is known as *psychophysics*.

As far as we know the only information carrying elements in the brain are action potentials elicited (or fired) from neurons. All information about the external world that is captured by the sense organ (in this case the ear) must therefore be represented by the activity of neurons. Auditory neuroscience is also concerned with how different features of sound, such as frequency, level, object identity and location, are represented in this neural activity, and how accurate this representation is. This *physiological* level gives more access to the actual mechanisms driving the perception of hearing. A complete understanding of hearing arguably requires being able to explain how the low and high levels are linked.

One of the most fundamental features of sound is frequency, which is very closely related to the perception of pitch. Sound is caused by the vibration of objects causing air molecules (or any other medium) to 'pulse', i.e. have alternating regions of low pressure (rarefaction) and high pressure (condensation). These waves of pulses move out from the vibrating body, with the molecules not moving a net amount. Frequency simply describes the speed with which these molecules move, or the number of pulses in a given distance or time. The pulses can have very complicated patterns, which reflect the nature of the object and its movement. There is therefore an evolutionary advantage to

being able to extract such information from sound, e.g. being able to identify the deep rumbling of an avalanche or the high pitched squeak of prey, will aid survival.

Any first step in such information extraction requires the complex pulses in sound to be broken down into smaller, constituent pieces; a particular sound can then be identified by the quantities of these 'building blocks' that are present. One useful metric to use is the speed of the pulses, or frequency of the sound. Slow pulses are described by low frequency and fast pulses by high. Sounds of one frequency are known as tones or pure tones. A complex sound (one which contains multiple frequency components) can simply be represented by the summation of its individual frequency components, and conversely, can be broken down into individual frequency components. A very useful way of representing this (borrowed from a description of heat transfer in engineering) is to describe a single frequency as a sinusoid, with the number of pulses represented by the number of periods in a given distance/time. All sounds can then be summed from and decomposed into sinusoids of various amplitude, frequency and phase. Such construction and deconstruction is known as Fourier analysis (Fourier, 1822). It turns out that the auditory system represents sound in a very similar way.

Any sound can be broken down into a continuum of infinitesimally narrow frequency components, but the auditory system is not perfect and can only represent these frequencies at a certain fidelity. An important problem in auditory research is to describe the accuracy with which it does this, also known as 'frequency resolution'. This ability forms the basis of many complex tasks such as picking out separate voices in a musical harmony, or more practically, identifying speech in a noisy environment.

Perceptually, one measure of this, known as 'frequency discrimination', is the smallest difference in frequency between sounds separated in time, that can be detected by the auditory system (Shower and Biddulph, 1931). This usually involves playing two tones consecutively and getting the person/animal to identify if they are the same or different in frequency. The just noticeable difference (JND) is the lowest frequency difference that can be accurately detected, and changes for different frequencies and levels of a pure tone.

Another more general approach, and one that is the basis of my research and this thesis, is to measure the ability to separately identify the individual sinusoidal components in a complex sound (Moore, 2003). This is commonly referred to as 'frequency selectivity', or simply frequency resolution or frequency analysis. Here the feature of interest is

whether, and under what circumstances, a sound (usually a pure tone) is detectable. The sound of which the detectability is being tested is often called the 'signal'.

In the rest of this *General introduction* I will be describing the development of the field of frequency selectivity research and the important concepts and discoveries along the way. Each sub-section describes a separate aspect of the field and wherever possible has been placed in a chronological and logical order. However, since all these aspects developed in parallel, it becomes difficult to present them linearly and at times knowledge is assumed that is not fully described until later on.

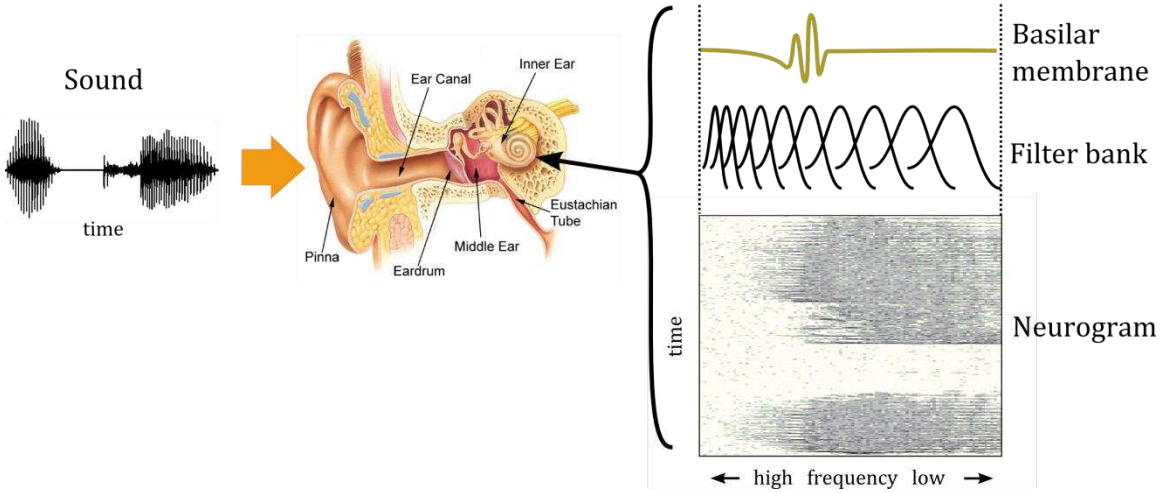


Figure 1.1

A schematic showing the constituent parts of the auditory periphery and how sound is transcribed into neural activity. Sound enters the ear and vibrates the tympanic membrane. These vibrations are then transferred to the cochlea where they spread along the BM according to frequency. This activity can be thought of as a filter bank; a series of overlapping filters than run along the BM. The vibration causes hair cells to move, causing auditory nerve cells to fire, demonstrated in the neurogram; darker colour means higher firing. This neurogram was produced by an auditory nerve model of 100 units (log spaced in frequency) in response to the sound wave shown. Middle ear image taken from http://imgarcade.com/1/ear-inside/.

1.1 The auditory periphery

For hundreds of years the outer, middle and inner ear, also known as the auditory periphery, has been thought of as a resonator system (Du Verney, 1683). This 'resonance' theory has been refined over time (Helmholtz, 1877; Barton, 1908) such that now we understand all the resonating features of the auditory periphery and how they result in a neural place code.

The outer ear comprises the pinna (the visible portion) and the ear canal (or meatus) (Figure 1.1). Sound enters the ear, travels down the ear canal and causes the ear drum (or tympanic membrane) to vibrate. The ear drum marks the divide between the outer and middle ear. Three small bones (or ossicles) connect the ear drum to the cochlea, in the inner ear, and efficiently transfer the vibrations between the two. The cochlea contains an elongated chamber bisected by the cochlea partition, which consists of the basilar membrane (BM) and organ of Corti. This chamber is curled into a spiral, which has no impact on its mechanics and is believed to be purely space saving, therefore often the cochlea is described as a straight chamber as if it had been 'unravelled'. It is filled with a near-incompressible fluid, and the mechanics of the cochlea are such that the vibrations of the ossicles cause pressure waves to form within this fluid. The BM is relatively narrow and stiff at the base (end closest to the ossicles) and wider and less stiff at the apex. It is free to vibrate and the pressure wave in the cochlear fluid causes it to move like a travelling wave; this wave grows from the basal end, gradually increases in amplitude, before abruptly stopping after it has reached maximum amplitude. The thickness, density, and shape of the BM is such that the travelling wave peaks at different points depending on the frequency of the sound, with higher frequency sounds causing a peak towards the base of the cochlea, and lower frequency sound towards the apex. Hair cells all along the organ or Corti convert the movement of the BM into neural activity, more specifically they release neurotransmitters that cause auditory nerve fibres (ANF) to produce action potentials, and so by observing where along the cochlea the neurons are firing one can work out the frequency characteristics of the sound. This gradual change from high to low frequency represented by the 'place' of neural activity is known as a tonotopic map (or cochleotopic), and is seen throughout the auditory system (Moore, 2003). From the cochlea the neural activity is propagated throughout the auditory system and the job of separating, analysing, and interpreting the sound begins.

1.2 The basilar membrane and its non-linearities

In many ways a major part of understanding frequency selectivity is understanding the mechanics of the cochleae. This is the first point at which the vibrations of sound are transferred into electrochemical signals, and therefore has a significant impact on the properties and limits of the system as a whole. So much so that it is widely accepted that psychophysical estimates of frequency selectivity, which are the result of the entire auditory system, 'reflect the filtering properties of the cochlea' (Moore, 1995).

One major feature is that the vibration of the cochlear partition is not just passive but has an active component. Von Békésy (1944) looked at the movement of the BM in dead cochlea in response to very loud sounds, and found that psychophysical bandwidth measures were far too sharp to be fully due to these passive mechanics. There must therefore be an active component, designed to sharpen tuning, that is only present when the cochlea is alive. This too was corroborated in other studies where a link between the physiological condition of the cochlea and tuning was found; the healthier the cochlea, the sharper the tuning, with dead cochlea tuning being the broadest (Rhode, 1971, 1973; Rhode and Robles, 1974; Rhode, 1978; Sellick et al., 1982). Early research focussed on the basal end of the cochlea, due to it being relatively accessible in animals, and this effect is particularly noticeable in this region. It was later shown that this active component is most probably due to the activity of several rows of hair cells on the organ of Corti known as the outer hair cells (OHC) (Ashmore, 1987; Smith et al., 1987; Hudspeth, 2008).

Moreover it was found that mechanical frequency selectivity varies with position along the BM (Von Békésy and Wever, 1960; Rhode, 1978). By playing a tone of a certain frequency and recording the movement of the BM in response, it is possible to build an input/output (I/O) function, describing the relative output level in response to a certain input level. It was found that the active component causes a gain in output for the region on the BM tuned to the frequency of the tone. This results in a non-linear I/O function. Regions further away do not receive this gain, and show linear I/O functions (Sachs and Abbas, 1974), as do dead cochleae. An example of each can be seen in Figure 1.2. For a given frequency the cochlear gain is linear for low levels of sound (in Figure 1.2 the red line is parallel to the blue line < 30 dB SPL (decibel sound pressure level)), above which it decreases in effectiveness. For very high levels, the active component has little effect and the passive mechanics of the BM dominate, resulting in no active gain. Psychophysical studies of such functions in humans show that the cut off level after

which gain is no longer as effective (the first 'kink' in the red line of Figure 1.2) is 20-55 dB SPL (Murugasu and Russell, 1995; Oxenham and Moore, 1997; Ruggero et al., 1997; Russell and Nilsen, 1997). The result of such localised gain is to sharpen tuning. Let us say, for example, a pure tone of 5 kHz is played. It will cause the BM to vibrate over a wide region, with maximum activity around the region representing 5 kHz. The active gain boosts the response over this 5 kHz region comparatively more than the rest of the BM. If we assume the system has a fixed resolution (or constant 'internal noise') then this relative increase at the 5 kHz region will make it more distinct from the activity on the rest of the BM. Sharpest tuning therefore occurs when gain is maximal, for healthy cochleae and low levels of sound. This also means measures of tuning are dependent on stimulus level. This is discussed in more detail in 1.4 Auditory filter shape.

In the previous paragraph I described the passive component of BM activity as 'linear' and the active component as 'non-linear'. It may be prudent to define what is meant by this. A linear system must have both 'homogeneity' and 'superposition'.

> *Homogeneity*: If the input of a system is altered by a scale factor, then the output must also be altered by the same scale factor, and be otherwise unchanged. For example, if the level of a sound is doubled, then the level of the internal response to that sound should also double, without adding any other frequency components. This results in a linear I/O function.
>
> *Superposition:* The output of a system to a number of simultaneously occurring inputs should be the same as the outputs of the system to each of those individual inputs, summed up. For example, a complex sound made up of two pure tones of frequency f_1 and f_2, with amplitudes A and B, respectively, should elicit the same response as A × the response to just frequency f_1, plus B × the response to just frequency f_2. This is expressed algebraically as:

$$\mathbf{S}\left(\sum_i a_i f_i\right) = \sum_i a_i \mathbf{S}(f_i) \qquad (1.1)$$

> where **S** is a function describing the response of the system, f_i is each individual frequency component in the complex sound, and a_i is its amplitude.

The active gain makes the system non-linear because for a given sound the boost that is applied, over regions of the BM, is frequency dependent. Effectively the values of a_i are not the same for both sides of *equation* (1.1).

Not only is the BM non-linear in this way, but it was found that the maximum gain that can be applied varies along the length of the BM. Hicks and Bacon (1999) measured gain psychophysically in humans and found that the maximum amount of gain increased with increasing tone frequency (and therefore more basal BM activity). Cooper and Yates (1994) examined activity of ANF of the guinea pig and inferred activity on the BM from it, using tonotopic maps described in Liberman (1982) and Robertson et al. (1980). They found a compressive non-linearity (so called because of the 'compressed' region of the I/O function (red line in Figure 1.2)) along most of the BM, which increased with frequency; there was none for frequencies below 500 Hz.

The non-linear nature of frequency selectivity is not just present in the early stages of the auditory system but is a running theme throughout (see 1.8.5 Central non-linearities). It has implications for any attempt to quantify tuning, and has led to the wide variety of methods and measurements described later in this chapter. If the system were linear then any method to measure bandwidth would produce the same answer. However, non-linearities mean different methods, or the same method with slightly different parameters, will give different results. Neither result is correct or incorrect, just affected differently by the non-linearities in the system. It is therefore very important to take note of method and parameters when looking at measures of tuning.

As mentioned previously the field of neuroscience has two main branches: psychophysics and physiology. The study of frequency selectivity is no exception, and the two branches have developed in parallel with much crossover between the two. However, below I describe the development of each separately, starting with psychophysics since this is where the first ground breaking studies were conducted.

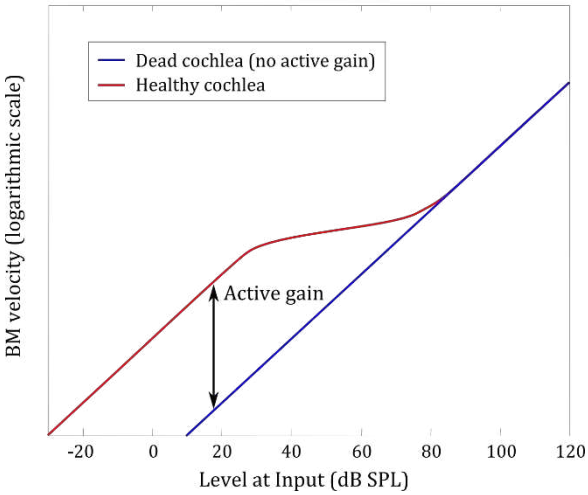


Figure 1.2

Example I/O functions for the cochlea. For a healthy cochlea, in which OHCs are intact and able to provide active gain, the I/O function for a pure tone will look like the red line. For low input levels (< 30 dB SPL) the active component provides gain, increasing the relative level at output. At intermediate input levels this gain drops off, creating the compressive region seen at 30-80 dB SPL input levels. At high levels (> 80 dB SPL) the cochlear gain is no longer effective and the system becomes linear once more. The blue line represents a linear system, so for a dead cochlea or for a region of the BM far away from the area tuned to the frequency of the test tone.

1.3 Psychophysics

1.3.1 The discovery of masking

At the core of almost all psychophysical investigations into frequency selectivity lies the phenomenon of masking. This is defined as 'the process by which the threshold of audibility of one sound is raised by the presence of another (masking) sound' (ANSI, 1994). Early psychophysical studies showed that a sound could be rendered less detectable by the presence of another sound played at the same time; they named this disruptive sound the 'masker'. Moreover the effectiveness of this masker to reduce the detectability of the signal is greater if the two are closer in frequency (Mayer, 1894; Wegel and Lane, 1924). These early experiments were conducted using pure tones, but it was shown that this was also true for maskers that were narrow bands of noise (sound that has energy over a small range of frequencies) (Fletcher and Munson, 1937). The amount of masking (another way of describing the decrease in signal detectability) is also dependent on the level of the masker, with a 1 dB increase in masker level resulting in a 1 dB increase in signal detection threshold (at least for most frequencies in the audible human range) (Hawkins Jr and Stevens, 1950). In other words, the threshold at which a signal can be detected in noise is a relatively constant signal-to-noise ratio (SNR). This led to the idea that frequency selectivity could be explored by manipulating the characteristics of the signal and masker, be it frequency, level, temporal, and observing the effect this has on the detectability of the signal.

Even now over a hundred years later the mechanisms behind masking are still not fully understood, but are believed to be multifaceted and not just caused by a single feature of the auditory system. One contributing factor is BM mechanics resulting in interference between stimuli called 'suppression', which is described in more detail later. However, central (i.e. neural) mechanisms also contribute (Costalupes et al., 1984). This is described in more detail later in 1.8.9 Mechanisms behind masking.

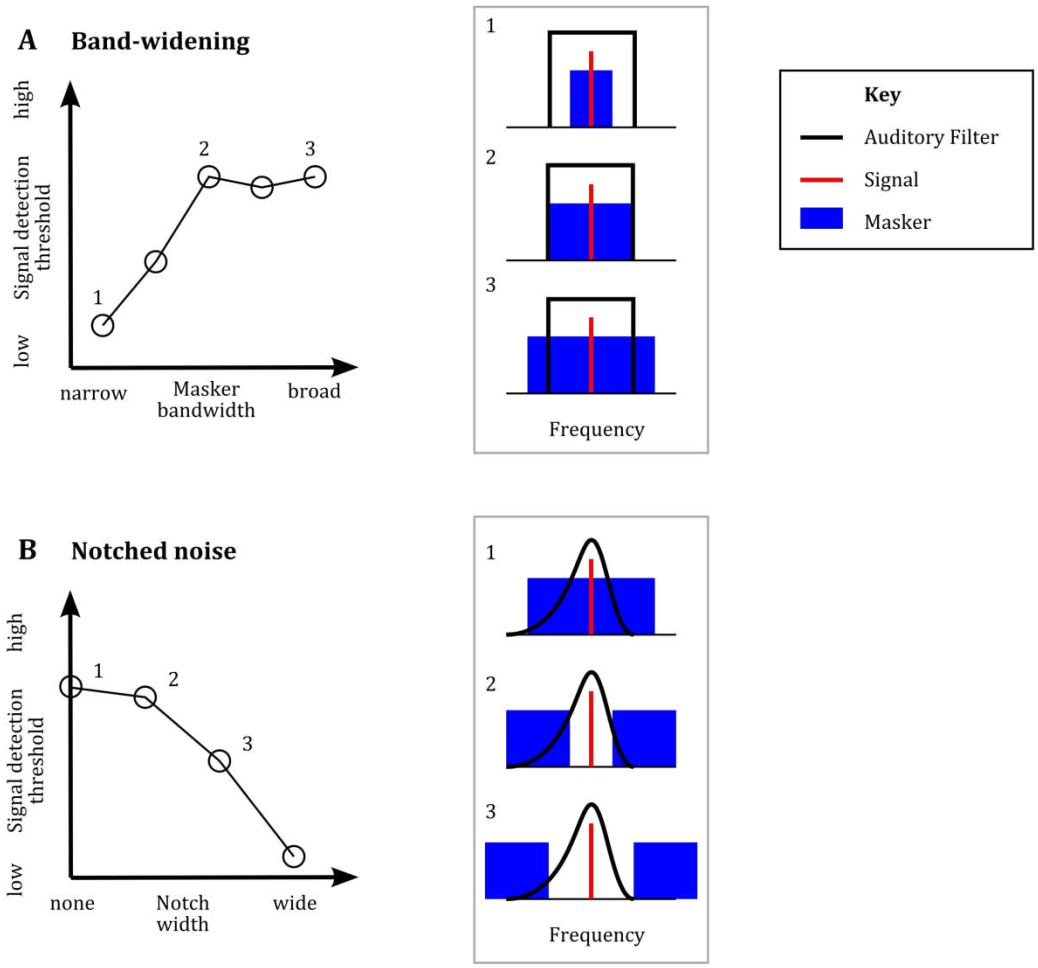


Figure 1.3

A *Detection threshold of the signal against bandwidth of masker: 1) Masker bandwidth low so low amount of masking, 2) Masker bandwidth equals CB so amount of masking reaches maximum, 3) Wider masker bandwidth adds no more energy into filter so signal detection level remains constant.* ***B*** *Detection threshold of signal against size of notch: 1) No notch so masking at maximum, 2 & 3) As notch width increases less masker energy falls into the filter, meaning lower detection thresholds of the signal.*

1.3.2 Power spectrum model

Arguably the birth of the field as we know it came with (Fletcher, 1940). Here Fletcher conducted a series of experiments in which he measured the detection threshold of a fixed level pure tone in the presence of a narrow band masker centred on the frequency of the signal. He observed that as the bandwidth of the masker increased (effectively getting louder) the detection threshold of the signal increased at a steady rate. However, at a certain point the detection threshold abruptly stopped increasing and remained constant for all larger bandwidths, despite the masker still getting louder (Figure 1.3 panel A). He dubbed this point the 'critical bandwidth' (CB) and used it as a measure of frequency selectivity. To explain this phenomenon Fletcher suggested that the auditory

periphery acts as a series of band-pass filters that overlap one another and run along the full length of the BM. These filters allow only a narrow range of frequencies through determined by the CB; all frequencies outside have no effect. The main factor determining the detection threshold of a signal is the amount of masker energy falling into the filter surrounding it, i.e. the more masker energy in the filter, the harder the signal is to detect. As the masker increases in bandwidth, with its spectrum level staying constant, more energy falls into the filter and the detection threshold of the signal increases. Once the bandwidth of the masker reaches the CB no further frequencies can pass into the filter and so the amount of effective masking energy reaches a maximum. Since the signal level is constant, the detection threshold also stays constant. In this way the detection threshold of the signal can also be thought of as a constant SNR at the output of this filter. He also assumed that when detecting a pure tone one always picks the filter with the best SNR at the output. Such a filter became known as an auditory filter (AF) and this set of assumptions became known as the 'power spectrum model' of masking, since only the long term power spectra of the stimuli are important and the phase and short term temporal effects are ignored (Patterson, 1986).

1.3.3 Band-widening and the critical ratio

The method described above is known as 'band-widening' and has been used extensively to measure AF bandwidth (Zwicker et al., 1957). Studies using this method reported much variability between subjects but in general CB increased with signal frequency, meaning higher frequencies have broader filters. The idea of the CB persists to this day despite many problems. One major problem, which does not just affect band-widening but many psychophysical methods, is off-frequency listening (Johnson-Davies and Patterson, 1979). This refers to the assumption that as the frequency characteristics of the masker change, the auditory system keeps using the same AF. If the filter with the highest SNR is always picked by the homunculus, or whatever mechanism detects the sound in higher order brain regions, then as the masker bandwidth is increased other neighbouring filters may give a better SNR. This would mean the CB wasn't just representing one filter but several different ones at different points, which would lead to much broader estimates of bandwidth. Another problem is that band-widening is very susceptible to stimulus variability (Patterson and Henning, 1977). Fletcher made a simplifying assumption that AFs are rectangular which we now know isn't true (see 1.4 Auditory filter shape for more details); they are in fact more like Gaussian functions with long sloping tails. The majority of power therefore passes through the centre (tip) of the filter and a relatively small amount through the edges (tails); masking energy

passing through the tails contributes a relatively small amount to the overall SNR at the output. Small short term fluctuations in the level and frequency of the masker therefore have a far greater impact on signal detection than the subtle changes in masking that occur as the masker bandwidth approaches the CB. Also the bandwidth of the masker has a significant effect on its envelope (the shape of a sound's amplitude in time); the narrower the band, the more its natural envelope fluctuates. Therefore the instantaneous level of narrow bands are much more variable and uncontrollable than broader bands, leading to much more variability in detection threshold. This problem is most notable when measuring the centre of the filter, meaning that both tip and tail measurements of the AF are dominated by fluctuations in noise. This causes a lot of variation between identical measurements and makes it difficult to give a definitive answer on CB. The problem is so severe that band-widening was actually damned with the description 'totally unsuited to measure the critical bandwidth' (Bos and de Boer, 1966).

Another method, that is a simplification of band-widening, is finding the 'critical ratio' (CR). This is done by measuring the detection threshold of a signal in the presence of a broadband masker. It is assumed that the AF is filled fully with masker energy and so the detection threshold of the signal, multiplied by a factor *K*, represents CB. This factor, however, has to be estimated using a more detailed method, such as band-widening, and was found to vary with signal frequency and across individuals (Patterson, 1986; Moore et al., 1997).

1.3.4 Masking with pure tones

A more direct approach to exploring the AF is to establish the detection threshold of a pure tone signal, of fixed frequency, in the presence of a tone masker, of various frequencies. In this way one can directly explore the graded change from no masking, when the masker frequency is far away from that of the signal, to maximum masking, when the two frequencies are the same (Chistovich, 1957; Houtgast, 1973; Zwicker, 1974). This has the added benefit of exploring the shape of the filter in more detail. However, this method also suffers from off-frequency listening and, additionally, from a consequence of the non-linearity of the BM known as 'two tone suppression' (Kemp and Chum, 1980; Ruggero et al., 1992). Because pure tones don't just elicit an infinitesimally small response on the BM, it is possible for the activity of two pure tones to interact with one another. At certain distances apart on the BM the activity caused by the pure tones cancel each other out, dampening the response to the stimulus. In the tone masking

experiment, therefore, certain masker frequencies will be damped mechanically on the BM and the detection threshold of the signal will be affected adversely, since it is not known exactly how it affects different stimuli. On top of this it was found that at some frequencies the pure tone masker and signal can combine in such a way that beats are audible. These beats can be used as cues for the presence of the signal even if the signal is inaudible, thus causing lower detection thresholds (Plomp and Steeneken, 1968; Rodenburg et al., 1974; Terhardt, 1974; Wightman et al., 1977; Patterson, 1986). A variant was also developed, using a narrow band masker as opposed to a tone, in an attempt to alleviate the problem of beats (Zwicker, 1954; Glasberg et al., 1982). The occurrence of beats was reduced, but they were not eliminated and the problem persists (Buus, 1985).

1.3.5 Rippled noise

Another method that emerged in the late 70's was to use rippled noise (RN), or comb-filtered noise, to probe bandwidth (Houtgast, 1974, 1977; Pick et al., 1977). The idea is to use broadband noise that has been modulated in frequency to have peaks and troughs, and to find the detection threshold of a pure tone signal for various modulation rates. The amount of masking depends on the precise relationship between the peaks and the troughs of the masker power spectrum and the filter. Generally speaking though, the higher the modulation rate, the more 'ripples' fall into the filter and the more masking occurs. By looking at the change in signal detection threshold with modulation rate, it is possible to work out the size of the filter. This method does not suffer from problems of beats and the effect of off-frequency listening is somewhat more complicated but thought to be less severe than band-widening or tone masking.

1.3.6 Notched noise masking

All the methods mentioned above have their faults, most notably that of off-frequency listening. Notched noise (NN) masking emerged to deal with this persisting problem, and is the method that has prevailed. It was first used by Webster et al. (1952) but didn't become a commonly used method until the mid-70's (Patterson, 1976; Patterson and Henning, 1977; Weber, 1977; Fidell et al., 1983; Glasberg and Moore, 1986; Dubno and Dirks, 1989). The basic idea is similar to that of band-widening, but inverted in frequency; instead of slowly 'filling' the AF with the masker, it is emptied. As in all other methods, the signal is a pure tone but here the masker consists of two narrow bands of noise separated by a gap or 'notch' in frequency (see Figure 1.3 panel B). This notch is usually centred on the frequency of the signal, and the detection threshold of the signal

is determined for notches of increasing width, i.e. the two bands of noise move further apart in frequency. As the notch increases, less of the masker energy falls inside the filter, causing less masking and the detection threshold of the signal drops. By looking at the rate at which the threshold drops with notch width one can infer the size, as well as the shape of the AF (see 1.4 Auditory filter shape). In some cases the masker is not two distinct bands that move further apart in frequency, but broadband noise with a notch of increasing size cut out; this is more commonly referred to as 'band stop' noise. The two band version has the benefit of constant masker power regardless of notch width. Also constant bandwidth means a constant amount of natural envelope fluctuation, removing any potential bias in detection thresholds.

When the notch width is very large, only a very small amount of the masker energy falls into the far tails of the filter. These tails attenuate almost all the energy so the changes in SNR at the output of the AF with notch width are very subtle. In band-widening these subtle differences are swamped by the small changes in masker energy passing through the centre of the filter. With notched noise the only masking energy passing through the filter is in these edges, so there is nothing to swamp it. This makes determining the edges of the filter far more accurate (Moore and Glasberg, 1983b). Also, since the noise bands symmetrically flank the signal, in frequency, the AF with the best SNR will always be the one centred on the signal. Off-frequency listening is therefore not a problem since all other filters will have a poorer SNR and thus will be ignored (Patterson and Nimmo Smith, 1980). (This is only true if AFs are also symmetric, which is technically not the case, although they are very close approximations. This is discussed in more detail in 1.4 Auditory filter shape below). Both these factors make the notched noise method better than tone masking, band-widening and the critical ratio (Patterson, 1986).

Another important aspect of the AF, which has a dramatic effect on estimates of frequency selectivity, is its shape. This line of research developed in parallel with the psychophysical methods described in this section, mainly in conjunction with notched noise masking, and is described in the following section.

1.4 Auditory filter shape

The width of the AF is the crucial feature when looking at frequency selectivity, and this is dependent on its shape. One benefit of describing filter shape, beyond obtaining an insight into the workings of the auditory system, is that fewer detection thresholds need be measured to obtain a bandwidth estimate. Also one can get an impression of how reliable the data are by how closely they match the stereotypical shape.

As mentioned previously, in his 1940 paper Fletcher simplified AF shape to a rectangle, but also found that regardless of shape, the bandwidth of the AF increased with signal frequency (SF) (Fletcher, 1953). He used a full spectrum broadband masker to mask pure tone signals of different frequencies. As the frequency of the signal increased so did the detection threshold. The masker was the same throughout so the difference must be caused by an increase in masking energy passing through the AF, therefore larger AFs. This is at least partly due to the decrease in cochlear gain with increasing frequency, described in 1.2 The basilar membrane and its non-linearities.

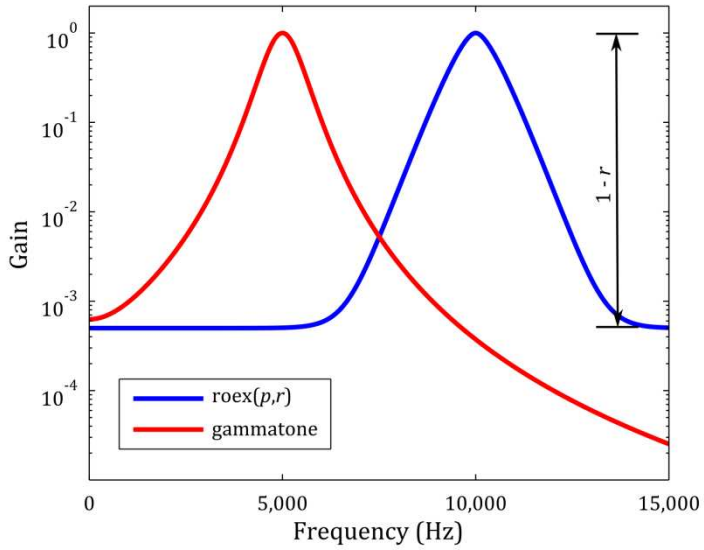


Figure 1.4

Examples of a roex(p,r) and a gammatone function. Gammatone parameters: a *= 1,* n *= 4,* b *= 700,* f *= 5 kHz. The order (n) determines the steepness of the sides, the bandwidth (b) determines how wide it is. Roex parameters:* p *= 30,* r *= 5e-4, CF = 10 kHz.* p *determines how steep the slopes are,* r *determines the position of the floor. The slopes appear straight due to the logarithmic scale on the 'Gain' axis.*

Schafer et al. (1950) first investigated the shape of the filter using band-widening and showed that the AF was not rectangular since there was no discontinuity in the detection threshold function. They dubbed the new AF shape a 'universal-resonance filter' with gradually sloping sides. This was corroborated by a notched noise study (Webster et al., 1952) and later a pure tone masking study (Small Jr, 1959). Swets et al. (1962) compared a number of potential filter shapes, namely rectangular, universal-resonance and Gaussian filter using band-widening, and showed that the critical bandwidth depended greatly on the assumed filter shape. Investigation into the precise AF shape was therefore critical. Patterson (1974) and Margolis and Small (1975) used low pass and high pass noise to investigate the shape, but detailed investigation didn't really take off until notched noise became established. Patterson (1976) used notched noise and showed that the Gaussian function was a good approximation to the AF close to the centre of the filter but that the tails dropped off too fast. Patterson and Nimmo Smith (1980) showed that a rounded exponential function would make a much better fit, and in Patterson et al. (1982) the roex(*p*,*r*) function was first suggested as the AF shape. This function is

$$W(g) = (1 - r)(1 + pg)\mathrm{e}^{-pg} + r \qquad (1.2)$$

where W is the filter weighting function, g is the deviation from the centre of the filter, p is the parameter defining the slopes of the function, and r defines the floor of the filter (see Figure 1.4). At the same time the symmetry of the filter was investigated by using asymmetrically spaced notch widths. It was found that the AF was slightly broader on the lower edge than the upper. This lead to a variant of the roex(*p*,*r*) with two p parameters; one defining the slope of each tail, allowing for an asymmetric AF. However, for moderate noise levels AF was found to be only very slightly asymmetric, on a linear frequency scale, such that an approximation of symmetry is entirely reasonable (Patterson, 1974; Patterson and Nimmo Smith, 1980; Patterson, 1986). More refinement led to the roex(*p*,*w*,*t*) and roex(*p*,*w*,*p*,*t*) functions in which the upper and lower edges, and lower edge only, respectively, of the roex function have two slope variables (Glasberg et al., 1984b; Glasberg and Moore, 2000). This creates 'skirts' allowing shallower slopes further away from the centre of the filter, and steeper slopes more centrally, although it has been argued that these additional filter functions are 'too flexible' to be of much use (Rosen et al., 1998).

These skirts are an influence of the theory of an active and a passive component to cochlear mechanics, described above in 1.2 The basilar membrane and its non-

linearities. In this situation the passive component can be thought of as a broad, low gain filter with shallow slopes, and the active component as a narrow, high gain filter with steep slopes. The active component would dominate at frequencies close to the signal, and the passive component for frequencies further away. This would lead to an AF with steep slopes near the centre, and shallower slopes further away. In Glasberg et al. (1999) they did away with the AF shape completely and attempted to directly fit a two filter model to notched noise data; such a model is known as a 'cascade filter' model. This, however, did not fit such data as well as a roex(p,r) based fitting procedure (Glasberg and Moore, 2000).

The roex is not the only modern function to describe AF shape. In Holdsworth et al. (1988) it was suggested that a gammatone function could be used to simulate AF shape. Such a function is simply the product of a sinusoid and a gamma distribution, and has the following formula in the time domain,

$$gammatone(t) = at^{n-1}\mathrm{e}^{-2\pi bt}\cos(2\pi ft + \phi) \quad (1.3)$$

where f is the centre frequency (in Hz), ϕ is the phase of the carrier (in radians), a is the amplitude, b is the filter bandwidth (in Hz), n is the filter order, and t is time (in seconds) (see Figure 1.4). This function has the advantage of having fewer parameters than the roex and has a fixed slight asymmetry determined by its order. Also it has a simple implementation in the time domain meaning that it is perfectly suited for use in filter bank models which require quick and effective filtering of a signal in the time domain. The disadvantage, however, is that because of its fixed shape and fewer parameters (which is an advantage for finding stable fits to data), it does not always fit notched noise data as well as the roex. The roex, gammatone functions, cascade models and all their variants are all used today to describe frequency selectivity for different implementations since they each have their advantages and disadvantages. A nice description of these can be found in (Lyon et al., 2010).

Another universal feature of AF shape is that of level dependence. The highly non-linear nature of the active filter described above means that the level of the fixed stimulus has an effect on the final AF shape. It was noted in several studies calculating filter shape at various fixed levels of signal and masker, that as level increases, AF get broader (Patterson, 1986; Moore, 1987; Glasberg and Moore, 1990; Rosen et al., 1992; Rosen and Baker, 1994; Baker et al., 1998; Baker and Rosen, 2006). This broadening is a result of a shallower slope on the lower edge of the AF, so as level increases asymmetry also increases. In addition this is a true change in AF shape and not just a non-linearity or

change in efficiency of the detector following the filter (Lutfi and Patterson, 1984). Since AF is affected by level, it is very important, when comparing bandwidth measures across studies, to make sure they use the same psychophysical experiment, with the same stimulus parameters.

Developments in AF shape were driven mainly by improvements in the notched noise method, which was refined over many years, exploring the effects of stimulus duration, frequency, level, temporal arrangements and many more (Moore and Glasberg, 1981; Moore et al., 1984; Moore et al., 1987). This led to the classic (and much referenced) measurement of frequency selectivity in humans, using simultaneous masking and fixed masker level, presented in Glasberg and Moore (1990). Two key features of masking experiments (that do not only apply to notched noise masking) concern the temporal arrangement of the stimuli, and which stimulus should be kept fixed in level. These have been a matter for debate and are described in the next two sub-sections.

1.5 Fixed signal versus fixed masker level

Detection threshold of a signal is usually determined psychophysically by varying the level of the stimuli and producing a psychometric function, which describes the relationship between stimulus level and proportion correct of the psychophysical task. Levels are set such that at one extreme the signal is not detectable and the subject performs at chance (i.e. proportion of trials that are correct is equivalent to if the subject was randomly guessing), and at the other extreme they correctly identify the signal 100 % of the time. The function transitions smoothly between the two extremes and detection threshold is (usually) taken to be the stimulus level at the steepest point of this psychometric function. In a two-interval two-alternative forced choice (2I2AFC) task (one of the standard psychophysical tasks), this transition usually takes the shape of a logistic function with minimum 50 % correct (since there are 2 intervals) and peak 100 %. Threshold is determined at 75 % correct. An example of such a psychometric function can be seen in Figure 3.3.

To create a psychometric function only the level of one stimulus needs to vary, the other is kept fixed. The question is which one to keep fixed: the signal or the masker? Keeping the signal level fixed is known as the 'isoresponse' approach and is believed to reflect taking the output SNR of the AF to determine signal detection. When the masker level is kept fixed this is known as the 'isolevel' approach and is believed to reflect using the SNR at the input to the AF to determine signal detection. In a linear system the two are simply inversions of one another, and it is not important which one is used. However, in a non-linear system (like the auditory system) they will give different answers and the choice becomes relevant (Eustaquio-Martin and Lopez-Poveda, 2011).

In early notched noise experiments there did not appear to be a difference between the bandwidth measures of each approach (Moore and Glasberg, 1981; Glasberg and Moore, 1982) and so due to easier implementation, the fixed masker level approach was generally adopted. However, in the 90's a series of papers were published asserting that it would be better to fix the signal level (Rosen et al., 1992; Rosen and Baker, 1994; Baker et al., 1998; Rosen et al., 1998). There were several reasons for this. One reason is that in the fixed masker level condition, maskers with very wide notches still determine detection even if they are too far away to have any impact. When a signal is detected without a masker, one still gets a psychometric function from the detection task. This is attributed to 'internal noise' causing variable success around threshold. If the notch is so wide that the masker has no impact on signal detection, the detection threshold will be

determined by this internal noise and not the masker. Therefore if the signal level is varied it becomes difficult to know exactly which of these is causing the masking when the notch is large and the stimulus masking low. If the signal level is fixed at a level above absolute threshold, then one can be sure that the signal is always detectable and that any effective masking is being done by the stimulus. Also Rosen et al. (1992), Rosen and Baker (1994), Baker et al. (1998) and Rosen et al. (1998) showed that filter shape was dependent on stimulus level (see 1.4 Auditory filter shape) and created a model to describe this. They found that fixing the signal level gave a much better fit to their model and a clearer explanation of cochlear mechanics. Also results matched up very closely with measurements made directly on the BM. Recent studies, however, have contested this view (Eustaquio-Martin and Lopez-Poveda, 2011; Lopez-Poveda and Eustaquio-Martin, 2013). They show, using a model of the auditory periphery, that when the signal level is fixed the notched noise method tends to underestimate the sharpness of the AF.

1.6 Forward versus simultaneous masking

A key feature of all masking experiments is that (at least) two sounds are played: a signal and masker. However, the two do not have to be presented at the same time for masking to occur. When the two are played at the same time it is known as simultaneous masking. If, however, the masker immediately precedes the signal, masking still occurs, and this is known as forward masking (Fwd) (Mayer, 1876; De Mare, 1940; Lüscher and Zwislocki, 1947; Munson and Gardner, 1950; Zwislocki et al., 1959; Plomp, 1964; Elliott, 1971; Widin and Viemeister, 1979; Jesteadt et al., 1982; Moore and Glasberg, 1983a; Nelson, 1991; Plack and Oxenham, 1998). Figure 2.3 shows an example of both temporal arrangements in a notched noise experiment. Masking can also occur if the masker directly follows the signal. This is known as backward masking but is very weak and poorly understood (Elliott, 1971). Forward masking is thought to be partly due to the mechanics of the BM. Short duration maskers (clicks ~ 100 μs) can lead to 'ringing' on the BM, in which activity persists beyond the end of the masker presentation (Duifhuis, 1970, 1973; Plack and Moore, 1990). This means that activity evoked by the signal and masker occur on the BM at the same time, effectively producing simultaneous masking (Sim). It is caused by the finite response time of the AFs, and is most prominent at low frequencies. This effect, however, is negligible for longer duration maskers and higher frequencies (Vogten, 1978; Gorga et al., 1980; Carlyon, 1988). Forward masking is therefore thought to be mainly neural in nature and its origins are a topic of much debate (see 1.8.7 Physiological characteristics of forward masking and 1.8.8 Theories behind psychophysical forward masking for more details).

What is known is that forward masking is not as effective as simultaneous masking, and decays with distance between masker offset and signal onset; this decay is roughly proportional to the log of the time distance (Moore and Glasberg, 1981; Moore and Glasberg, 1983a; Oxenham and Moore, 1995). Duration also has an impact, with the amount of masking increasing for masker lengths up to at least 50 ms (Fastl, 1976), and up to 200 ms in some studies (Kidd Jr and Feth, 1982; Zwicker, 1984). The effectiveness of the masker also increases with increasing masker level, but this relationship is not linear (like simultaneous masking) (Moore and Glasberg, 1983a), and a 1 dB increase in masker level only increases the detection threshold of the signal by about 0.6 dB (Jesteadt et al., 1982), down to as little as ~ 0.2 dB SPL (Plack and Oxenham, 1998). The function describing the change in detection threshold with masker level is known as a 'growth of forward masking' function.

In every psychophysical experiment in which the two temporal arrangements have been used (not just for notched noise but also for tone masking and others), forward masked estimates of bandwidth have been narrower than simultaneous (Houtgast, 1973, 1974; Leshowitz and Lindstrom, 1977; Wightman et al., 1977; Vogten, 1978; O'Loughlin and Moore, 1981; Glasberg et al., 1984a; Moore et al., 1984; Weber and Patterson, 1984). In one study forward masking was only 17 % narrower (Moore and Glasberg, 1981), but in others the difference has been more than a factor of 2 (Houtgast, 1977; Moore, 1978). This sometimes substantial difference has been attributed to the suppression or 'two tone inhibition', described in 1.3.4 Masking with pure tones above (Sachs and Kiang, 1968; Heinz et al., 2002). The idea is that, just like with tone masking, the interaction between signal and masker on the BM mechanically damps the activity in certain regions of the BM causing less masking, leading to wider bandwidth estimates. Because the amount and effect of suppression is not well understood, recent notched noise experiments tend to favour the forward masking arrangement since this does not suffer from this phenomenon. In addition a comparison between notched noise and rippled noise experiments using both forward and simultaneous masking arrangements (Glasberg et al., 1984a), showed no difference between the methods with simultaneous masking. However, the notched noise forward masked estimates were significantly narrower than both simultaneous masked arrangements and the forward masked rippled noise.

1.7 Behavioural measures in animals

All of the psychophysical experiments described above were conducted in humans, however, psychophysical experiments have also been conducted in animals. The attraction is not simply to learn more about frequency selectivity across the animal kingdom, but also the potential for physiological experiments that are not possible in humans. Comparing both in the same animal model has the potential to help us understand much more about the mechanisms of auditory frequency resolution. Conducting a psychophysical experiment in an animal is much more difficult than in a human, since one needs to train the animal to learn a relatively complex and at times abstract task through operant conditioning, instead of simply telling them what to do. Also motivating the animal to do the task is difficult since monetary reward, or simply doing the investigator a favour, does not work with animals as it does with humans. Furthermore even without these disadvantages, much psychophysical work in humans selects the highest performing/highly trained subjects; the length of training and relative limit in population size makes this unrealistic for animal research.

Many tasks have been developed to obtain signal discrimination responses of an animal, each often tailored for that species. Examples include lever pressing, go/no-go and localisation tasks. Also various ways of motivating the animal, such as rewards for correct behaviour, punishment for incorrect behaviour, food/water regulation combined with food/water reward for correct behaviour, and many more. Each method has its own characteristics of response, which must be taken into account when comparing studies. For example, shock avoidance tasks, where a shock can only occur following inaction, tend to encourage false alarm responses since a miss usually means shock and pain. However, tasks are usually designed to minimise cognitive load (so that any incorrect responses are a result of detection ability and not task confusion) and minimise bias.

All the methods described in previous sections for measuring frequency selectivity in humans have also been tried in animals, coupled with a behavioural task described in the previous paragraph. Band-widening has been used in a range of animals such as the cat (Watson, 1963; Pickles, 1975), chinchilla (Miller, 1964; Clark et al., 1975; Seaton and Trahiotis, 1975), rat (Gourevitch, 1965; Gourevitch, 1970), monkey (Gourevitch, 1970), dolphin (Au and Moore, 1990) and even goldfish (Fay, 1974). The critical ratio has been favoured in many studies with animals, due to it requiring much fewer measurements therefore much quicker and easier implementation. Such studies have been conducted

in the cat (Watson, 1963; Pickles, 1975; Nienhuys and Clark, 1979), chinchilla (Miller, 1964; Seaton and Trahiotis, 1975; Niemiec et al., 1992), mouse (Ehret, 1975), macaque (Gourevitch, 1970), parakeet (Dooling and Saunders, 1975), dolphin (Johnson, 1971; Au and Moore, 1990) and even beluga whale (Johnson et al., 1989). The results of some of these can be seen plotted in Figure 1.5, compared to a human estimate from (Green et al., 1959). This demonstrates the general trend of animals having broader bandwidth estimates than humans.

Tone masking experiments have been conducted in the gerbil (Burkey and Gans, 1991), parakeet (Saunders et al., 1978b) (also using a narrowband masker (Saunders et al., 1978a)) and budgie (Dooling and Searcy, 1985). Each of these only used simultaneous masking, however. One study in the patas monkey (Smith et al., 1987) did use a forward tone masker but did not measure the same using simultaneous masking, so no comparison between the two is possible. Three studies did compare forward with simultaneous tone masking in the chinchilla (McGee et al., 1976), macaque (Serafin et al., 1982) and parakeet (Kuhn and Saunders, 1980). The first did not find a difference between the two stimulus arrangements but in the latter two, forward masked bandwidths were narrower than simultaneous. This matches up with the findings of human psychophysics. In Saunders et al. (1979) band-widening, critical ratio and simultaneous tone masking were all used to measure bandwidth in the parakeet, and the results were very similar across the board. Even more remarkably the same experiments were conducted in humans and the results were similar between method and the two species. They concluded therefore that the mechanisms governing frequency selectivity are the same in the auditory system of both species.

All of these methods have their problems, which have been described in detail in the previous sections. There have also been a few notched noise experiments conducted in animals. In Niemiec et al. (1992) forward masked notched noise was compared to critical ratio, band-narrowing (band-widening but in reverse) and rippled noise in the chinchilla. They found that the critical ratio gave much broader AF bandwidths than human ones (Glasberg and Moore, 1990), and the estimates got progressively narrower for band-narrowing and rippled noise. Notched noise yielded the narrowest bandwidths, which were similar to results in humans. In Halpern and Dallos (1986) notched noise was used to investigate frequency selectivity in the chinchilla. They even used both forward and simultaneous masking. The results, however, were inconclusive since their detection threshold versus notch width plots were non-monotonic (resulting in an AF with a 'dip' in the centre), which is almost unheard of. They did find, though, that

bandwidth widened with increasing level of masker (the masker level was fixed). One of the most relevant studies, for this thesis, is Evans et al. (1992). Evans managed to train guinea pigs to perform a behavioural notched noise task and used this (as well as rippled noise) to estimate filter bandwidths for several frequencies. The relationship between signal frequency and bandwidth from this study can be seen plotted in Figure 1.6. Another set of (currently unpublished) behavioural bandwidth measures, that are relevant for a later chapter in this thesis, are plotted against signal frequency in Figure 3.9 (filled red circles). They are of the ferret, measured using a continuous (and therefore simultaneous) notched noise masker, made by other researchers at the Institute (Sollini and Alves-Pinto, unpublished).

All three of the notched noise studies just mentioned used simultaneous masking with fixed masker, which is not regarded as the best method to use in human psychophysics. Behavioural research in animals lags behind that of human studies, in no small measure due to how much more difficult and time consuming it is. Also results are much more variable. For further detailed study of frequency selectivity it would be useful to apply an up-to-date method, such as that described in Oxenham and Shera (2003). Here symmetric and asymmetric notched noise maskers are used in both forward and simultaneous arrangements, with signal level kept fixed. Also roex functions are fitted to the resulting detection threshold plots.

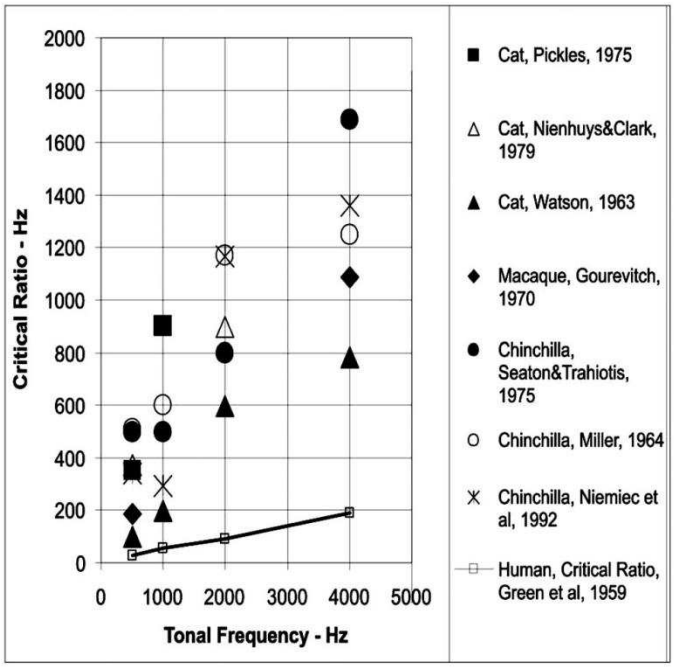


Figure 1.5

Critical ratio plotted against tone frequency for different mammalian species. Figure taken from Yost and Shofner (2009) (adapted from Fay (1988)).

Overall the aim of psychophysical research into frequency selectivity is to measure the ability of the entire auditory system of an organism. Masking experiments have been developed and refined over the decades to effectively measure AFs; the concept through which frequency selectivity is described. Because of the linear, one-to-one relationship between masker level and detection threshold (at least for simultaneous masking; roughly two-to-one for forward masking), these psychophysical bandwidths have widely been regarded to reflect cochlear tuning. However, to investigate whether this is true, and see if this bandwidth representation is maintained throughout the auditory system, a physiological investigation is required.

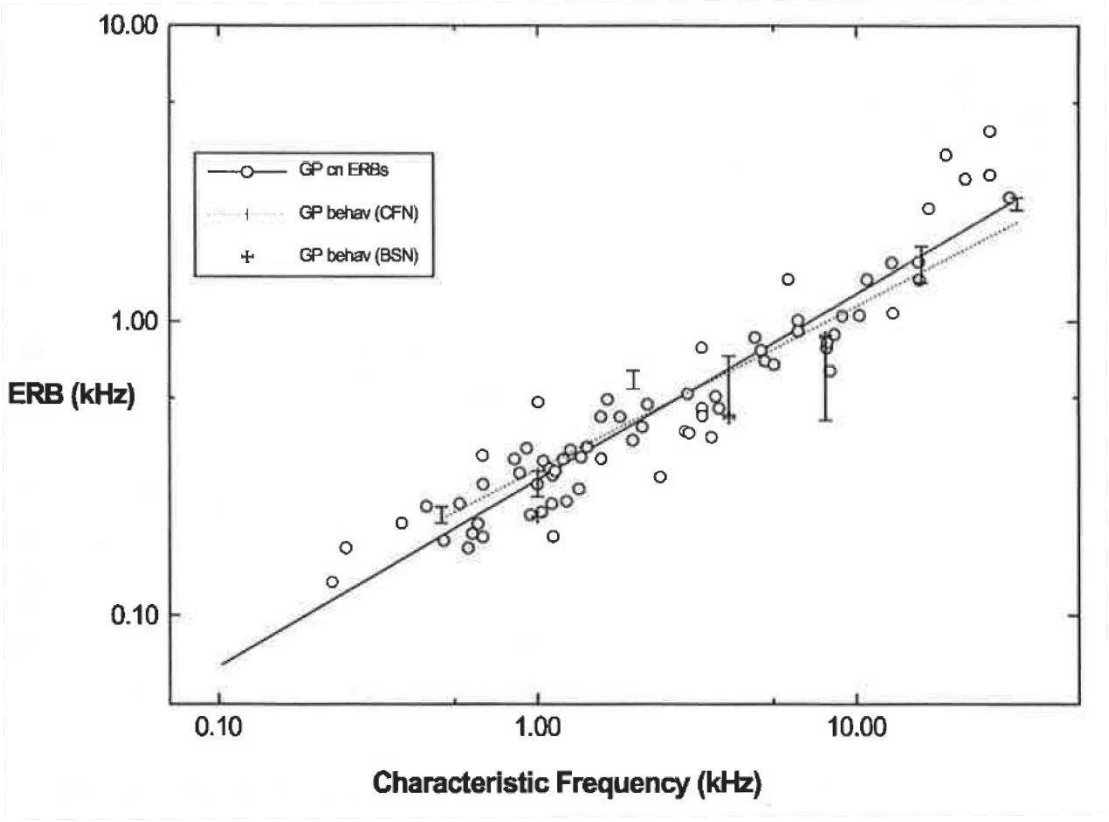


Figure 1.6

CF vs ERB for guinea pig data. Open circles denote AN bandwidth estimates. Error bars and crosses denote psychophysical results from rippled noise (CFN; comb-filtered noise) and notched noise (BSN; band stop noise) respectively. Dotted line shows line of best fit for both psychophysical sets of points. Taken from Evans (2001).

1.8 Physiology

Another major branch of auditory neuroscience is physiology, in which neural activity evoked by sound is investigated. More specifically the activity of a neuron (or population of neurons) in response to auditory stimuli is explored, in an attempt to learn more about the neuron(s)'s role in auditory perception. This usually involves measuring electrophysiological activity from one or several brain regions that collectively form the auditory system. These extend from the cochlea in the periphery up to cortex. Each area has its own neuronal architecture and connections to other areas, which help it to perform specific functions in processing the incoming signal. Further on I will be discussing experiments in different auditory structures and so to provide some context, here is a brief description of the auditory system.

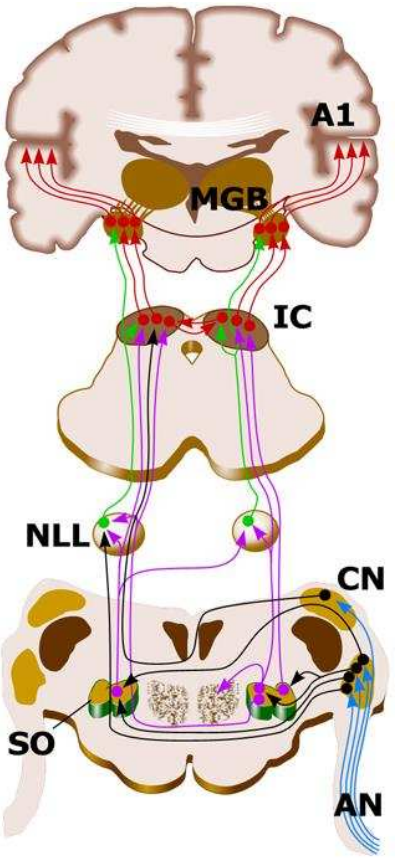


Figure 1.7

The main ascending pathways in the mammalian auditory system. AN is the auditory nerve coming directly from the cochlea. CN is the cochlea nucleus, with ventral and dorsal regions denoted separately. SO is the superior olivary complex. NLL the nuclei of the lateral lemniscus. IC is the two nuclei of the inferior colliculus. MGB is the medial geniculate body in the thalamus. A1 is primary auditory cortex. Image taken from http://www.urmc.rochester.edu/labs/davis-lab/projects/functional_pathways_in_the_auditory_system.

1.8.1 The central auditory system

Figure 1.7 shows the nuclei and the main ascending connections in the auditory pathway. The auditory system is unique among many sensory modalities in that there are several stages between the peripheral sense organ and the fore brain. Significant pre-processing occurs here before a sound can be identified and interpreted by higher brain regions. Figure 1.7 only shows some of the connections between the nuclei. There are many more on the ascending pathway and substantial efferent connections from cortex back down to allow feedback and higher level control. These efferent connections even go down as far as the OHCs in the cochlea allowing active manipulation of cochlear mechanics. However, in this study only the afferent pathway up to primary auditory cortex (A1) is of interest and this is described below. The pathway is actually made up of two interconnected pathways, one on each side of the body, with duplicate regions at each stage. The following description is only of one half.

In the cochlea the mechanical energy of sound is first converted into electrical energy in the hair cells. From here the first connection up the auditory pathway is the auditory nerve (AN). This is a thick bundle of neurons that relay the neural activity from the cochlea to the cochlea nucleus (CN). It consists of high and low spontaneously firing cells that carry different types of information about sound. Many physiological experiments consider the AN response to fully represent cochlear activity. The next stage is the CN, which is actually split into dorsal (DCN) and ventral (VCN) regions. Here orderly and abundant input from AN creates several topographic maps of frequency equivalent to cochlear activity. The VCN is the more primary-like region, thought to mainly act as a relay for ascending information. DCN has extensive efferent connections from higher brain regions, as well as complex inhibitory circuits and a high proportion of intrinsic connections, leading to the belief that it is involved with more complex auditory processing. For example, such inhibitory circuits could reject unwanted signals, enhance weak ones, enhance the SNR in a noisy signal or have a disinhibitory role. Next along is the superior olivary complex (SOC). This is split into lateral and medial olives, the former being involved in calculating inter-aural level difference, and the latter inter-aural time difference, both of which are important for sound localisation. Little is known about the role of the lateral lemniscus (LL) but it is present in all mammals and has neurons very similar to collicular ones. Next comes the inferior colliculus (IC). This is one of the largest sub cortical nuclei in the vertebrate brain and almost all ascending neurons terminate here. It also has connections to almost every known auditory brain region, making it the most interconnected auditory structure. This suggests that it plays

a vital role in both the ascending and descending pathways. It is split into 3 regions: central, lateral and dorsal nuclei. The central nucleus is exclusively auditory and contains several laminae of tonotopically organised, non-frequency-overlapping neurons. It is essential for normal hearing and is the most primary-like region of the IC. The lateral nucleus is the target of considerable non-auditory input, leading to the idea that it is involved with multisensory integration. The dorsal region has the most projections from the cerebral cortex and its role is unknown. The medial geniculate body (MGB) is the auditory portion of the thalamus and is split into several sub-nuclei by neuronal architecture and connections. Its role is multi-facetted and complex but has mainly ascending projections. The final stage is auditory cortex, the most afferently innervated region being A1; from here auditory activity propagates throughout cortex. Its structure (like all cerebral cortex) is laminar, with complex interconnections both laterally and longitudinally. It is very complex and not well understood, but is thought to deal with high level processing. For a more detailed description of all the auditory brain regions just described, see Schreiner and Winer (2005).

Each of these areas has been explored and recorded from using a variety of methods, teaching us much about the inner workings of the auditory system. Bandwidth measures of neurons have also been made in an attempt to see how they fit together to manifest the overall ability of the system. Before I go into too much detail on this, it is important to explain a few commonly used methods that help to characterise multiple aspects of neuronal activity (and which have formed the cornerstone of many fields of neurophysiological research, not just frequency selectivity).

1.8.2 Frequency response area

The first successful recording of auditory neurons was in the AN of the guinea pig in 1954 (Tasaki, 1954). Much of the study looked at the shapes of action potentials and firing properties in response to pure tones. It was also the first time a frequency response area (FRA), also known as an excitatory frequency tuning curve (eFTC), was recorded. An example, from a single neuron recorded in this study, can be seen in Figure 1.8 panel A. Such a plot is created by recording the firing rate of a neuron (or population of neurons) in response to a pure tone of various frequencies and levels. The frequency of the tone is plotted on the horizontal axis (in Hz or kHz) and the level on the vertical (in dB SPL). Firing rate is represented by colour. It illustrates the following key characteristics of a neuron:

Threshold: The lowest level of signal that still elicits an increased firing rate. This is simply the level of the tip of the FRA (× in Figure 1.8 panel A).

Characteristic frequency (CF): The CF is the frequency that still elicits a response at threshold. Simply the frequency of the tip of the FRA.

Best frequency (BF): The frequency that elicits the maximum firing rate.

Bandwidth: This is measured in a number of different ways, one of which is width of the FRA at a certain level above threshold, e.g. 10 dB in Figure 1.8 panel A. This is described in more detail below.

Most neurons are tuned to a certain frequency, resulting from the main innervations they receive from the cochlea, but they also respond to others. This is especially true for higher sound level stimuli. The conditions under which a neuron fires reveals a lot about its role, and certain types of cells have distinct FRA shape that allow them to be identified. Figure 1.8 panel A shows the classic 'V' shape of a primary-like neuron, but more complex neurons have much more complicated FRA shape.

Bandwidth is most commonly measured from an FRA by calculating the Q_{10}. This is described be the following simple formula:

$$Q_{10} = \frac{BW_{10}}{CF} \quad \text{where} \quad BW_{10} = width\ of\ FRA\ 10\ dB\ above\ threshold. \tag{1.4}$$

Variants of this also exist as Q_n where n describes the amount of dB above threshold the width of the FRA is measured at. Another measure, more common in psychophysics, is the equivalent rectangular bandwidth (ERB). This equates to finding the width of the rectangular filter that passes the same amount of energy, and has the same gain as the AF. For the FRA this is described by

$$ERB = \frac{Area\ of\ FRA}{Threshold}. \tag{1.5}$$

Since the area of the FRA is in units of level × frequency, and threshold is in level, ERB is a frequency value. Q_{ERB} is also sometimes seen, which is simply ERB divided by CF, but this is fairly rare.

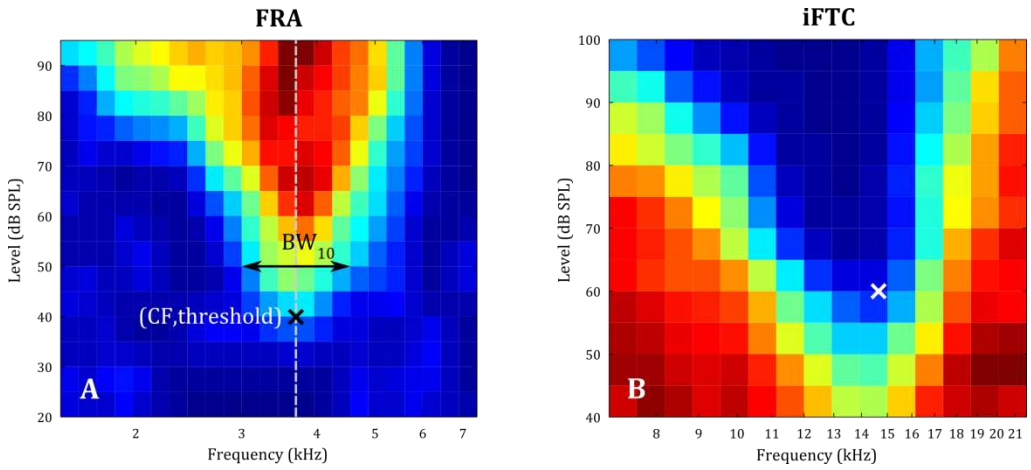


Figure 1.8

Example FRA and iFTC. ***A*** *Plot of firing rate in response to a tone of various frequencies and levels; blue = low, red = high. × marks the point representing CF and threshold. BW_{10} is the width of the FRA 10 dB above threshold. Grey dashed line indicates RLF.* ***B*** *Firing rate in response to tone masker with various frequencies and levels. White × marks the frequency and level of the signal tone. Blue region indicates masking.*

1.8.3 Rate against level function

Another useful plot of neural activity that is used to characterise certain properties of a neuron, is the rate against level function (RLF). This is a plot of firing rate against level of a single frequency stimulus, and is actually a subset of the FRA, indicated by the dashed line in Figure 1.8. The RLF is usually made using the CF of the unit. An example can be seen in Figure 2.6 panel B. Key features include threshold, saturation (maximum firing rate), slope, dynamic range (range of levels between threshold and saturation) and the shape (monotonic/non-monotonic).

1.8.4 Tone-on-tone masking

This is similar to the tone masking described in 1.3.4 Masking with pure tones, whereby a pure tone masker of various frequencies and levels is played in conjunction with a pure tone signal of fixed frequency and level (usually CF, just above threshold). The firing rate in response to the signal is measured and plotted against masker frequency and level, and is often referred to as an inhibitory frequency tuning curve (iFTC). An example can be seen in Figure 1.8 panel B. This method allows the exploration of masking by inhibition, since any reduction in response to the signal caused by the masker is masking. In some cases the masker can actually cause a higher firing rate in response to the signal; this is known as facilitation. Important features are the same as

in FRA (i.e. threshold, CF, BF, bandwidth, shape) except here decrease in firing rate is the area of interest.

1.8.5 Central non-linearities

As mentioned previously one of the key discoveries, which significantly shaped auditory research, is that with simultaneous masking, a 1dB increase in masker level results in a 1 dB increase in detection threshold of a signal (Hawkins Jr and Stevens, 1950). This implies that the auditory system is linear, however we know that cochlear mechanics are highly non-linear, and the fact that forward masking exists proves that the system can't be linear. So what exactly is happening in the central auditory system? The answer is a number of non-linear processes that manifest in overall linear behaviour.

A highly useful tool in exploring these non-linearities is the RLF of a neuron. In many neurons it is a monotonic, sigmoid shape with a surprisingly narrow dynamic range over which firing rate varies. In the AN, RLFs are monotonic, but have regions where firing rates increase rapidly, and regions where they increase more slowly. This is thought to reflect the linear (rapid) and compressive (slowly increasing) regions of the BM I/O function (see 1.2 The basilar membrane and its non-linearities*)*. In central neurons, no link has been made of RLF shape with cochlear non-linearity. However, it is common to observe strong non-linearities such as non-monotonicity, which are often thought to reflect inhibition. Therefore RLFs across the auditory system appear to be hallmarks of many different sources of non-linearity.

One such non-linearity is the auditory system's solution to the 'dynamic range problem' with encoding the level of a stimulus (Viemeister, 1988), caused by the surprisingly small dynamic range of neurons mentioned above. Neurons use neurotransmitters and the movement of different ions, e.g. Na^+ and K^+, to produce and propagate action potentials; these molecules/ions become depleted during firing and must be replenished. The maximum rate at which this can occur is limited by the biology of the neuron, and the range over which it can accurately change its firing rate is known as the dynamic range. The majority of ANFs have low thresholds and saturate quickly, giving a dynamic range of just 35 dB (Sachs and Abbas, 1974; Evans and Palmer, 1980). Other neurons with slightly higher thresholds tend to have larger dynamic ranges, since most of the input sound levels they respond to are in the compressive region of the BM I/O function, meaning the dynamic range of the inputs from their OHCs is already wider. However, even with these characteristics they can still not account for the full ~120 dB range of human hearing through firing rate alone. To get around this neurons do not

encode rate in a fixed firing rate, but their firing rate adapts to the stimulus. Dean et al. (2005) recorded RLF from neurons in the guinea pig IC in response to broadband noise of different levels. Level was varied randomly with the large majority (80 %) coming from a narrow (12 dB) range of levels. This was repeated for several different ranges of level from 33 - 81 dB SPL. The RLFs show a clear reduction in the maximum firing rate, and the dynamic range shifts towards higher levels, with increasing level of input. This shows that the neurons are adapting their firing rates to the most common input level in a highly non-linear manner.

1.8.6 Shifts in RLFs

A very similar phenomenon occurs in tone-in-noise masking experiments, that has revealed much about the origins of masking. A pure tone signal (set at the CF of a neuron) is simultaneously masked by a broadband noise, and an RLF is recorded. This is repeated for several different levels of masker, and the effect on RLF observed. In a completely linear system one would expect to get 'excitatory masking', in which the response of the neuron in the signal + masker condition is simply the sum of the neuron's response to the signal and masker individually. This would result in the floor of the RLF rising (the magnitude dependent on the level of the masker) but all other features remaining the same. An example can be seen in Figure 1.9 *(a)*. However, if there is non-linear masking, the dynamic range of the RLF might shift to a higher SNR, and the amount of shift is dependent on the level of the masker. An example can be seen in Figure 1.9 *(b)*. (In the figure, taken from Delgutte (1989), it is referred to as 'suppressive' masking, although such a shift can also be caused by other non-linear mechanisms. Also theoretically there could be a shift to a lower SNR, or a change in shape, depending on the nature of the non-linearity). In fact neurons throughout the auditory system display both kinds of masking, with most excitatory masking occurring in the AN (Delgutte, 1989), and being almost completely absent in A1. Since this shift in RLF is present at every stage of the auditory system, it has been used as an indicator of the amount of masking that might be occurring at each stage.

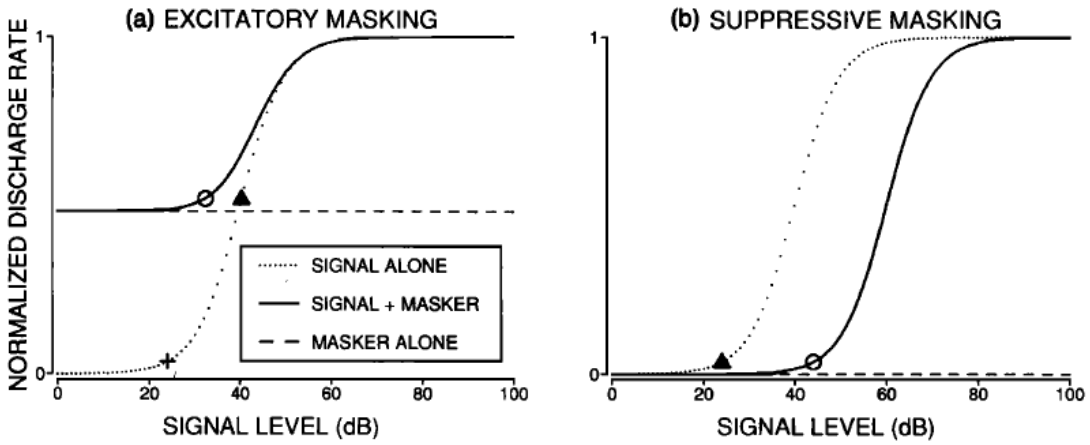


Figure 1.9

Schematic of the effect of excitatory and suppressive masking on a neuron's RLF (taken from Delgutte (1989)).

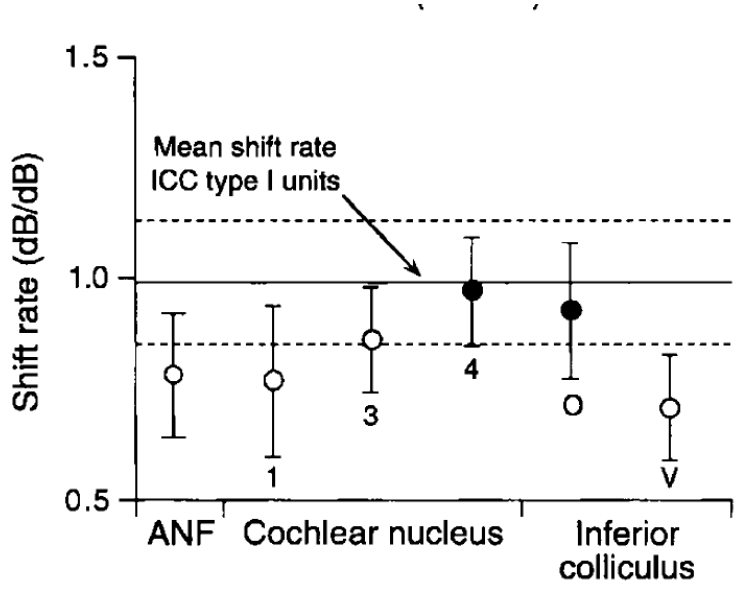


Figure 1.10

Growth of simultaneous masking in different brain areas. The RLF of neurons in a tone-in-noise experiment shift roughly linearly with increasing masker level. This plot show the amount it shifts per 1 dB increase in masker level. Labels 1, 3 and 4 relate to type I, III and IV neurons in CN. Solid line refers to results for IC neurons with an I shaped FRA, dashed lines are ± 1 standard deviation. O and V refer to neurons with that shaped FRA. Taken from Ramachandran et al. (2000).

Studies in the AN of the cat, looking at the response to a pure tone in the presence of a simultaneous broadband masker, showed a shift in RLF of 0.61 dB (Costalupes et al., 1984) and 0.79 dB (Gibson et al., 1985) per 1 dB increase in masker level. Gibson et al. (1985) also recorded from neurons in the CN and found that the shift in VCN and DCN type I and III units (more primary) is 0.86 dB/dB, and 1.04 dB/dB for DCN type IV (more complex). In the IC of the guinea pig Rees and Palmer (1988) found the shift to be 0.97 dB/dB, and in the decerebrate cat Ramachandran et al. (2000) found the shift in I type IC neurons to be 1.04 dB/dB. Figure 1.10 shows a plot of the results from (Ramachandran et al., 2000). Open symbols show neurons with solely excitatory input, whereas the solid line and filled symbols show neurons with both excitatory and inhibitory input. In A1 RLF was found to shift linearly with masker level (i.e. 1 dB/1 dB) as seen in psychophysics (Phillips and Cynader, 1985; Phillips, 1990; Ehret and Schreiner, 2000). These results show a clear trend that the effect of masking stimulus on RLF develops up the auditory pathway, reaching its maximum as early as the IC. From this point onwards the rate of shift in RLF can fully account for the change in masking seen psychophysically, i.e. the 1dB/dB change in threshold with masker level. Also, inhibitory inputs seem to be a major source contributing to masking.

This analysis has also been conducted using non-simultaneous masking (a mixture of both forward and backward masking). In Costalupes et al. (1984) a non-simultaneous masker was also used to measure the RLF of cat ANFs, and was found to shift at a markedly smaller rate than with the simultaneous masker. This was also seen in guinea pig IC neurons in Rees and Palmer (1988), where RLF with a simultaneous masker shifted at 0.86 dB/dB and only 0.31 dB/dB with a non-simultaneous masker. (Rates were calculated by fitting straight lines to the data points above 0 dB noise spectrum level in Figure 9 (a) of the paper). In Gibson et al. (1985) a method similar to non-simultaneous masking was used, in which a pure tone was played continuously at a fixed level, which was abruptly reduced for a period of 200 ms; they called it an 'interrupted tone'. The firing rate in this lower level region was then taken, and the process repeated for various continuous tone levels, producing an RLF. This is equivalent to having a fixed level tone signal, with a varying level non-simultaneous tone masker of the same frequency. They found that the rate of shift in RLF was 0.33 dB/dB in AN, 0.37 dB/dB in VCN, and 0.47 dB/dB in DCN. These together suggest that the difference between the two forms of masking is caused by suppression on the BM, since the behaviour of the two centrally is very similar, and that this is still evident at least to the level of the IC.

Shifts in RLF are useful indicators for masking, but they only describe the suppressive type and not excitatory (see Figure 1.9). To see the full amount of masking at different stages of the auditory system it is necessary to use a signal detection theory (SDT) approach (see 1.8.11 Signal detection theory section) (Young and Barta, 1986). Such studies have been conducted, looking at the growth of forward masking in the AN using SDT (Relkin and Pelli, 1987). Relkin and Turner (1988) found that detection threshold of a pure tone signal increased at 0.1 – 0.5 dB per 1 dB increase in masker level, whereas the psychophysical rates of Jesteadt et al. (1982) and Plack and Oxenham (1998) were 0.6 dB and > 0.2 dB respectively. This potentially tenuous relationship was corroborated in Turner et al. (1994), in which a direct comparison between psychophysical and physiological (in AN) growth of forward masking showed the latter to be 64 % or 36 % of the former (depending on envelope on/off ramps). These studies clearly indicate that forward masking is not just manifested in the periphery and that central mechanisms contribute. By extension, the similarity in behaviour between the shifts in simultaneously and non-simultaneously masked RLFs as one moves up the auditory system, may suggest that central mechanisms also contribute to simultaneous masking.

So what exactly are these central mechanisms and how is forward masking manifested? Many physiological experiments have been conducted to measure forward masking, and help answer this question.

1.8.7 Physiological characteristics of forward masking

In Harris and Dallos (1979) the activity of ANFs in the guinea pig was measured in response to a tone-on-tone forward masking experiment (see 1.8.4 Tone-on-tone masking). As with psychophysical studies, there was a logarithmic relationship between masking and time between masker offset and signal onset (1.6 Forward versus simultaneous masking). This was independent of CF. They also found that in the same ANFs the iFTC (with just above threshold, CF signal) was very similar to its FRA. A similar finding was made in Brosch and Schreiner (1997) looking at forward masking in A1; tone inhibition persisted for 53 – 430 ms with iFTC matching up closely with FRA, reducing in size with increased distance between masker and signal. Also the amount of masking was dependent on the level of the masker. Indeed Nelson et al. (2009) recorded from IC neurons in the awake macaque and found that some key features of growth of forward masking (maximum amount of masking, and recovery from masking) were already seen at the level of the IC. Additionally, forward masking already had a wide dynamic range at this level. So individual neurons display the same characteristics of

forward masking as seen psychophysically, and at various stages of the auditory system. In addition masking without the inclusion of suppression on the BM (i.e. non-simultaneous), seems to reflect the excitatory characteristics of individual neurons. So what kind of mechanism could produce this forward masking so early on in the auditory system? There are two main theories, which are described in the section below.

1.8.8 Theories behind psychophysical forward masking

One theory is that psychophysical forward masking is caused by neural adaptation (Smith, 1977; Kidd and Feth, 1981; Jesteadt et al., 1982; Bacon, 1996; Nelson and Swain, 1996), potentially already in ANF (Duifhuis, 1973; Smith, 1979). Adaptation is the phenomenon whereby the response of neurons to a constant stimulus diminishes over time (Kiang, 1965), and can develop over several seconds or even minutes (Javel, 1996). It is thought to be due to intrinsic mechanisms within neurons, synaptic depression (where neurotransmitters become depleted over time), or interactions between neurons (e.g. inhibitory feedback loops). Adaptation in the ANF cannot be the sole cause since it does not account for the full amount of forward masking seen psychophysically (as described in the previous section). Also cochlear implant patients show similar amounts of forward masking as normal hearing listeners, which means it cannot be solely manifested in the periphery (Shannon, 1990).

Another theory is that it is caused by neural activity that persists in the central auditory system after the stimuli have ceased (Plomp, 1964; Penner, 1975; Zwicker, 1984; Moore et al., 1988; Oxenham and Moore, 1994). This is more commonly referred to as the 'temporal window model' because it can be modelled by a BM non-linearity followed by integrating over a time invariant, linear temporal window (Festen and Plomp, 1983; Moore et al., 1988; Plack and Moore, 1990; Oxenham and Moore, 1994; Plack and Oxenham, 1998). In this model the only effective non-linearity is that in the periphery, and overall the auditory system acts linearly after that, at least in growth of masking. This model has the added benefit of also describing simultaneous masking, encapsulating the differences between the two temporal arrangements, such as amount and growth of masking.

Which of these is the mechanism behind forward masking is still debated. Nelson et al. (2009) claims that adaptation and offset inhibition can completely account for forward masking instead of the persistent excitation of the temporal window model. In Oxenham (2001) a temporal window model was shown to fit more closely to psychophysical data

of growth of forward masking than adaptation, although they did not discount the possibility of adaptation being the cause.

1.8.9 Mechanisms behind masking

It is clear that masking isn't caused by one mechanism, but is instead the cumulative effect of several. Suppression on the BM makes a sizeable contribution and is believed to fully represent the difference between forward and simultaneous masking. The amount of masking increases throughout the central auditory system and is most likely caused by neural adaptation, or persistent excitatory activity. Factors contributing to this include network interactions, synaptic depletion and intrinsic mechanisms in neurons. Another contributing factor is the activity of the medial olivocochlear system (MOC) (Guinan Jr, 1996, 2006). This system includes efferent connections from the SOC to the OHCs in the cochlea, which allows direct central control over cochlear gain. It has a time constant of roughly 100 ms and turns down the automatic gain on the BM.

Such mechanisms are thought to aid many fundamental auditory functions such as noise cancellation, helping pick out useful auditory information from background noise (Winslow and Sachs, 1988; Kawase et al., 1993). There is evidence that this ability increases up the auditory pathway (as does masking) providing a noise free representation of salient sounds in A1 (Rabinowitz et al., 2013). Also similar non-linear adaptive processes seen in the visual system have been shown to improve the efficiency of neural coding (Brenner et al., 2000; Fairhall et al., 2001), so this may have the same effect in the auditory system.

Regardless of the mechanisms behind it, masking has been a very useful tool in exploring frequency selectivity. The following section describes the various methods (including masking) that have been used to explore frequency selectivity physiologically, and what they have taught us.

1.8.10 Physiological bandwidth measurements

FRA

Due to its relative simplicity leading to quick and easy implementation, the FRA has been the most popular way to describe frequency selectivity in neurons. It has been used to measure bandwidth in the AN of many species, such as the cat (Kiang, 1965; Wilson and Evans, 1971; Liberman, 1978), squirrel monkey (Geisler et al., 1974), gerbil (Ohlemiller and Echteler, 1990), mouse (Taberner and Liberman, 2005), pigeon (Sachs and Abbas, 1974), starling (Manley et al., 1985), guinea pig (Evans, 1972; Tsuji and

Liberman, 1997; Evans, 2001) and ferret (Sumner and Palmer, 2012). Many of these can be seen in Figure 1.11. The results across mammalian species are all very similar, showing a roughly linear (for log CF) increase in bandwidth with CF of similar rate. All these studies were conducted in anaesthetised animals since it is still very difficult to access the AN with chronic implants. The CN, too, has been extensively studied in this manner since its various types of cells have very distinct response characteristics which allow them to be identified without histological verification. Probably the most relevant study in the CN to the study in this thesis is Sayles and Winter (2010), which investigates bandwidth of single-units (SU) in the guinea pig. This will be explored in more detail later.

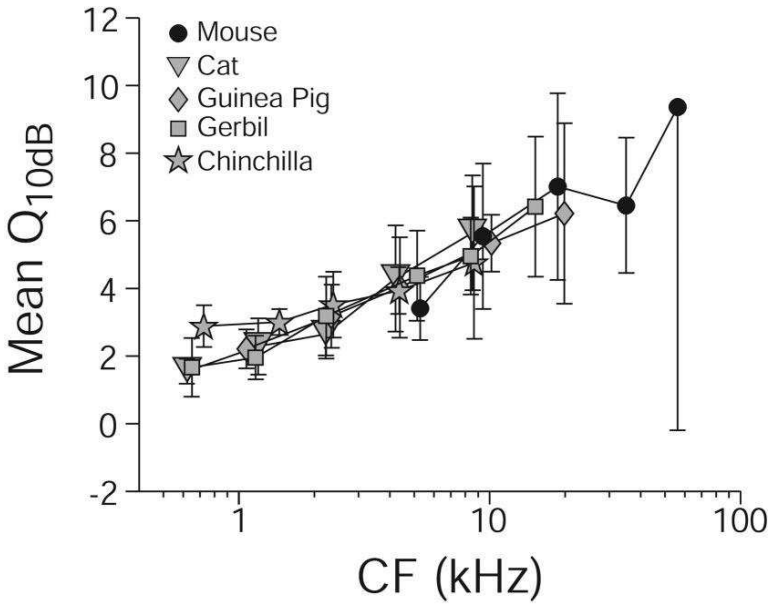


Figure 1.11

Physiological bandwidth estimates in the AN of 5 mammalian species, made using FRA. Data are from Taberner and Liberman (2005) for mouse, Liberman (1978) for cat, Tsuji and Liberman (1997) for guinea pig, Liberman (unpublished data) for chinchilla, and Ohlemiller and Echteler (1990) for gerbil. Figure taken from Taberner and Liberman (2005).

Much of what we know about the IC has come as a result of recording FRAs. Such experiments have been conducted in a range of species such as the cat (Rose et al., 1963; Gersuni et al., 1971; Langner and Schreiner, 1988; Schreiner and Langner, 1997), mouse (Ehret and Moffat, 1985; Egorova et al., 2001; Ehret et al., 2003), bat (Wenstrup et al., 1986; Casseday and Covey, 1992), chinchilla (Wang et al., 1996), rat (Hernandez et al., 2005) and guinea pig (LeBeau et al., 2001; Palmer et al., 2013). Schreiner and Langner (1997), showed us that the IC consists of discrete laminae of neurons arranged topographically in frequency and also other characteristics such as best modulation frequency, onset latency, bandwidth and binaural interaction type (Langner and Schreiner, 1988; Egorova and Akimov, 2013). Hernandez et al. (2005) investigated the bandwidth of IC neurons in the rat, looking at Q_{10}, and compared this to a whole host of other brain regions and mammalian species, namely the mouse (Egorova et al., 2001), guinea pig (Syka et al., 2000), cat (Ramachandran et al., 1999) and albino rat (Kelly et al., 1991). These can be seen in Figure 1.12. They show that the IC displays similar frequency selectivity across species (except albino rat) and that this is similar to the AN (Møller, 1978; Carlier and Pujol, 1982) and other brain regions in the rat. The growth of Q_{10} with CF is not as linear here as for the AN data in Figure 1.11, growing much more rapidly at higher CF. This trend may, however, only be true for the rat since it is mainly limited to these plots. Palmer et al. (2013) is possibly the most comprehensive look at FRA characteristics of guinea pig IC neurons. They compared responses to over 2000 units from a number of anaesthetic and experiment types, and even calculated bandwidth with both the Q_{10} and ERB for each. They found that FRA shape in IC can generally be split into a number of different categories, which can help to identify the role of a cell; V-shaped, non-monotonic Vs, narrow, closed, tilt down, tilt up and double-peaked. V shaped are the most primary like cells receiving only excitatory input, all other shapes are the result of complex excitatory and inhibitory inputs.

FRAs recorded in A1 units, however, have shown much less homogeneous characteristics. Sutter (2000) showed that there is a huge variety in FRA shape, such that they should be thought of more as a continuum than fitting into discrete groups. FRA bandwidth investigations in A1 have also shown mixed, and at times conflicting, results. Suga (1977) showed that the bat A1 has a strong topographic arrangement of frequency and amplitude. Sally and Kelly (1988) and Phillips and Irvine (1981) each investigated the Q_{10} of A1 neurons in the albino rat and cat respectively (see Figure 1.12). They both found that Q_{10} increased with frequency, in a very similar manner to experiments in lower brain regions described above.

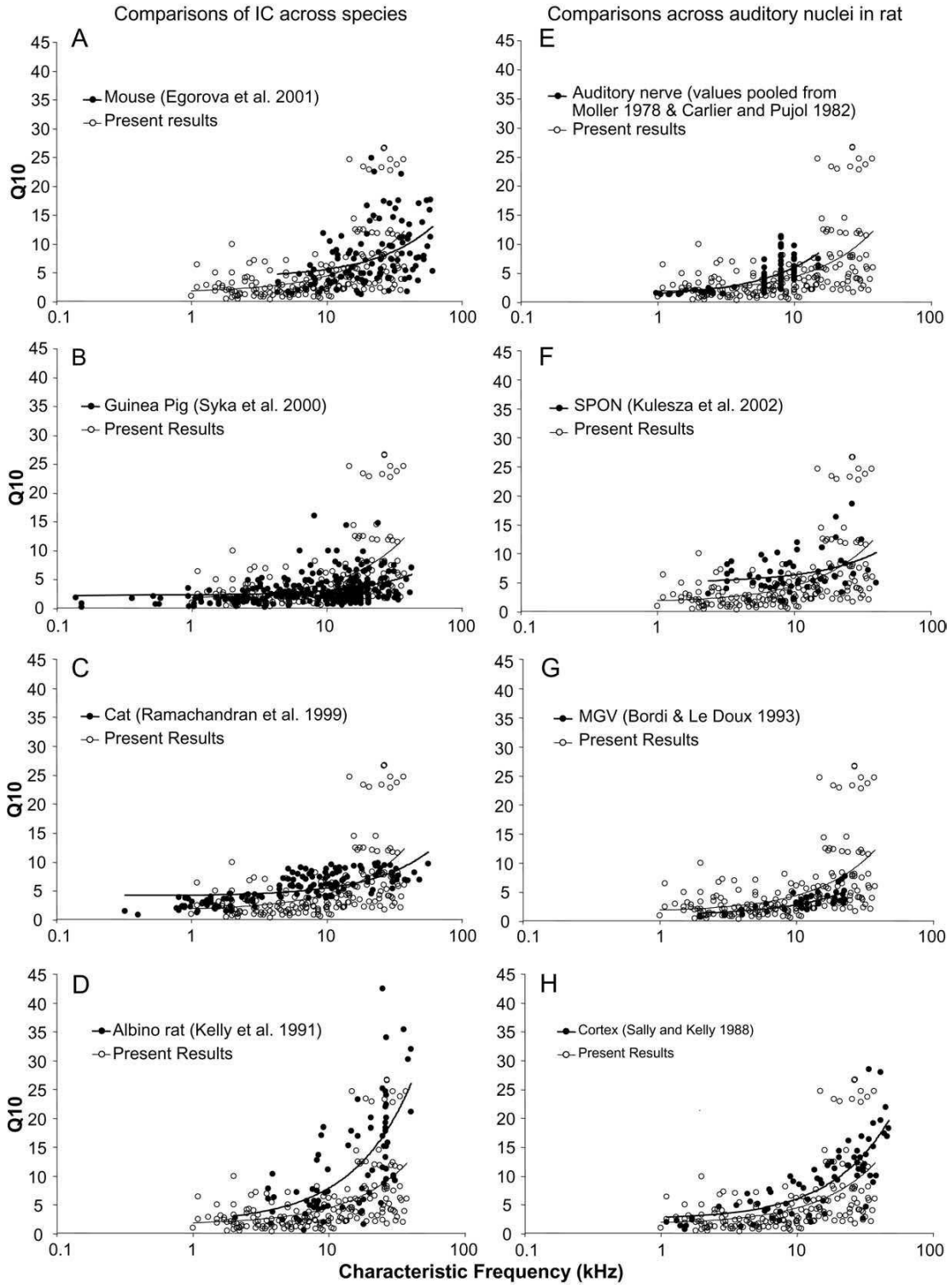


Figure 1.12

Physiological bandwidth measures of several mammalian species in the IC and other brain areas, made using FRA. 'Present results' are for the rat. Figure taken from Hernandez (2005).

Several studies have found tonotopicity in the A1 of the anaesthetised guinea pig (Wallace et al., 2000), cat (Phillips, 1985; Phillips et al., 1994), and awake marmoset (Bendor and Wang, 2008). However, Schreiner and Mendelson (1990) and Schreiner (1995) seemed to show that no tonotopicity exists in the cat A1, instead there is a non-uniform spatial distribution of receptive fields. They suggested that cortex acts like a 'self-organising mapping algorithm' and that there is no inherent frequency or bandwidth structure. Another experiment in A1 of the macaque (Recanzone et al., 2000) showed that there was no increase in Q_{10} with frequency, although this was in an awake behaving animals so is not directly comparable. Another study in awake behaving animals (Galván et al., 2001) used the local field potential (LFP) to measure bandwidth in guinea pig A1. They found that CF and tuning changed on a daily basis, suggesting that frequency and bandwidth isn't even fixed. Such seemingly contradictory findings show how complex and inexplicable cortex is compared to lower brain regions.

Despite the relative simplicity of collecting FRA making it the most common method for describing physiological frequency selectivity, there is still speculation about what exactly it is measuring, and whether it is comparable to psychophysical methods. The first obvious difference between the two is that one utilises masking mechanisms, involving complex inhibitory and excitatory interactions, both within and between neurons, whereas the other simply uses excitatory input. Mechanisms behind masking are not fully understood but are known to be multifaceted and involve widespread interactions. Also maskers generally contain a much wider range of frequencies than the signal, so in masking experiments a much wider range of the cochlea is activated. Whether both these approaches measure the same feature of frequency selectivity is unknown, and seems somewhat doubtful in light of all the non-linearities observed in the auditory system. One such example is the BM I/O function (see 1.2 The basilar membrane and its non-linearities). Input sound levels over a range of ~20-55 to ~90 dB SPL are known to be compressed in the cochlea, and are represented by a disproportionally narrow range of firing rates. Since FRA are produced using signal levels spanning a wide range of levels, they will certainly be affected by this non-linearity. In masking experiments, however, only the detection threshold of the signal is important. By fixing the level of one stimulus in the centre of either the linear region of the BM I/O function from ~0-20 dB SPL, or the centre of the compressive region (see Figure 1.2), both signal and masker should be in the same portion of the I/O at threshold. Therefore at threshold the relationship between the two should be linear, and therefore this non-linearity should have no effect. Masking experiments have also been

conducted on physiological data to measure bandwidth, and explore the relationship between the pure tone and masking approaches.

Masking

Experiments using tone-on-tone masking (see 1.8.4 Tone-on-tone masking) have been conducted in most auditory regions in several species, such as the cat (Greenwood and Maruyama, 1965; Ramachandran et al., 1999), bat (Suga, 1969; Vater et al., 1992; Lu and Jen, 2001), starling (Nieder and Klump, 1999) and guinea pig (Ehret and Merzenich, 1988; Prijs, 1989). Such experiments have shown that neurons have inhibitory side bands, or inhibitory regions flanking CF, which sharpen their tuning. These are evidence of the non-linear mechanisms throughout the central auditory system whose role it is to sharpen tuning. As well as the band-widening experiments described in the previous paragraph, rippled noise has also been used in an attempt to use one method for both psychophysical and physiological studies (Kate et al., 1974; Bilsen et al., 1975). However, as far as I am aware the notched noise has not been attempted at the SU level.

It has, however, been used to measure bandwidth in animals physiologically at a much broader level. Auditory brainstem responses (ABR) in response to notched noise stimuli have been used to give a measure of global brainstem frequency selectivity (Popov et al., 1997; Gall and Lucas, 2010; Henry and Lucas, 2010b, a; Lina and Lauer, 2013).

There is evidence that measuring bandwidth using excitatory stimuli alone (e.g. FRA) is not equivalent to measuring bandwidth using masking methods. Ehret and Schreiner (1997) is a comprehensive study of bandwidth measured using FRA, band-widening and the critical ratio, in the A1 of the cat. They measured the CR and the CB (using band-widening) with several different spectrum levels, effectively creating 'tuning curves' using these two masking methods. They found that the tuning curves of the masking experiments differed substantially in shape from those measured using the FRA, and concluded that the two methods aren't measuring the same thing. In Egorova and Ehret (2008) FRA and iFTC of guinea pig IC neurons were measured and compared to the CB measured in the same cells. They were shown to be substantially different in shape and bandwidth, showing that tone masking too, is measuring something different to band-widening. In both studies the masking experiments used simultaneous masking. If, in fact, the only non-linearity that affected results was cochlear in origin, then two-tone suppression should cause the only difference between pure tone and masking results. Such suppression has the effect of broadening bandwidth estimates (see 1.3.4 Masking with pure tones) and so masking bandwidths would be systematically broader than

pure tone ones. This is, however, not the case in either study just mentioned, providing more evidence that central non-linearities exist and have an effect on bandwidth estimates.

In Ehret and Schreiner (1997) the neural CBs matched up with ones measured psychophysically using band-widening, from which they concluded that bandwidth does exist at the neural level but that it is independent of FRA. However, the mean firing rate in response to the stimuli was used to determine detection, which does not encompass all aspects of neural firing that may be used to determine psychophysical detectability (Young and Barta, 1986). Therefore unless this relationship is corroborated in a study using SDT, or a similar approach, it cannot be definitively said that the CB, as measured psychophysically, exists at the neural level.

Other methods

There have been several methods used to measure bandwidth physiologically and there is not one 'industry standard'. One popular method is reverse correlation. This involves playing either white noise, random tones or spectrally rippled noise and recording the spiking of the neuron in response. Simply reverse correlating the spike train (sequence of 0's and 1's, where a 1 represents a spike) with the spectro-temporal representation of the sound gives a spectro-temporal response field (STRF). All random parts of the stimulus will cancel out and only those features that trigger spiking will remain. This gives information about the excitatory and inhibitory characteristics of the unit, not only in frequency but time. It can be used to measure bandwidth and has been used in several species (De Boer, 1967; De Boer and Jongkees, 1968; De Boer and Kuyper, 1968; De Boer, 1969; Ruggero, 1973; Miller et al., 2002; Escabí and Read, 2003). It has even been used in humans through intracranial depth electrodes over auditory cortex (Bitterman et al., 2008). This study found bandwidth to be far narrower than AN measurements in other animals and far narrower than expected.

Another method that is emerging is to use the otoacoustic emissions (OAE) of the ear to measure bandwidth in the cochlea. This is the phenomenon whereby the cochlea produces a small audible response to sound stimuli, thought to be a by-product of cochlear amplification caused by the OHCs (Lilaonitkul and Guinan, 2009). It was found that stimulus-frequency otoacoustic emissions (SFOAE) could be used to measure bandwidth since the delay of the OAE after stimulus presentation is related to bandwidth. This provides a non-invasive physiological measure of bandwidth that can be used in humans. Shera et al. (2002) is such a study and they found that frequency

selectivity was much finer in humans than previously suspected. The findings were disputed by Ruggero and Temchin (2005), but have since been corroborated by two further studies; Shera et al. (2010a) and Joris et al. (2011). This technique has even been used to measure bandwidth in the tiger, an animal in which invasive techniques would also be impossible (Bergevin et al., 2012).

1.8.11 Signal detection theory

A fundamental assumption that must be made in all physiological studies, especially those using masking, and those making comparisons with psychophysical thresholds, is what constitutes 'detection of a signal' in a neuron? In behavioural studies the subject indicates whether they heard the signal or not, but with neurons there are spikes or other electrophysiological activity. Increasingly studies look at the precise pattern of spikes and relate this to the stimulus (this is known as temporal coding and is discussed in more detail in 2.2.7 Analysis*:* Different approaches) but most simply look at the mean firing rate of the neuron around the time of signal presentation. This mean firing rate is typically compared to the spontaneous rate of the neuron (mean firing rate in the absence of any stimulus) and some threshold criteria relative to spontaneous rate is used to determine when the two are different enough for the signal to be deemed 'detected'. In the real world, however, animals must make decisions from single events and cannot wait for several repeats to average the response. Therefore detection of a signal must be unambiguous on a single trial. Just using mean firing rates throws away valuable information that could influence detectability, namely variance. Say, for example, a neuron had a very precise firing rate with very little variability in the absence of any stimulus. If there was a small increase in firing rate in response to a sound, then looking at just the firing rate of the neuron after a single trial one could be fairly sure whether the sound was present or not. Figure 1.13 panel A shows an example of such a situation. Now if there was a second neuron with a very variable firing rate in quiet, then the same small increase in response to the sound would make much less of a difference. Looking at just the firing rate of the neuron one could not be very certain if the sound was present or not (see Figure 1.13 panel B). The traditional approach would deem the signal to be just as detectable in each neuron but in reality that isn't true.

An entire field of research is devoted to just this problem, known as signal detection theory (Green and Swets, 1966). This treats responses when a signal is present and responses when a signal is absent as samples from different probability distributions. By finding these distributions and determining the amount of overlap, the detectability of a

signal can be quantified. This approach was developed and used extensively in engineering, as well as psychophysics. (For a more comprehensive description see Macmillan and Creelman (2004)). However its use in physiological neuroscience has been rather limited despite its benefits. Relkin and Pelli (1987) used SDT on neural data to build neurometric functions to determine the detection threshold of a signal. They did this by recording the response of a neuron to a masker and signal, then also to just the masker on its own. This was repeated several times creating a distribution of responses for each case, and SDT used to describe how discriminable the signal was. (The idea was actually borrowed from visual sciences (Tolhurst et al., 1983; Parker and Hawken, 1985; Britten et al., 1996)). Since then it has been used in a number of studies (Young and Barta, 1986; Palmer et al., 2000; Nelson et al., 2009; Alves-Pinto et al., 2010), even specifically for the purpose of comparing psychophysical and physiological data (Stüttgen et al., 2011).

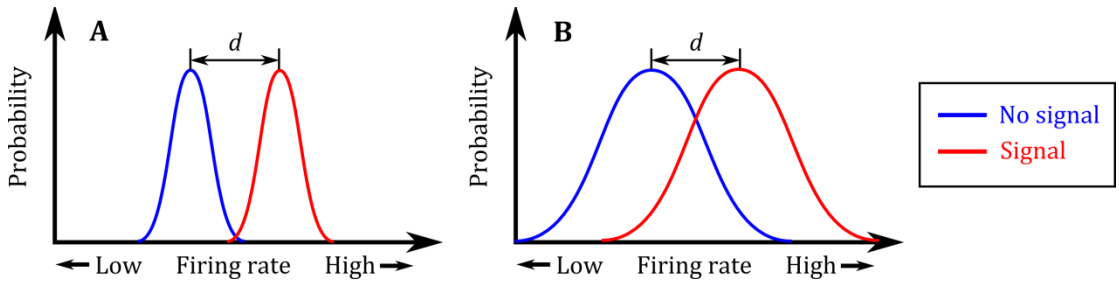


Figure 1.13

Signal detection theory demonstration. ***A*** *Example of a neuron with precise firing, i.e. low variance. No signal (blue) and signal (red) distributions have very small overlap so discrimination between the two from a single trial is good.* ***B*** *Example neuron with erratic firing, i.e. high variance. Difference between mean firing rates of no signal and signal firing rate distributions is the same but there is much more overlap. This makes discrimination more difficult.*

1.9 Comparison of frequency selectivity using psychophysics and physiology

One of the fundamental goals of neuroscience research is to find out how neural activity in the brain leads to perception. How exactly does the firing of neurons result in our sense of the world? Frequency selectivity is no exception, as highlighted by the review King (1998). To begin to answer this question it is first necessary to observe the activity of neurons and the perception they represent, and begin to find relationships between the two. Only then will a causal relationship be investigable. In the world of frequency selectivity this means comparing bandwidth measurements made physiologically to ones made psychophysically, and there have been a number of studies which aimed to make a quantitative comparison of the two in the same species.

Wilson and Evans (1971) compared psychophysically measured bandwidths using rippled noise, with FRA measured bandwidths in the AN, of the cat. Pickles (1975) used simultaneous band-widening to psychophysically measure bandwidth, and also compared this to FRA measured bandwidths in AN, also in the cat. They found that the behavioural CB was broader than AN bandwidth. This finding was corroborated by a further study (Pickles, 1979) behaviourally measuring bandwidth with band-widening and rippled noise, and comparing this to FRA measured bandwidth in the cat AN. They found that the behavioural measures were roughly 3 times broader. These studies, therefore, suggest that band-widening cannot be equated to FRA measured bandwidth, since they are significantly different. Pickles (1979) also measured bandwidth psychophysically with tone masking. He compared this to the FRA measured bandwidths in AN, and found that the results of these two methods matched up fairly closely.

However, due to the highly non-linear nature of the auditory system, slight differences in the way in which measurements are made, lead to significantly different results. It is therefore necessary to use as similar a method as possible when comparing psychophysical and physiological bandwidths. An attempt to do this was made in Ehret and Merzenich (1985) where the CB of cat IC cells was measured using band-widening, and compared to the behavioural CBs of Pickles (1975) and Pickles (1979). They measured bandwidth at several different masker levels and the results can be seen in Figure 1.14. They found that bandwidth measures were level invariant, and did not match up closely to behavioural measures (despite the authors concluding otherwise). The scatter of results is large and there is a possibility that this caused the conclusion of

level invariance, but at least quantitative comparisons using the same method in the same species were being made. Later in Ehret and Schreiner (1997) they found neural CB of A1 cells matched up much more closely with behavioural CB. Another attempt to compare psychophysical and physiological bandwidth measures using the same method was made in Salvi et al. (1982). Here the same pure tone masking experiment was used and bandwidths, using auditory brain stem evoked potentials and behaviour, were measured in the chinchilla. Here the results were found to be very similar, suggesting that frequency selectivity is established at the periphery.

All the behavioural methods just mentioned, however, suffer from problems; tone masking and band-widening both suffer from off-frequency listening, which confounds results, and band-widening suffers from other major problems leading to it being described as 'totally unsuited to measure the critical bandwidth' (Bos and de Boer, 1966). Any true quantitative comparison would need to be made using notched noise masking, the currently accepted 'best practice' for measuring bandwidth psychophysically. One such study (which is highly relevant for the next chapter) is Evans (2001). This study was conducted in the guinea pig, comparing psychophysically measured notched noise and rippled noise bandwidths to FRA measured bandwidths in the AN using pure tones. Evans found a very close match between the two data sets and concluded that frequency selectivity is therefore established at the periphery.

Another study which used notched noise masking, this time on physiological data, was Fishman and Steinschneider (2006). They recorded from multiple neurons in the A1 of the awake macaque using what they called a ‘two noise masking’ approach (very similar to simultaneous notched noise masking except with narrower (50 Hz) noise bands) and calculated bandwidth using roex functions as AF. They compared these results to behavioural ones from the macaque in Gourevitch (1970) (band-widening) and human (Glasberg and Moore, 1990) (notched noise). All measures matched up fairly closely and they concluded that therefore a representation of perceptual frequency resolution is available at the level of A1.

Although these studies both used notched noise masking, they failed to make a comparison between identical methods in the same species, therefore suffering the caveat of any relationship being influenced by non-linearities in the auditory system, and species specific differences. Also, neither of the studies used SDT in the physiological measurements, making them less comparable to the psychophysical

results. A comparison of behavioural and neural data in the same species using notched noise masking would therefore be very informative.

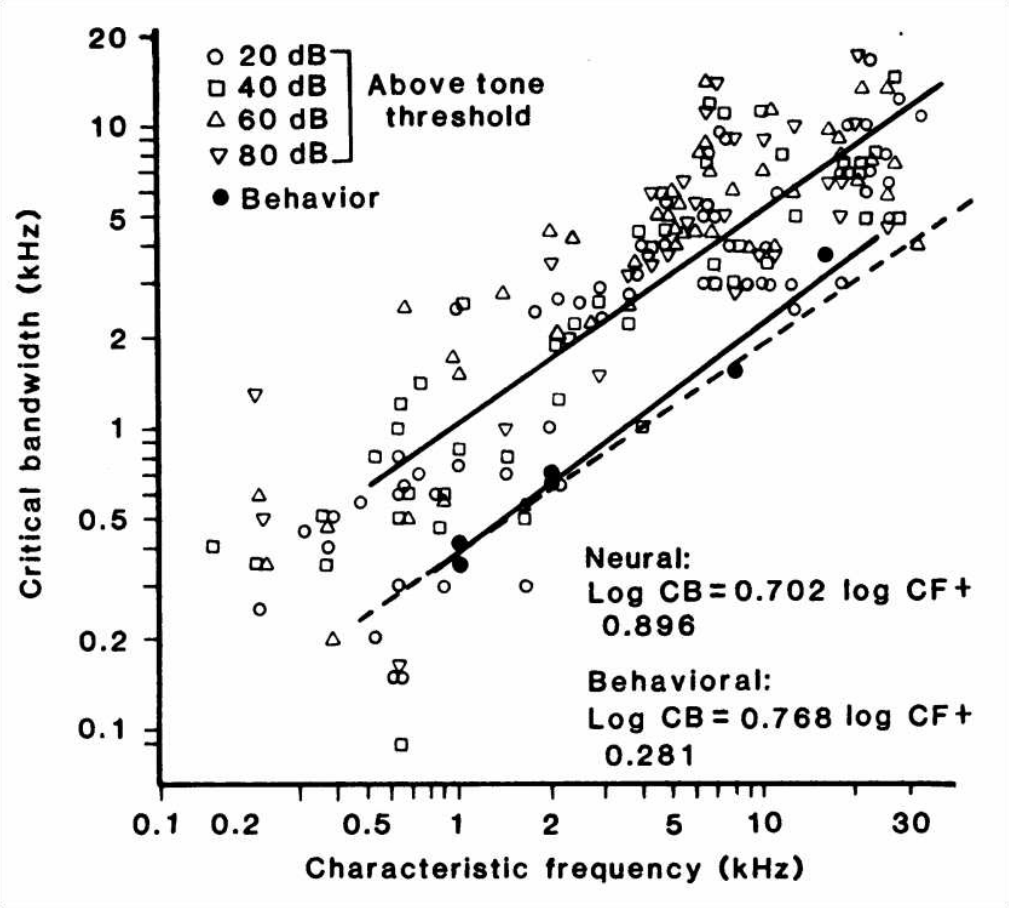


Figure 1.14

Figure from Ehret and Merzenich (1985) showing bandwidth measures using the band-widening technique vs frequency, measured both behaviourally and neurally. Open symbols show results for SU from the IC of the cat with the value in dB referring to the level of the signal above threshold of the unit. The solid lines are the regression lines of the data; the upper for the neural data, the lower for the behavioural data taken from Pickles (1975) and Pickles (1979). The dashed line is the neural regression line adjusted to compensate for "overmasking".

1.10 Summary

The field of frequency selectivity has developed along two fairly distinct lines: psychophysics and physiology. The first uses mainly masking experiments to measure perceptual bandwidths, and the latter mainly uses mean firing rates in response to pure tones. Due to the highly non-linear nature of the auditory system, different methods for measuring frequency selectivity give different answers, none of which can be said to be right or wrong. Also although the nature and origins of masking have been extensively explored, they are still not fully understood. Non-linearities of the BM play a large role, but non-linear central mechanisms must also contribute; this is evidenced by tone-in-noise representations of masking evolving up the auditory system, the fact that forward masking exists, and the mixed correspondence between psychophysical bandwidths as well as physiological bandwidths made using different methods in different brain regions. Despite this, psychophysical methods have developed over the years and notched noise, forward masking, with fixed signal level is a method commonly used in human psychophysics nowadays; it has however never successfully been implemented in animal behavioural studies. In order to understand the neural origins of perceptual frequency selectivity it would be advantageous to compare psychophysical and physiological measurements made at different stages of the auditory system. Attempts to do this in the past have suffered from using outdated, problematic methods, or comparing bandwidth measures made using fundamentally different methods. What is needed is a direct comparison between the two subfields using the same most up-to-date approach for both.

1.11 Aims and Objectives

There are two main aims to the study in this thesis, both leading towards a unifying method to bridge the gap between psychophysics and physiology in frequency selectivity research:

1) Physiologically measure frequency selectivity using a modern psychophysical method.
2) Use an up-to-date psychophysical method, to measure bandwidth behaviourally in animals.

The reasons for these aims have been outlined in this chapter. In order to bridge the two sub-fields, it is advantageous to use one method to measure bandwidth in both. This allows for a comparison of results without the caveat of methodological differences. Also previous attempts to do this used psychophysical methods that are now outdated, so it becomes necessary to repeat the process with an up-to-date methods. Also great advances have been made in human psychophysics that have never been implemented in animals. Almost all physiological data (at least at the single neuron level) cannot be collected in humans and must be collected in animals. Therefore using an up-to-date psychophysical method to measure behavioural and neural bandwidths in animals would be a very useful step toward this general aim.

To do this I will be conducting two main experiments:

1) Measure neural bandwidth, in the guinea pig, using notched noise with fixed signal, in both forward and simultaneous masking arrangements. Do this in the IC and A1.
2) Train ferrets to perform a notched noise task, with fixed signal and both forward and simultaneous masking arrangements.

As far as I am aware bandwidth has never been measured neurally using notched noise in such a way before. Seeing how such measures compare to behavioural results of the guinea pig (from the literature) as well as neural bandwidth measured using different methods, will teach us a lot. Also it can answer questions about the origins and limits of frequency selectivity. Guinea pigs will be used since they are an ideal experimental animal due to their size and similarity in hearing to humans. Any results could therefore potentially tell us something about human hearing. I will be recording in IC since it is one of the most crucial low level auditory brain structures and can tell us much about the auditory periphery. I will be recording in A1 since this is thought to be the seat of

perception, and we could learn much about an elusive brain region. Also I will be measuring bandwidth with both forward and simultaneous masking to see if the difference between the two, seen so commonly in psychophysics, is present at the neural level. I will also be using an SDT approach since this uses more of the available neural information to establish signal detection, and yields results far more comparable to psychophysics.

Very few animal behavioural studies have managed to successfully use forward masking to obtain bandwidth estimates. This is the standard method used in human psychophysics so to bring animal studies up-to-date, it is important to develop a successful method to do this. I will be using a novel technique to try to obtain both forward, and simultaneously, masked estimates of bandwidth using the notched noise method in the ferret. I will be using the ferret since it is a very good behavioural animal that predominantly uses hearing to hunt for prey. This makes its dependence on hearing more similar to humans than many prey animals that might be used for such tasks. I will also be using both fixed signal, and fixed masker, arrangements to see if a difference can be found in their results.

1.12 Structure of thesis

The remainder of this thesis is broken down into two data chapters and a summary. The first data chapter describes the physiological work I conducted in the guinea pig. It is comprised of two experiments; one in the IC, and the other in A1. Since there is much crossover between the two experiments, the chapter contains one introduction and one methods section for both. The introduction is a concise summary of the elements of the general introduction that are relevant to this chapter, since I feel this helps with clarity. A separate results and discussion section follows for each experiment.

The second data chapter describes the psychophysical work I conducted in the ferret. This too begins with a concise introduction, summarising relevant elements of the general introduction. One of the main aims of this work was to establish and develop a novel behavioural method to measure frequency selectivity. I have included the results of its testing and validation in the methods section. The results section therefore only contains details of the frequency selectivity measurements made using this method. The discussion section refers to both the methods and results sections.

The final chapter is a summary of the two data chapters, as well as a description of what we have learned from them and what future work can follow on.

CHAPTER 2: Guinea pig physiology

2.1 Introduction

Frequency selectivity is the ability to resolve the individual frequency components in a complex sound, and is a fundamental part of hearing. It is vital for identifying subtly different sounds, without which complex tasks, such as communication through speech, would be impossible. For over a century it has been known that the ear works as a frequency filter (Helmholtz, 1877), but it wasn't until the seminal study of Fletcher (1940), describing the auditory system as a sequence of overlapping band-pass filters, that research in this field truly began. Since then it has progressed along two distinct lines: psychophysics and physiology.

In psychophysics the main metric used is the threshold of a tone (or narrowband noise) in the presence of a masking sound. The idea being that the closer a sound is in frequency to another, the more of a masking influence it has (Wegel and Lane, 1924), and by exploring this relationship it is possible to infer the shape and size of the auditory filters. Non-linearities throughout the auditory system mean that the method used to measure AF bandwidth has an effect on the result, and therefore its choice and precise parameters become very important. Notched noise masking has become the favoured method, since it does not suffer from problems arising from off-frequency listening (Johnson-Davies and Patterson, 1979; Patterson and Nimmo Smith, 1980; O'Loughlin and Moore, 1981) and is more accurate than other methods such as band-widening (Bos and de Boer, 1966; Patterson and Henning, 1977). In addition, the forward masking stimulus arrangement has come to the fore since it does not suffer from some confounds caused by non-linearities of the cochlea, such as suppression, as simultaneous masking does (Sachs and Kiang, 1968; Heinz et al., 2002). This tends to make simultaneously masked bandwidth estimates broader than forward ones

(Houtgast, 1974; Leshowitz and Lindstrom, 1977; Moore and Glasberg, 1981; Moore, 1987; Oxenham and Shera, 2003). Although there is still some controversy over the issue (Eustaquio-Martin and Lopez-Poveda, 2011; Lopez-Poveda and Eustaquio-Martin, 2013), the current standard in human psychophysics is to determine detection thresholds by keeping the signal level fixed and varying the masker level (analogous to fixing the output level of the filter) (Rosen et al., 1992; Rosen and Baker, 1994; Baker et al., 1998; Rosen et al., 1998). These developments have mainly been made through human studies, but some animal studies have also contributed (Pickles, 1975; McGee et al., 1976; Kuhn and Saunders, 1980; Serafin et al., 1982).

Frequency selectivity has been explored physiologically in most (if not all) known auditory brain regions, including the IC and A1, predominantly in animals. The main metric used is the firing rate of a cell (or population of cells) in response to a pure tone stimulus (Tasaki, 1954; Katsuki et al., 1958; Hernandez et al., 2005; Taberner and Liberman, 2005). Tone-on-tone masking is another frequently used method (Ehret and Merzenich, 1988; Ramachandran et al., 1999), but despite both using masking, neither approach can be said to measure bandwidth in the same way as psychophysical methods (Ehret and Schreiner, 1997). In addition, almost all physiological investigations only use the mean firing rates of cells and ignore their variability, thereby discarding important information when exploring the detectability of sound (Relkin and Pelli, 1987). If we are to understand the neural basis of frequency selectivity, it would be useful to compare the performance of neurons against that of the subject as a whole (be it human or animal). This means comparing psychophysical and physiological measures of bandwidth, and due to the non-linear nature of the auditory system, using the same method for both.

There have been some attempts to apply the same psychophysical method to neural and behavioural data in the same animal (Pickles, 1979; Ehret and Merzenich, 1985), but they met with little success. However, since then techniques for both collecting and analysing psychophysical and physiological data, and our understanding of frequency selectivity has grown. Therefore here I conducted a series of experiments using notched noise stimuli, similar to those used in Oxenham and Shera (2003), on single-units and multi-units (MU), as well as LFPs, in the IC and A1 of the guinea pig. The methods are the current standard in human psychophysics and have yet to be applied to animal physiology. In addition, a signal detection theory approach will be used (Green and Swets, 1966; Relkin and Pelli, 1987; Alves-Pinto et al., 2010) which takes into account all

aspects of neuronal activity that might give rise to discriminability, and makes the results even more comparable to those in psychophysics.

2.2 Methods

2.2.1 Subjects

Experiments were performed on 60 guinea pigs (Cavia porecellus) (24 tricolour, 36 albino) of both sexes (10 male, 50 female) weighing 265-823 g. All tricolour guinea pigs were bred in house whereas all albinos (Dunkin-Hartley) were obtained from Harlan Laboratories. All experiments were carried out in accordance with UK Home Office regulations.

2.2.2 Surgical procedure

All neurophysiological recordings were made in non-recovery experiments on anaesthetised animals. An experiment began with the animal being anaesthetised with an intra-peritoneal (i.p) injection of Urethane (0.9 – 1.3 $g.kg^{-1}$ in a 20% solution, Sigma), at a dose of approximately 4.5 $ml.kg^{-1}$, followed by an intramuscular (i.m) injection of 0.2 ml Hypnorm (Fentanyl citrate 0.315 $mg.ml^{-1}$, Fluanisone 10$mg.ml^{-1}$, Janssen). Afterward, a subcutaneous injection of 0.2 ml Atropine Sulphate (600$\mu g.ml^{-1}$, Phoenix Pharma) was administered to suppress bronchial secretions. Anaesthesia was maintained throughout the experiment by supplementary 0.2 ml i.m injections of Hypnorm whenever the forepaw reflex was evident. This was checked at least every 30 minutes.

Once anaesthesia was established, the animal was transferred to the sound attenuating chamber (Whittingham acoustics Ltd) and the ears, head and throat shaved. Next, a tracheotomy was performed; the trachea was exposed, an incision made, and then cannulated with a polythene tube (2.4 mm external diameter, 2 mm internal diameter). This allowed the animal to be artificially respired with 100% oxygen (BOC Healthcare), mixed with air, throughout the experiment. The animal's core temperature was monitored and maintained at 38 °C by a thermal blanket and rectal thermometer connected to a homeothermic control unit (Harvard Apparatus). The heartbeat was monitored with an electrocardiogram (VetSpec), via two electrodes inserted into the skin on either side of the ribcage, and maintained at 300-360 beats per minute. End tidal CO_2 level was monitored by a sensor in the ventilator (Harvard Apparatus) and maintained at 28-38 mmHg by adjusting the volume and rate of flow.

Both tragi were removed to enable access to the ear canals and any obstructions cleared. Next, the animal's head was placed into a stereotaxic frame and hollow Perspex speculae positioned in the ear canals, such that the tympanic membrane was visible. The skull was exposed by making a midline incision along the scalp, the skin retracted and the

periosteum and temporalis muscles removed. The auditory bullae were exposed and a small hole drilled into each. Long polythene tubes (>20 cm) were inserted (0.5 mm internal diameter) and sealed using petroleum jelly (Vaseline). This helped equalise middle ear pressure since this cannot be done by the animal under anaesthesia. In order to increase the stability of recordings a small incision was made into the dura mater above the posterior cranial fossa, which allows cerebrospinal fluid to flow freely, minimising brain movement due to breathing.

If the IC was the target, then a small rectangular craniotomy was made on the top of the skull, 10-14 mm behind bregma and 0-2.5 mm lateral from midline, on the right hand side. This allowed access to the right IC from directly above. If the target was A1 then a roughly 6×6 mm craniotomy was made on the right hand side of the skull using coordinates described in Wallace et al. (2000). In both cases, a dental drill (Minitech Dentimex) was used to cut though the skull, the dura mater removed and the exposed cortex covered with agar (1.5% agar in 0.9% saline). This protected the cortex from desiccation and improved stability.

2.2.3 Recording equipment

Extracellular recordings of neural activity were made using glass-coated tungsten electrodes, made in-house according to the methods in Bullock et al. (1988). Tip size was in the range of 2-12 μm to achieve high impedance for high signal to noise ratio recordings. Between 1 and 4 of these electrodes were mounted onto a printed circuit board allowing simultaneous, independent recording from each. In some experiments 16 channel multi-electrode arrays (Neuronexus; A series), in both 1×16 single shank and 4×4 four shank arrangements, were used instead. These have electrode sites 15 μm in diameter, so have lower impedance, and were only used to collect MU and LFP.

Microelectrode boards were mounted onto a 16 channel high impedance Tucker Davis Technologies (TDT) head stage (RA16AC), which band-pass filters the signal (0.16 – 6000 Hz) before passing it onto a pre-amplifier (TDT; RA16PA) to be digitised. Next it passed a digital signal processor (TDT; RX7) under the control of Jan Schnupp's BrainWare, where it was further filtered and amplified. Action potentials were deemed to have occurred when the amplitude of the band-pass filtered (450-6000 Hz) signal crossed a certain threshold. This threshold was manually set by the experimenter to only pick out SU or MU activity. In addition the original unfiltered signal was recorded to allow LFPs to be extracted offline, by low pass filtering at 300 Hz. All analyses on recordings was done in Matlab (Mathworks).

2.2.4 Stimulus delivery

All stimuli were generated by a digital circuit, created in RPvdsEx (TDT software), loaded onto an RX8 processor (TDT; System 3). Stimulus parameters and presentation order were controlled by Jan Schnupp's BrainWare on a Windows PC. Once generated the stimuli were passed to a 24-bit sigma-delta digital-to-analog converter, before being attenuated by programmable attenuators (TDT; PA5). Finally the signal passed through a power attenuator, which would reduce the output by 20 dB, before reaching the custom-modified tweeters (RadioShack; 40-1377) inserted into the hollow specula, creating a sealed system. The maximum output of the system was roughly 100 dB SPL (120 dB if the power attenuators were set to 0) and this was calibrated at the beginning of each experiment using a condenser microphone (Brüel and Kjaer; 4134). The microphone was placed a few millimetres from the tympanic membrane, and was fitted with a 1mm calibrated probe tube. The calibration consisted of 20 presentations of 50 ms long white noise with no attenuation, the recordings were averaged, and the Fourier transform taken and adjusted for the characteristics of the microphone. This process was repeated for each ear and the results checked to ensure the frequency response was flat.

2.2.5 Data collection

The head stage for the multi-electrode array was attached to the stereotaxic frame by a micro-drive unit (Burleigh) and a micro-manipulator arm (Kopf), allowing electrode placement to within 1μm. In order to do so, the skull of the guinea pig was levelled and the electrode array inserted into cortex from above, vertically dorso-ventrally, 11.5 mm behind bregma and 2 mm lateral of midline (Rapisarda and Bacchelli, 1977). The electrode was advanced to ~7 mm below the surface of the brain until auditory responsive units were found. Since IC is ~2-3 mm across and ~4-5 mm long, all penetrations were made in this range around the coordinates stated above. In a small number of experiments the exact location of electrode penetrations was verified histologically (see Figure 2.1). For A1 experiments visual landmarks were used; the micro-manipulator was mounted at a 45° angle to vertical and advanced into the region defined as A1 in Wallace et al. (2000) using blood vessels as guidance. Figure 2.2 shows an extract from the paper depicting this region and beside it a photo from an experiment with the area I deemed to be A1. Penetrations were made across this whole region and CF of units used to corroborate location.

For both brain regions the protocol for searching for units was similar; repeated 50 ms bursts of ~70 dB SPL broadband noise were played (at 450 ms intervals in IC, 800 ms in A1) and the electrode advanced until a responsive unit was found. Broadband noise was used because it has energy at all frequencies and is therefore the most likely stimulus to find auditory responsive neurons. The various sets of stimuli would then be played until either the unit stopped responding, or the full repertoire was recorded. At this point the electrode would be advanced further. Once the driven activity in response to the auditory signal ceased, the electrode was withdrawn and another penetration made in a different location.

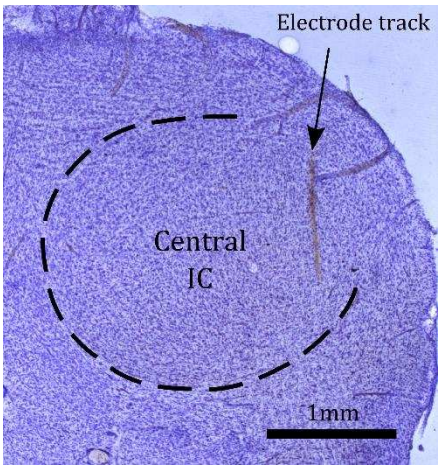


Figure 2.1

Nissl stained slice of central IC from one experiment. Central IC and an electrode track are highlighted.

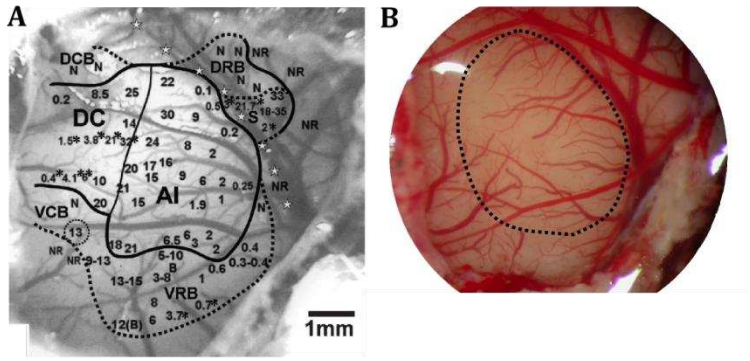


Figure 2.2

A *Extract from Wallace et al. (2000) showing the location of A1.* ***B*** *Image of cortex from one experiment depicting the area identified as A1.*

2.2.6 Stimulus sets

Once a neuron (or unit) was located, several sets of stimuli were presented, each evaluating a different characteristic of the unit. Below is a description of these, arranged in the order they were presented. In each set, different stimulus conditions were presented in randomly interleaved order. The procedure was the same for both IC and A1, except in IC the stimulus presentation rate was 450 ms, while in A1 it was 800 ms. This is because it takes longer for response activity to dissipate in cortex (Ulanovsky et al., 2004).

Frequency response area

A series of pure tones, 10 ms in duration (5 ms cosine squared on/off ramps), of various frequencies and levels were presented to the unit. The range of levels was usually 20 to 100 dB SPL in 5 or 10 dB steps, and the range of frequencies 200 – 20,000 Hz in 0.05 to 0.25 octave steps. The number of action potentials elicited by the unit in a window around the time of tone presentation was then calculated for each frequency and level combination. This window was chosen to be the region over which the unit displayed a firing rate above its spontaneous firing rate. These were then plotted in an FRA (see Figure 2.6 panel A for an example and 1.8.10 Physiological bandwidth measurements*:* FRA for more details). The CF was determined as the unique frequency which elicited an elevated firing rate (above spontaneous rate) for the lowest sound level, i.e. the frequency at the tip of the FRA.

Rate against level function

A 10 ms pure tone, with frequency set to CF, was presented 15 times at various levels, spanning the response range of the unit in 5 dB steps. The firing rate of the unit was calculated by taking the spike count in a small window around the time of signal presentation, and calculating the average firing rate for each stimulus level. Again, this window was chosen to be the region over which the unit displayed a firing rate above its spontaneous firing rate. Plotting firing rate against level gives the RLF. The threshold of the unit was deemed to be the point at which the RLF met the spontaneous firing rate and was determined visually (see Figure 2.6 panel B).

Notched noise masking

The signal was set as a 10 ms pure tone (5 ms cosine squared on/off ramps) with frequency at CF and level fixed at ~5-10 dB above threshold. Masker level varied between presentations, in 5 dB steps over the range 30-100 dB SPL, with an additional no masker condition. Level was varied in this arrangement since it is the current

standard in human psychophysics (Rosen and Baker, 1994). The characteristics of the masker were similar to those used in Oxenham and Shera (2003); each noise band was 0.25 × CF in width, centred symmetrically about CF. Although AFs are known to be asymmetric, at moderate sound levels they are approximately symmetric, so only using symmetrically distributed notches should not cause a problem (Patterson, 1974; Patterson and Nimmo Smith, 1980; Patterson, 1986). Five different notch widths were used for each unit, with ΔF (defined as the gap in frequency between CF and the closest edge of each noise band) linearly spanning the range $0 \leq 0.4 \times CF$. Masker duration was 150 ms, inclusive of 5 ms on/off ramps. This, and the duration of the signal (10 ms, inc. 5 ms on/off ramps), were established through validation experiments to maximise the amount of forward masking, while minimising the total stimulus presentation time. Forward masking was usually attempted first, with the signal immediately following masker, before simultaneous masking, with signal appearing 20 ms before the end the masker (130 ms after onset). Figure 2.3 shows a schematic of the stimuli. Each masker level and notch width combination was presented with the signal 30 times and without the signal a further 30 times.

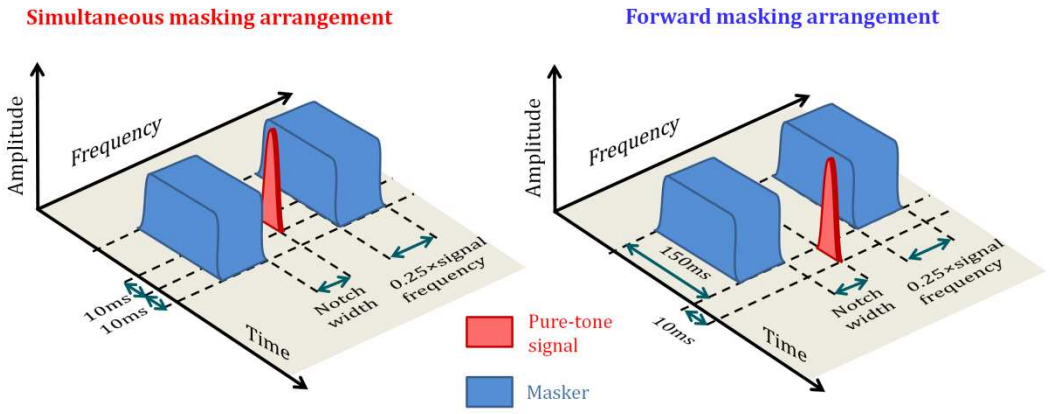


Figure 2.3

Temporal and spectral characteristics of masker and signal in the forward and simultaneous masking arrangements.

2.2.7 Analysis

In the notched noise experiment recording, the unit's response when the signal is both present and absent is the key to using SDT, but a problem still remains: what feature of the unit's response encodes information about frequency selectivity? The practical implication of this question is how to quantify the difference between the two conditions.

Different approaches

A number of different possibilities were considered, falling into two categories: a rate code or a temporal code. In a rate code only the number of spikes elicited by the unit encodes information, while in a temporal code the precise timing of these spikes is also informative. A rate code is simpler and robust to noise, both internal and external, whereas a temporal code has the capacity to carry much more information. One approach widely used in neuroscience which can be used in the context of a temporal code (Borst and Theunissen, 1999; Fonollosa et al., 2012) is information theory, but the relatively low presentation number means this is unsuitable for these data. Linear discriminant analysis (Duda et al., 2001) was another approach that was investigated but suffers from the same problem.

Another multi-dimensional spike train classifying approach which did work was k-means clustering (Ferrington et al., 1988; Helm et al., 2013). In this approach, the spike train is split into bins, and the number of spikes in each bin is calculated. Each bin is treated as a separate dimension and the value in each bin defines the spike train's location in multi-dimensional space. An iterative process is then used to split the data into two clusters. One parameter that has a large impact on the performance of the classifier is the size of the bins. A number of different bin sizes over the range 1-20 ms were tried and the average proportion correct for all stimulus conditions across all units for each of these time bins is shown in Figure 2.4. This plot shows a clear trend towards larger bin sizes yielding higher proportions correct, suggesting there is little information in precise spike timing. In most cases the easiest conditions, i.e. the ones with no masker, never reached 100% correct, which suggests that looking at timing actually hinders correct classification.

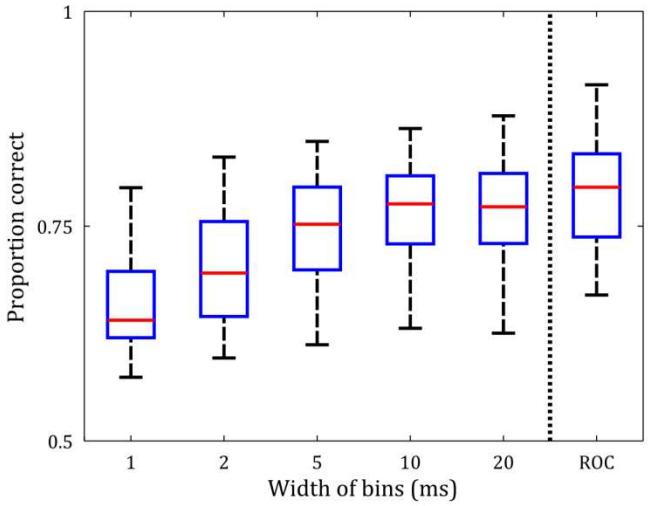


Figure 2.4

Proportion correct across all stimulus conditions, across all units, using the k-means clustering classifier with various time bin sizes, versus the ROC rate code classifier.

This theory is corroborated when looking at the higher proportion correct using a receiver operator characteristic (ROC) classifier, which uses a rate code. Here the spike count of each spike train was calculated in a window around the signal presentation time. These spike counts were then used to create an ROC curve, and the area under the curve (AUROC) taken to represent how discriminable the signal case was from the no signal case. This is a method widely used in SDT and has the advantage of being simple and computationally cheap to implement, and does not make assumptions about the statistics of the underlying distributions. It gives a score between 0.5 and 1, and can therefore be thought of as the proportion correct that would be obtained in an equivalent 2I2AFC task (Green and Swets, 1966; Macmillan and Creelman, 1991). For these reasons it was decided to use spike count with AUROC to represent how detectable the signal is in each stimulus condition. For further details on this approach see the 1.8.11 Signal detection theory.

Notched noise analysis

The first step in analysing the notched noise SU and MU data was to calculate the spike count in a window around the time of signal presentation, for each spike train from every stimulus condition. Figure 2.5 panels A and B show an example of these spike trains. The window was tailored for each unit to be the time over which the difference between signal and no signal presentations was maximal. A post stimulus time stimulus histogram (PSTH) was created using spike trains from all trials in which the signal was

presented, and a separate PSTH made from trials with no signal. The window was chosen to be the range, around the signal presentation, over which the two crossed. Since a rate code is used, this represents the area of maximum difference between the two. Figure 2.6 panels C and D shows examples. For the LFP data, the root mean square (RMS) in a 70 ms window after signal presentation was always taken. This represents the power of the LFP in response to the signal.

The AUROC was then calculated using these values, and neurometric functions built from them, as if they were proportion correct values; one for each notch width (see Figure 2.5 panel C). These are similar to psychometric functions, only based on neural data instead of psychophysical. Detection threshold for a given notched noise masker was taken as the SNR at which the neurometric function intercepted 75% correct. This is a common criterion since it is the point at which the psychometric function is steepest in a 2 alternative choice case (Wichmann and Hill, 2001). Next the AF function was fitted to these threshold SNRs (see Figure 2.5 panel D) using the following procedure:

1) A stereotypical spectral representation of the stimuli was made. This consisted of a vector of 20,000 elements, each representing a 1 Hz bin in frequency space from 0 – 20 kHz, one for each stimulus. All frequency bins that were present in the stimulus were represented by a value of 1; the rest were filled with 0. For example, a 10 kHz, 0.1 × CF notched noise masker would have 1s in bins 6,501 – 9,000 and 11,001 – 13,500, and 0s elsewhere.
2) Initial parameters for the AF function were estimated, and the resulting AF function integrated over the spectral representations described in the previous step. The result was an SNR value for each notch width which could be compared to the threshold plot.
3) The AF parameters were then manually tailored so that the simulated threshold plot was close to the experimental one.
4) An optimisation procedure was then used to estimate the parameters of the AF that best fit the data; the summed squared difference between the simulated and experimental threshold plots, in dB, was minimised.

The two parameter roex(*p*,*r*) function (Patterson et al., 1982) was used to represent the AF since this has the fewest parameters, increasing the stability of the optimisation. The roex function with just one slope parameter for both sides, i.e. symmetric filter, was used since the levels of stimuli are very close to threshold so the asymmetry of the AF should be negligible (Patterson, 1974; Patterson and Nimmo Smith, 1980; Patterson,

1986). Also since only symmetric notches were used, there is no benefit in using a higher order asymmetric function (Glasberg and Moore, 1990) (see 1.4 Auditory filter shape for more details). The final step in estimating bandwidth involved calculating the equivalent rectangular bandwidth of the fitted function. This was done by finding the width of the rectangular filter with equal gain and passing the same amount of energy as the auditory filter, i.e. the width of the rectangle with same height and area as the auditory filter shape in frequency space (see Figure 2.5 panel E). A two dimensional confidence region (spanning CF and ERB) was calculated for each ERB estimate by a bootstrapping method. Thirty random data points were selected, with replacement, from the 'signal present' spike counts, and thirty from the 'no signal'. CF and ERB were then calculated for these data by fitting a roex function, using the process described above. This was repeated 3000 times, and the ellipse encapsulating 95 % of the resulting data points used to represent the 95 % confidence region.

FRA analysis

In addition to the notched noise method, an ERB was calculated from the FRA. This is a method more commonly used to estimate filter width in physiology, and can therefore act as a good within-unit comparison between the two approaches. The edge of the area of the FRA over which the cell elicits an elevated firing rate is used to represent the AF. An algorithm was used to automatically calculate this edge using the following criteria: for each signal frequency, the lowest signal level that yielded a firing rate 1) at least 1 spike higher than, 2) at least 10% higher than, and 3) at least 4 standard deviations above the spontaneous firing rate of the unit. The script then fitted a roex(p,r) function, with independent slope parameters for each side, to the edge of the FRA. ERB was calculated in the same way as above.

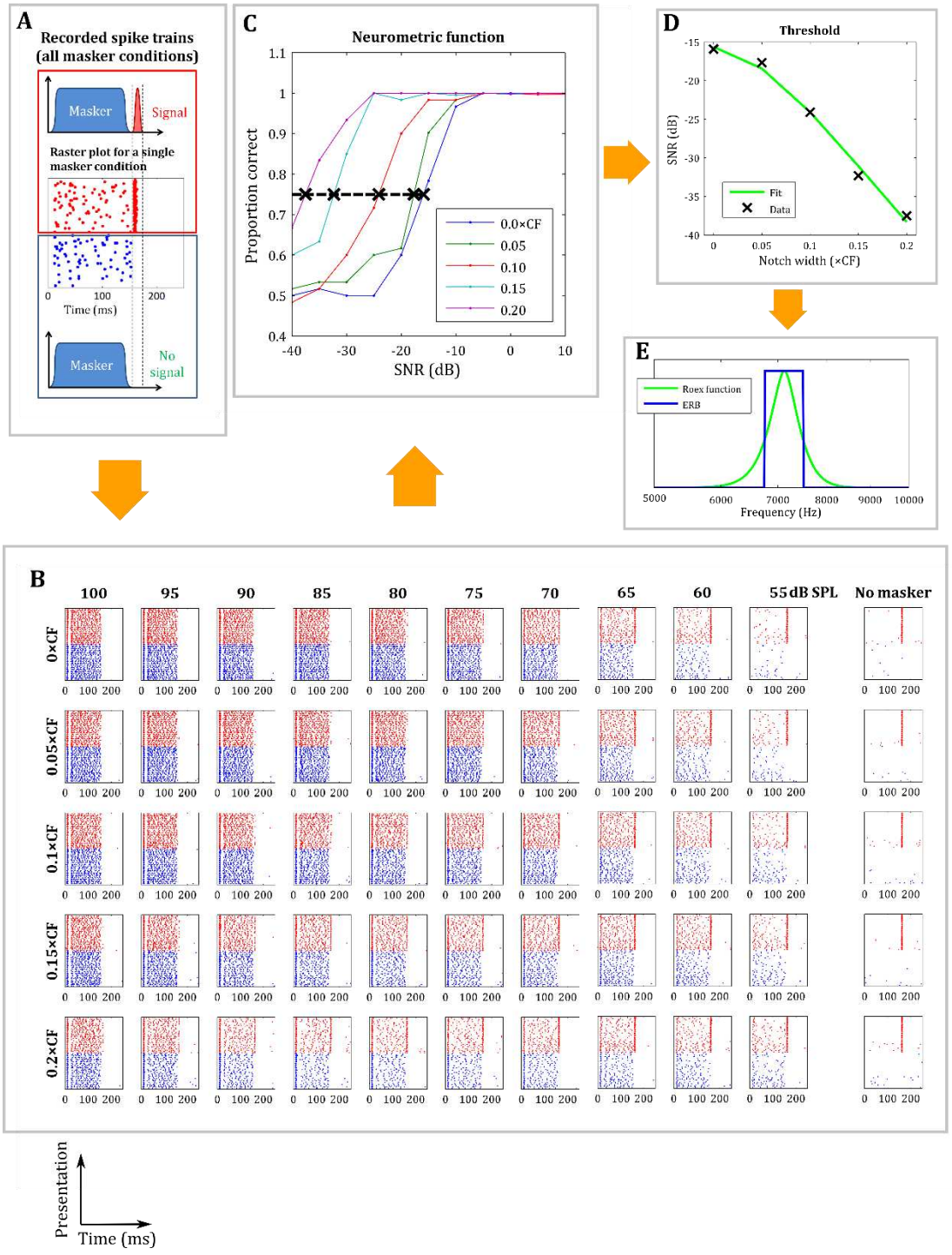


Figure 2.5

A Single masker condition, showing all signal and non-signal presentations separated. Dotted line shows spike count window. B Raster plot for all masker conditions. C Neurometric function obtained from the raster plots by calculating the AUROC for each masker condition. D Threshold SNRs taken from neurometric function (black crosses) with auditory filter fit. E Auditory filter fit, in this case roex(p,r), with ERB.

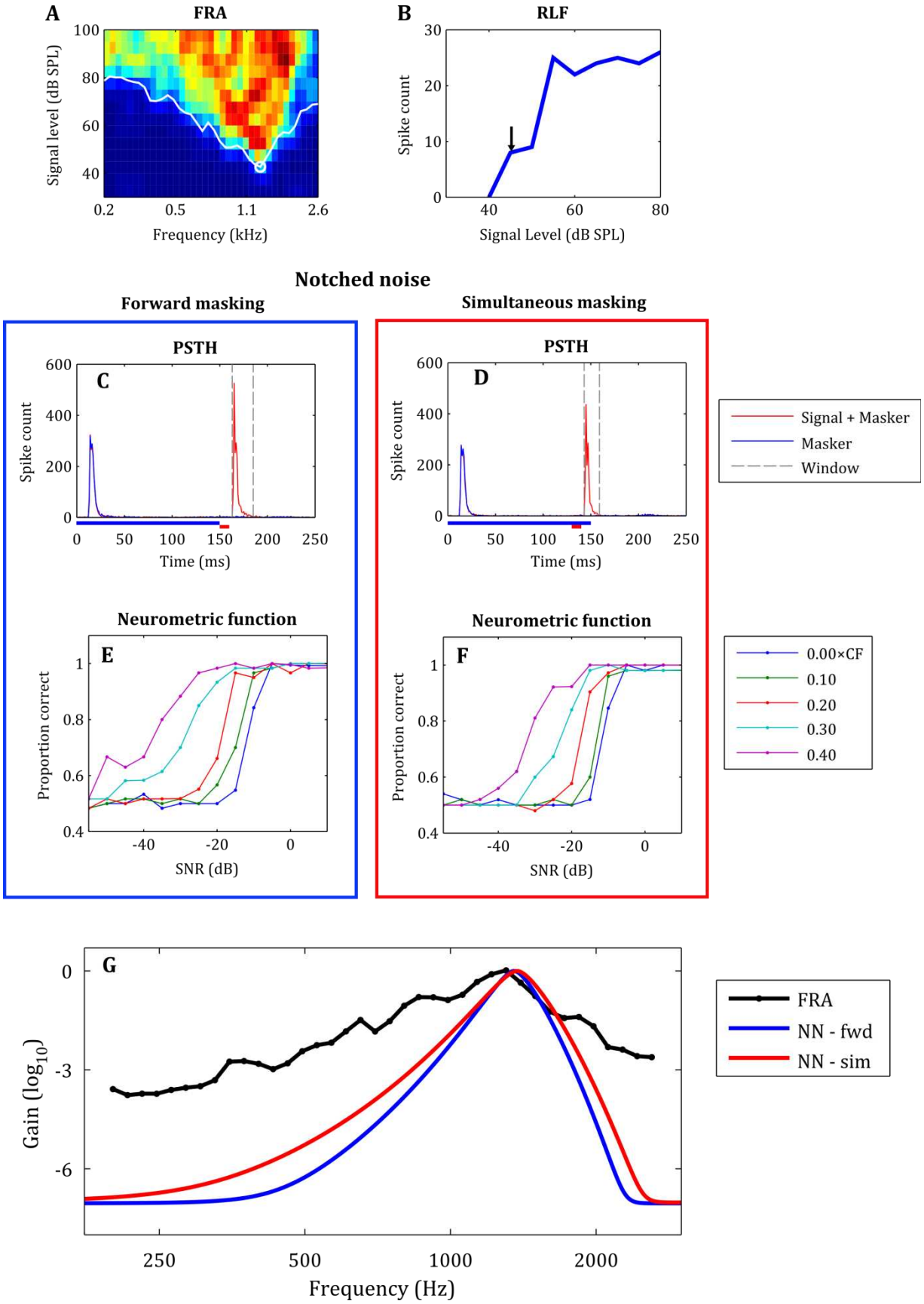


Figure 2.6

Example of plots for a since AI unit. ***A*** *FRA with edge (white line) and CF (white circle). Red = high, Blue = low firing.* ***B*** *RLF at CF (10,500 Hz) with signal level used for notched noise marked by black arrow.* ***C*** *&* ***D*** *PSTHs and* ***E*** *&* ***F*** *neurometric functions for forward and simultaneous notched noise masking. Blue PSTH = masker only presentations, Red = masker + signal, grey dashed lines = response window. Each neurometric function is for a different notch width.* ***G*** *Filter estimates obtained for one unit from edge of FRA (black), forward masked (blue) and simultaneously masked (red) notched noise. The stereotypical filter used is the roex(p,r).*

2.3 Results - IC

A total of 139 single-units and 46 multi-units were successfully recorded from the guinea pig IC. Neuronexus electrodes were used to record 7 of the MU; results from these sites were not different to those recorded with tungsten electrodes so they have been combined. Further MU data were extracted from 80 of the SU recordings by, independently for each unit, 1) creating a histogram of the raw electrophysiological recordings, 2) fitting a Gaussian function weighted towards lower voltages (this fits the function more to the recorded noise, and ignores the extreme values representing spikes), 3) redefine threshold as the 3 standard deviation point and recalculate spike events, and 4) exclude all SU spikes. The noise floor (created by background electrophysiological activity and recording equipment error) was very Gaussian in nature and this meant the method worked remarkably well at isolating salient spiking activity.

Figure 2.6 shows the suite of data that could be obtained from one unit. This example is of an SU (actually taken from A1). The FRA and RLF plots (panels A and B) were used to set the frequency and level of the signal, respectively, in notched noise (NN) experiments. Levels and notch widths of the masker were set independently for each unit to 1) ensure each neurometric function had a 75 % crossing, 2) maximise the spread of the neurometric functions, i.e. the differences in detection thresholds, and 3) minimise the number of different masker levels to keep experiment time low. Masker levels were always in 5 dB steps, and the same for all notch widths in a single NN experiment. Five notch widths were always used, spaced linearly from 0 × CF to 0.1 - 0.4 × CF; all were symmetric about CF. The reason for individually modifying the stimulus parameters for each unit, and the biggest problem facing this experiment, was the time constraint. It takes a long time to collect a neurometric function, since each point must be calculated from 60 stimulus presentations, each lasting 200 ms, with 250 ms between each to allow neural activity to settle. Enough points need to be collected to build a neurometric function, and enough neurometric functions need to be collected to describe the AF. In IC, when recording with a glass coated tungsten electrode, it is considered very good to "hold" a unit for ~1 hour without it suddenly bursting, or the electrode drifting away, or the unit simply stopping responding. In order to make a useful comparison between the different bandwidth measuring methods, it is important to complete all measures in the same unit to prevent any findings being swamped by inter unit variability (of which there is much). Therefore the experiment was a balance between keeping the duration low enough to fit all the measurement

methods, while ensuring enough data was collected to still obtain decent measurements for each unit. Figure 2.6 panel E shows the final product: three filter estimates yielding an ERB each. Notice how the FRA estimated filter is much broader than the NN ones, and how the simultaneously masked filter is broader than the forward one. This will be discussed in more detail later.

2.3.1 Unit characteristics

The response characteristics of all the units that yielded a notched noise ERB estimate were investigated and units were split into two categories depending on whether they fired only to the onset of the masker or showed a sustained response, firing throughout the masker presentation. There were also a number of units that showed onset behaviour for high masker levels which gradually changed into a sustained response as the level of the masker dropped. These categorisations are summarised in Table 2.1. Most SU and MU showed sustained responses. The proportion was higher for MU which is unsurprising since it represents the combined activity of many cells.

Threshold was calculated for each unit, defined as the lowest level at which the unit had a firing rate more than 3 standard deviations above its mean spontaneous firing rate. Assuming a normal distribution this equates to a 99.73 % confidence of elevated firing. Figure 2.7 panels A and C show these thresholds plotted against the CF of the unit for SU and MU respectively. Both sets of units have thresholds distributed fairly evenly across the range of hearing of the guinea pig, defined by the superimposed audiograms. Figure 2.8 shows a similar plot for over 2000 SU recorded from the guinea pig IC, from Palmer et al. (2013). The units in the current study (Figure 2.7) fit in well with this large collection, so would appear to reflect typical IC cells. In addition, panels C and D show the level of the signal used in the notched noise experiments plotted against its frequency. These levels are on average 8 dB above the thresholds of the SU (mean 7.7 dB, standard deviation 11.4 dB) and 6 dB for MU (mean 6.07 dB, standard deviation 11.21 dB). Also, the ERBs estimated using the different stimulus temporal arrangements seem to be spread out evenly across units, indicating there is no bias of threshold or CF in the notched noise results.

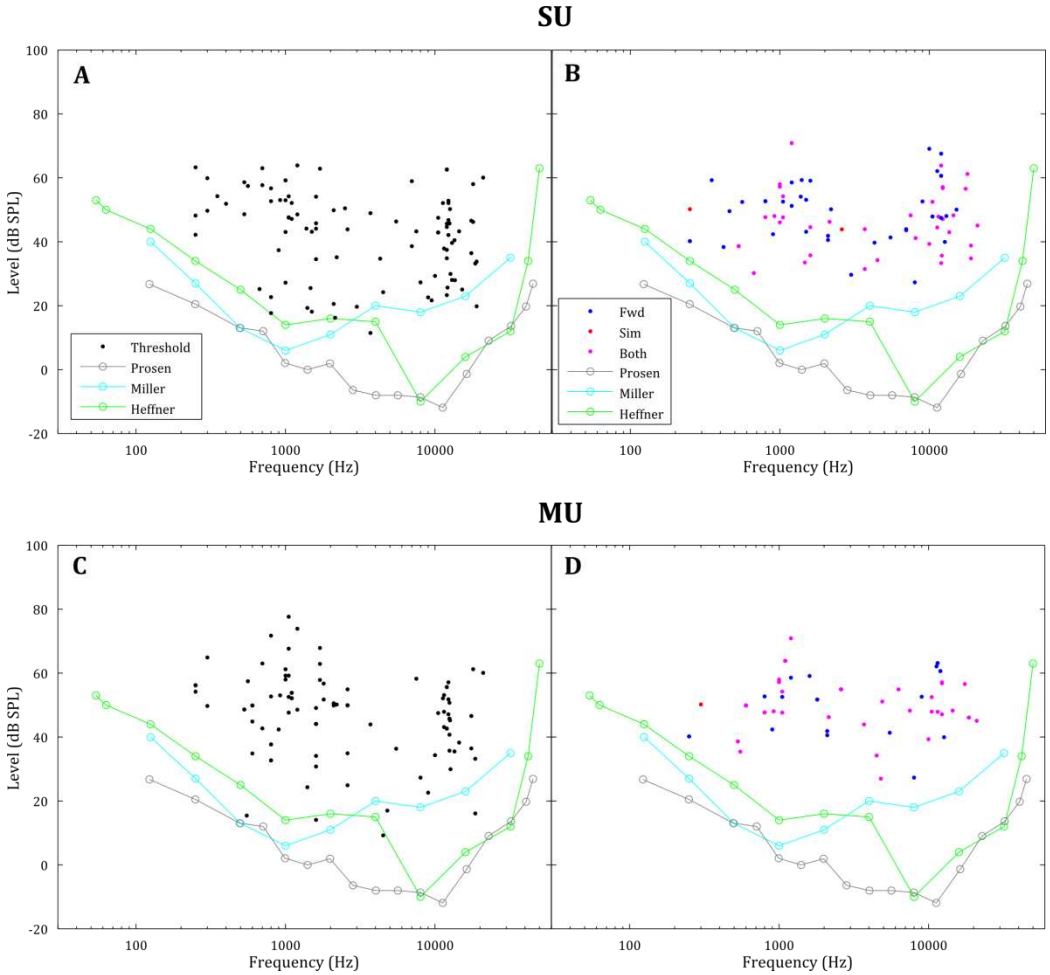


Figure 2.7

A *&* ***C*** *Threshold of units derived from rate level functions plotted against CF, for SU and MU respectively.* ***B*** *&* ***D*** *Fixed level of signal for each notched noise experiment; blue = units yielding just a forward masked ERB, red = just simultaneous, and magenta = both. Audiograms of the guinea pig taken from Miller and Murray (1966), Heffner et al. (1971) and Prosen et al. (1978).*

	SU		MU	
Onset	12	(23 %)	7	(20 %)
Sustained	35	(67 %)	28	(80 %)
Both	5	(10 %)	0	(0 %)

Table 2.1

Number of units that yielded a notched noise ERB with particular response characteristics. Units characterised according to their response to the 150 ms long masker. See text for further detail.

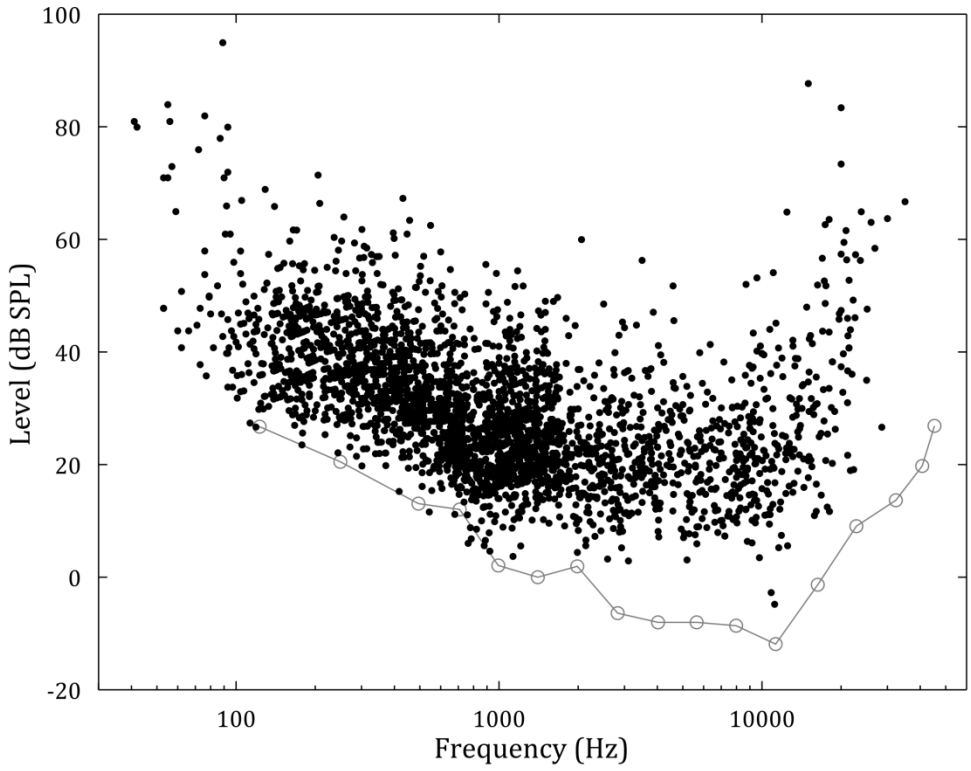


Figure 2.8

Threshold against CF for 2,826 SU from guinea pig IC from Palmer et al. (2013) (not plotted in paper), compared to guinea pig audiogram of Prosen et al. (1978).

2.3.2 SU results

In total, 45 forward and 28 simultaneously masked notched noise ERB estimates were obtained from SU activity in 17 animals. These can be seen in Figure 2.9. Panel A shows the ERB of each unit plotted against its CF for both notched noise methods. Panels B and C show the 95 % confidence regions for these points.

The clearest feature of these data is a linear increase, in log-log space, of ERB with CF, which is true for both NN methods. As a result, the lines of best fit were calculated by minimising the least squared mean error between the data points and a straight line, in log-log space, using the confidence regions to weight each point. This can be described by the following equation:

$$\underset{a,b}{\operatorname{argmin}} \sum_{i=1}^{N} w_i \left(ERB_i - aCF_i - b\right) \tag{2.1}$$

where

$$w_i = 1 \Big/ \sum_{j=1}^{3000} \sqrt{\left(ERB_i - erb_j\right)^2 + \left(CF_i - cf_j\right)^2} \qquad (2.2)$$

where CF_i and ERB_i are the CF and ERB of a particular unit, i = 1, ..., N are all the units, a is the gradient and b the intercept of the line of best fit, cf_j and erb_j the CF and ERB of a bootstrapped simulated unit. The forward masked data have a slightly wider scatter than the simultaneous, which is borne out by the lower R^2 value (0.62 to 0.89). Inspection of panel B, however, shows that those points furthest away from the line of best fit have very large confidence regions, all of which encompass the line. Also there seems to be little difference in the lines of best fit for the two methods.

Additional data have been plotted in panel A showing bandwidth estimates from a number of previous studies in different auditory brain regions of the guinea pig, using different methods: IC (Palmer) uses pure tone FRAs of 2217 SUs in the IC (Palmer et al., 2013); VCN (Sayles) uses roex(p) fits to pure tone FRAs of 432 SUs in the ventral cochlear nucleus (Sayles and Winter, 2010); ANF (Evans) uses roex(p) fits to pure tone FRAs of 80 SUs in auditory nerve fibres (Evans, 2001); Behaviour (Evans) uses psychophysical data, obtained by a lever-press task with both notched noise and rippled noise stimuli, from 10 guinea pigs (Evans et al., 1992). These are all in remarkably close agreement with each other and with notched noise data collected in this study. This is particularly true for high frequency units (> ~5 kHz), with lower CF units having potentially slightly narrower bandwidth estimates than these previous studies. Interestingly, the bandwidth estimates that are most disparate to my data are from Palmer's IC units, which are from the same brain area, although this difference is marginal. Looking more closely at the parameters of the lines in Table 2.2, the slopes of the NN lines are actually between that of Palmer et al. (2013) and those of Evans (2001) and Sayles and Winter (2010). This means that the rate at which the bandwidth increases with frequency is largest for Palmer et al. (2013) and smallest for Sayles and Winter (2010). The NN estimates grow at a rate between these two.

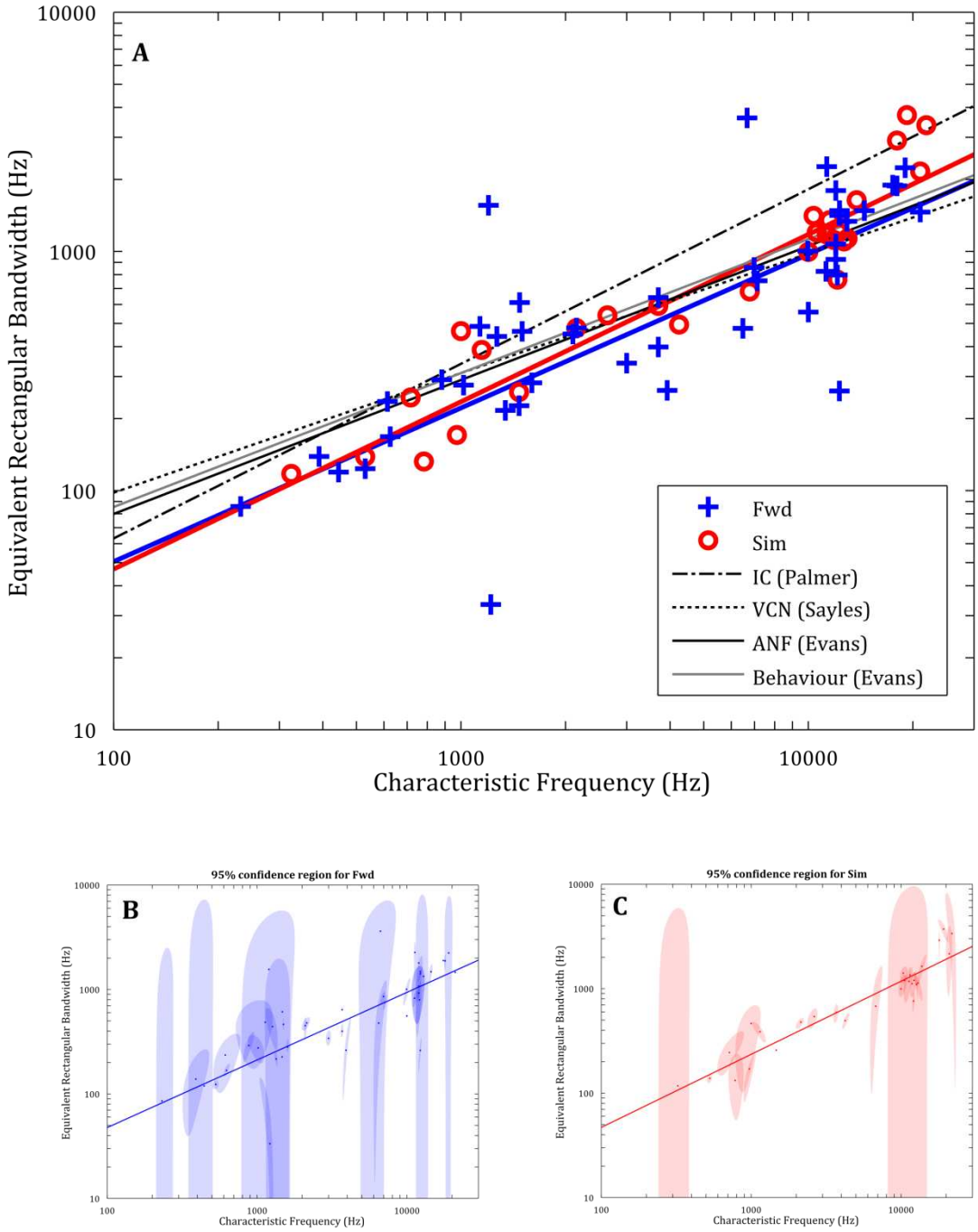


Figure 2.9

A *Forward and simultaneously masked notched noise ERB estimates plotted against CF of the SU (blue +, red o respectively) with lines of best fit (*R^2 *= 0.62 and 0.89,* χ^2_{red} *= 23.1 and 5.0 respectively). Lines of best fit for CF vs ERB plots for SU in the guinea pig from Evans (2001), Sayles and Winter (2010), Palmer et al. (2013). Also line of best fit for CF vs ERB plot from guinea pig psychophysics (Evans, 2001).* ***B*** *&* ***C*** *Data points in panel* ***A*** *with respective 95 % confidence regions calculated through bootstrapping.*

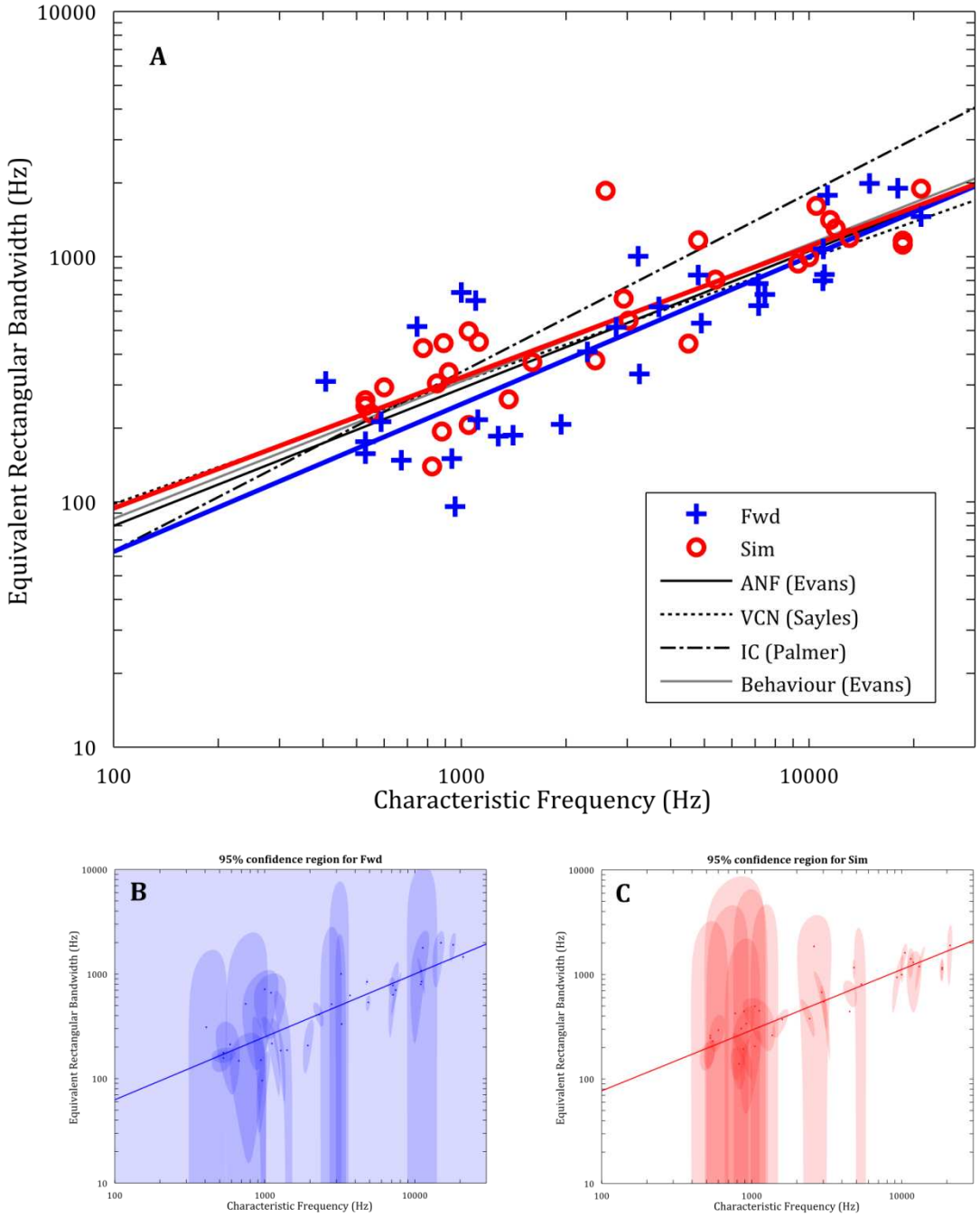


Figure 2.10

Similar to Figure 2.9 but for MU activity. ***A*** *Forward and simultaneously masked lines of best fit,* R^2 *= 0.68 and 0.75,* χ^2_{red} *= 4.3 and 2.8 respectively.*

2.3.3 MU results

Figure 2.10 shows plots of a similar nature for MU data. In total 31 forward and 31 simultaneously masked notched noise ERB estimates were made from MU activity in 13 animals. These data show a similar relationship to the SU; a clear linear increase in ERB with CF, very close agreement with data from both temporal arrangements and previous studies, with similar goodness of fits (R^2 of 0.68 and 0.75 for Fwd and Sim respectively). The rate of bandwidth increase is slightly smaller than with the SU (Table 2.2; forward NN slope is 0.60, simultaneous NN slope is 0.53), becoming more similar to the results of Evans and Sayles. Confidence regions are generally larger for MU, reflected in their lines of best fit having smaller reduced chi-squared values (Greenwood and Nikulin, 1996) - 4.3 and 2.8 for Fwd and Sim MU, versus 23.1 and 5.0 for SU. This is unsurprising since MU activity provides a wider range of firing rates leading to higher variability.

2.3.4 LFP results

LFPs were extracted from 23 forward, and 17 simultaneously masked notched noise estimates from both MU and SU in 3 animals. These can be seen in Figure 2.11, along with confidence regions for each. The lines of best fit for these data show the closest agreement with one another (Fwd and Sim), and with the previous studies (other than Palmer et al. (2013)). The difference to the results for the other data types, however, is very slight and the LFP data have the poorest goodness of fit (R^2 of 0.67 and 0.63 for Fwd and Sim) so this is unlikely to be meaningful; this poorer fit is most probably due to the relatively low number of data points.

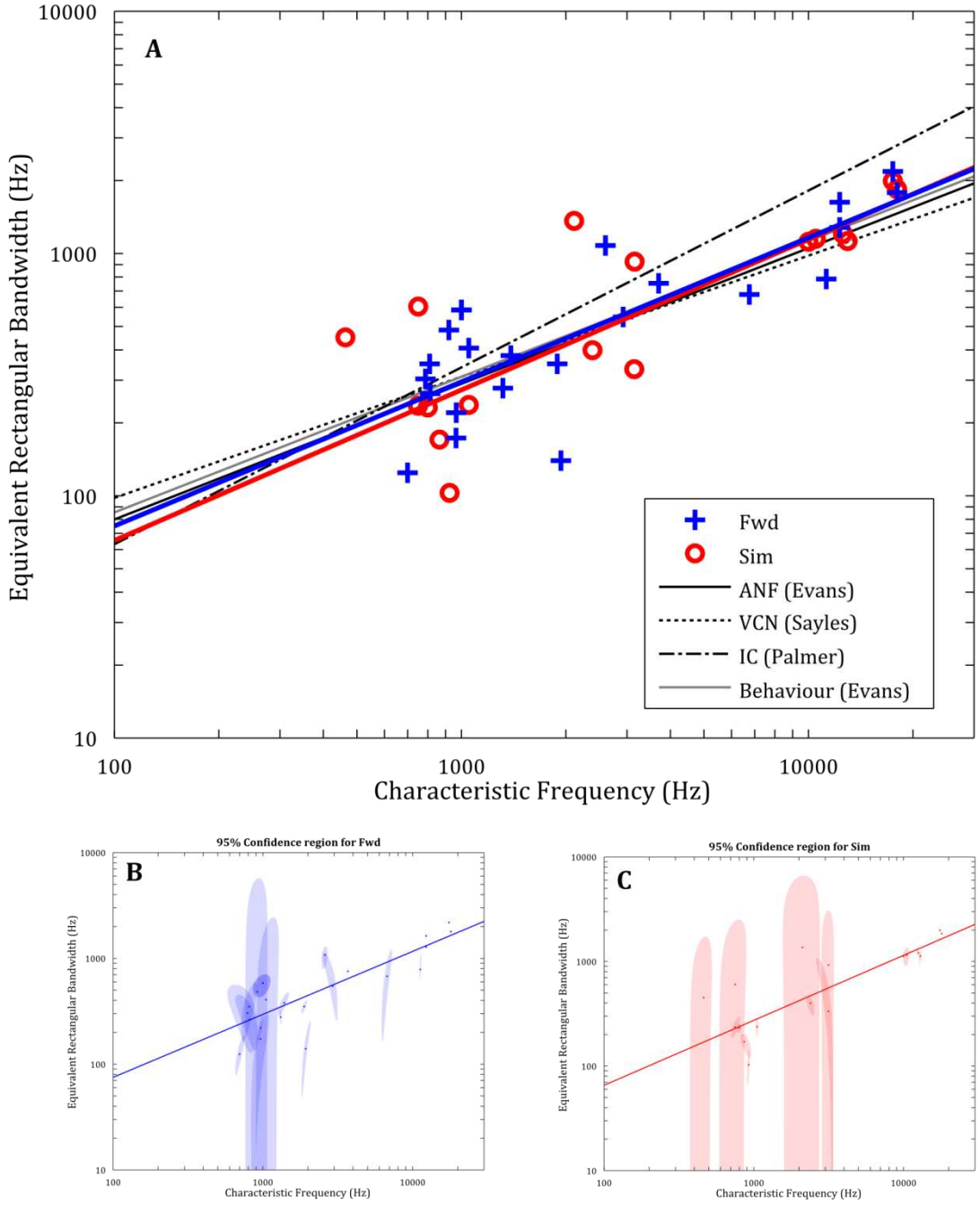


Figure 2.11

Similar to Figure 2.9 but for LFP activity. ***A*** *Forward and simultaneously masked lines of best fit,* R^2 *= 0.67 and 0.63,* χ^2_{red} *= 9.9 and 10.3 respectively.*

		$ERB = a \times CF^b$	
		a	**b**
Evans (Psychophysics)		6.45	0.56
Evans (ANF)		6.03	0.56
Sayles (VCN)		9.77	0.50
Palmer (IC)		2.19	0.73
SU	**Fwd**	2.63	0.64
	Sim	1.86	0.70
	FRA	10.47	0.60
MU	**Fwd**	3.89	0.60
	Sim	8.13	0.53
	FRA	30.20	0.51
LFP	**Fwd**	4.90	0.59
	Sim	3.72	0.62
	FRA	12.88	0.63

Table 2.2

Parameters of line describing the mean relationship between CF and ERB for the different data sets: Evans (2001), Sayles and Winter (2010), Palmer et al. (2013), and data from this chapter. CF and ERB are in Hz (hence the difference between equations described in original papers). Equation described in table is equivalent to $log_{10}(ERB) = b \times log_{10}(CF) + log_{10}(a)$.

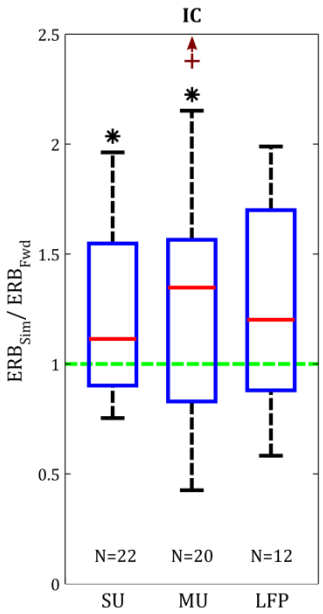


Figure 2.12

Boxplot showing the median (red line) and quartiles (blue box) of the ratio between Fwd and Sim ERB estimates in the same unit. Sim ERBs are significantly wider than Fwd ones for SU and MU (Wilcoxon signed-rank, p-value 0.03 for both) but not for LFP (Wilcoxon signed-rank, p-value 0.13). Burgundy cross represents outlier off screen.

2.3.5 Forward vs Simultaneous masking

In human psychophysics it has long been noted that forward masked notched noise ERB estimates tend to be narrower than simultaneously masked ones (Houtgast, 1974; Leshowitz and Lindstrom, 1977; Wightman et al., 1977; Moore, 1978; Vogten, 1978; O'Loughlin and Moore, 1981; Weber and Patterson, 1984; Oxenham and Shera, 2003). If this were true for the physiological recordings in this study, one might expect the lines of best fit for the simultaneously masked ERBs to sit above those for the forward masked ones. This is, however, not the case for any of the SU, MU or LFP data in Figure 2.9, Figure 2.10 and Figure 2.11. One possibility is that the difference between the two is being overshadowed by the effects of natural inter-unit variability and different stimulus parameters for each unit. Therefore to test this theory without such confounds it would be necessary to compare both measurements made in the same unit with the same stimulus parameters. Fortunately, these conditions were satisfied for 22 SU, 20 MU and 12 LFP. For each of these units the simultaneously masked ERB measurement was divided by the forward masked one, and Figure 2.12 shows box plots of the results. A value larger than 1 indicates that the Sim measurement is broader, values less than 1

mean the Fwd one is. For both SU and MU the ERB ratios are significantly larger than 1 (sign rank p-values of 0.03 each), meaning the simultaneously masked ERBs are indeed broader. There is no significant difference for the LFP data (sign rank p-value of 0.13), however, this is most probably due to the low number of recordings (only 12) making a rejection of equality unlikely.

The majority of units, however, showed even more non-linear and complex behaviour than this. Examples of the firing rate in response to the masker and signal plots for these units can be seen in the bottom panels of Figure 2.14. Teasing out the complex masking behaviour leading to these types of plots is much more difficult and beyond the capabilities of simply looking at their RLFs. It seems fairly remarkable that such units all yielded clean and clearly separated neurometric functions that gave an estimate of bandwidth.

2.3.6 Mechanisms underlying masking in IC

The benefit of using SDT to describe neural responses is that it uses both variability as well as mean firing rates to determine signal detection, as well as providing a result that is more comparable to psychometric data. A limitation is that it hides how masking occurs; is it driven by an increase in variability or a change in mean firing rate? Is the response to the signal simply being swamped by the response to the masker? Or is the response to the signal actively being suppressed? To attempt to answer these questions it becomes necessary to return to the original firing rate responses to the stimuli.

Figure 2.13 panel A shows a plot of firing rate against proportion correct, for the forward masked notched noise experiment in a representative SU. Each point represents the mean firing rate of the unit in response to the signal, in the trials when the signal was present, for a single masker condition (i.e. level and notch width). This is equivalent to taking the mean firing rate of all the red trials in Figure 2.5 panel A, over the grey dotted window shown. Each point in the plot of Figure 2.13 panel A represents this mean firing rate for one of the boxes in Figure 2.5 panel B. The colour represents the notch width. Typically in physiology it is just this mean firing rate that is used to represent the unit's response, and plots such as this provide a good comparison between the SDT and non-SDT approach. The plot in Figure 2.13 panel A follows a very clear pattern; as firing rate increases, proportion correct increases linearly from 50 %, before plateauing or 'saturating' at 100 %. This is true for all notch widths and the relationship is tight with very little variability. Every SU showed this close linear relationship. (Some did not plateau at high firing rates, but this is due to those units

never reaching 100 % correct). To show this, a LOWESS local regression model (Cleveland, 1981) was used to calculate the line of best fit for the signal firing rate versus proportion correct plot (Figure 2.13 panel A) for each SU. This is a method that gives a line of best fit for scattered data and has been used extensively in many fields of research (Karikoski et al., 1998; Seki and Eggermont, 2002; Hen et al., 2004). The mean r^2, the squared difference between a data point and the line of best fit, is calculated from this. A histogram of these 'mean difference' values is plotted in Figure 2.13 panel C. Almost all of the units (44/45) have a mean r^2 less than 0.005, showing that there is very little deviation from this mean, linear, saturating pattern.

The region of interest for this study, however, is the firing rates at threshold; activity elsewhere will have no effect on the overall outcome. To see if the lack of overall variability is also evident in this crucial region, the firing rates for each notch width of a single-unit were taken at threshold, and the coefficient of variation calculated; this is the standard deviation of the 5 firing rates divided by their mean. (The SU have such a wide range of firing rates, and higher firing rates leads to higher variance, that simply looking at the variance of these threshold firing rates does not truly explain their variability from the mean. Incidentally the inverse of the coefficient of variance is another definition of SNR (Smith, 2003)). A histogram of the coefficient of variation values for all SU can be seen in Figure 2.13 panel E. They too show a very low amount of variability, with most of the units (40/45) having a coefficient of variation value of less than 0.15. This shows that mean firing rate in response to the signal is the sole driving force behind signal detection, i.e. variability in firing rate responses across trials, or complex masking behaviour when the signal is not present, play no part. In effect, using SDT gives the same answer as the traditional approach of simply looking at mean firing rate.

The same exercise can be undertaken looking at the response of the unit to the masker, instead of the signal, against proportion correct, i.e. in Figure 2.5 panel A, calculate the firing rate over a window from 0 ms to the first dotted grey line, instead of between the two dotted grey lines. Such an investigation would show how closely a unit's response to the masker is related to the amount of masking. For example, if masking were solely due to depletion of neurotransmitters, and changes in ionic gradients that underlie action potentials, then one would expect there to be an inverse linear relationship between signal detection and response to the masker. In this case one would expect a mirror image of the plot in Figure 2.13 panel A. For the unit in panel A, the masker firing rate against proportion correct plot is shown in panel B. In this plot there is a clear and systematic spread across notch widths, with the widest notch reaching threshold at the

lowest firing rate. This 'spread', which is not seen in the signal response, is seen in varying degrees for all units when looking at the response to the masker. As a result both the variability around the line of best fit (Figure 2.13 panel C) and across notch width at threshold (Figure 2.13 panel E) is larger than for the signal (this is shown by a higher mean and broader range of values).

A potential confound between the results of looking at the response to the signal and the response to the masker, is that the length of the window over which the firing rate is calculated is different. For example, the signal is presented for 10 ms and the masker for 150 ms, so the firing rates in response to the masker will have much higher variability since there is more time for spikes to occur. To check whether this has an effect on the results, the spike count over a short window (the same length as for the signal in the same trial) at the start of the masker presentation was taken. Results were exactly the same as seen in Figure 2.13 panels C and E. The process was repeated for a short window at the end of the masker presentation, to see how response to the masker directly before signal presentation affected proportion correct. For onset units this region had a firing rate of 0 and so showed no discernible relationship. For the sustained units, all firing rate vs proportion correct plots (Figure 2.13 panel B) were shifted towards lower firing rates, but the variability plots were almost identical to those in Figure 2.13 panels C and E.

The plots in panels A and B came from a forward masked experiment in the same unit, and showed typical behaviour for all SU. The r^2 and coefficient of variance values from this SU appear in the bars marked by the arrows in panels C and E; they are typical for all SU. Results for the simultaneous masked arrangement were also very similar, with tight agreement between signal firing rate and proportion correct, and a much looser relationship when looking at the masker firing rate (panels D and F). When looking at the variability of the entire curve (panel D), the relationship for the signal response is not quite as close as elsewhere, probably because it actually represents the signal + masker response (since masking is simultaneous) and the masker response shows higher variability. This increased variability is, however, not seen at detection threshold (panel F). So for simultaneous masking, as well as forward masking, the mean firing rate in response to the signal is the only important feature of neural activity for measuring detection threshold. In other words using the RLF in place of the neurometric function gives the same measurement of bandwidth.

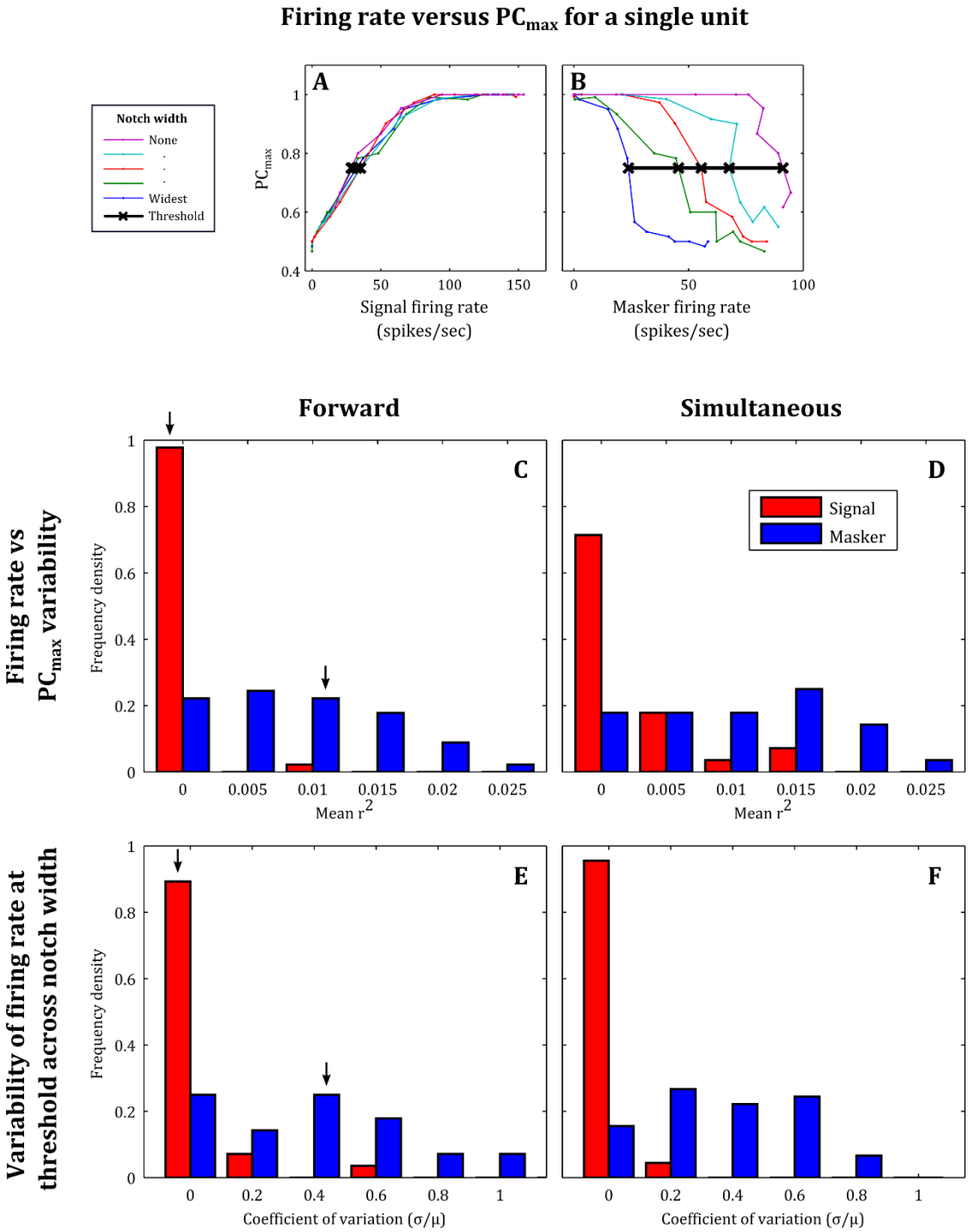


Figure 2.13

Firing rate versus proportion correct. ***A*** *&* ***B*** *Proportion correct against firing rate in response to the signal and masker, respectively, for a single-unit. Results for each notch width plotted separately. × = firing rate at threshold (75 % correct).* ***C*** *&* ***D*** *Histogram of mean* r^2 *difference between LOWESS trend line and data, for firing rate vs* PC_{max} *plots (curves in panels A & B).* ***E*** *&* ***F*** *Histogram of coefficient of variation for firing rates of different notches at threshold (× in panels A & B). Red = signal firing rates, blue = masker firing rates. Arrows indicate where the statistics of the unit in panel A & B appears. See text for more details.*

Looking now at just the response to the masker and signal (ignoring proportion correct), we can learn more about the nature of masking that is occurring. In the case where all the masking seemed to be occurring within the unit, one would expect an inverse linear relationship between the two, i.e. as the response to the masker increases, the response to the signal should decrease by the same amount. Figure 2.14 panel C shows a plot for such a unit. Only a small proportion of the SU (11/45 in Fwd, 5/28 in Sim) showed this simple linear relationship between the two. We can further investigate the cause of this type of behaviour, and the nature of masking, by looking at RLFs in response to the masker and to the signal, for different stimulus conditions (see 1.8.3 Rate against level function for more details and examples). Figure 2.14 panels A and B show such an RLF, showing the same firing rates plotted in panel C, only plotted against masker level instead. Looking at the response to the signal first (panel A), there is a clear shift in the RLF towards higher masker levels with increasing notch width. This shift is in the direction that one would expect since larger notches means less masking, and therefore a higher masker level is required to elicit the same firing rate. Also the fact that there is this systematic shift, and the floor of the RLFs do not rise with masker level, suggests that masking is suppressive and not excitatory. Forward masking was used in this example (and the same pattern was seen in every forward masked unit (see Figure 2.14 panel D for another example)) so the result is relatively unsurprising. However, the systematic shift was also seen in every simultaneously masked case, with only 4/28 showing a rise in the floor of the RLF with increasing masker level; for all the others the floor was constant across notch width. So only a very small proportion showed any signs of excitatory masking.

Since every RLF of the response to the signal has this systematic shift with notch width, the only way a 'linear' masker vs signal firing rate plot could occur (Figure 2.14 panel C) is if the RLF of the response to the masker has the same systematic shift, but in the opposite direction. Figure 2.14 panel B shows such a plot and is the link between the plots in panels A and C. Above, it was shown that using the RLF of the response to the signal in place of a neurometric function would yield the same estimates of bandwidth. In the case of the 'linear' unit described here, this would mean that using the RLF of the response to the masker would also yield the same answer, since the salient feature is the difference in SNR between notch widths, and this is the same for both RLFs. Such a unit would therefore display linear behaviour overall since bandwidth measurements would be invariant to method. This is, however, only true for a small proportion of the units, represented by those which fall into the '0' bar in Figure 2.13 panels C-F.

The majority of units showed a large variation across notch width in their firing rate in response to the masker vs proportion correct plots (as described above and shown in Figure 2.13 panel B). This too was reflected in their firing rate in response to the masker vs signal plot; an example is shown in Figure 2.14 panel F. The same systematic spread or 'separated' curves for different notch width is seen here. In total 33 % (15/45) forward and 32 % (9/28) simultaneous masked units displayed this clear systematic spread; 71 % (5/7) of these units showed a 'spread' in both forward and simultaneous temporal arrangements. For all of these, the reason for this spread was an invariance across notch width in RLF for the response to the masker, an example of which can be seen in Figure 2.14 panel E. For these units the response to the masker seems to be frequency invariant, and suggests that most (if not all) of the masking driving frequency selectivity has occurred before reaching this unit. This may be due to network interactions, such as inhibitory inputs directly into the unit, or masking occurring more peripherally and the excitatory input to the unit already showing the frequency selectivity. Estimates of bandwidth made by looking at the excitatory input of the masker would effectively be infinite. The behaviour of these units is therefore non-linear.

A higher proportion of units, however, showed much more complex firing rate in response to the masker vs signal plots, some of which can be seen in the lower panels of Figure 2.14. These represent 42 % (19/45) of the forward and 50 % (14/28) of the simultaneously masked units. Much more complicated masking interactions occur for these units than can be extracted by looking at the RLF and firing rate plots, reflecting a much higher degree of non-linearity. It is remarkable that such units still display frequency selectivity, and yield estimates of bandwidth. There also appears to be no relationship between CF and ERB (and their confidence intervals (CI)) of a unit, and these classifications.

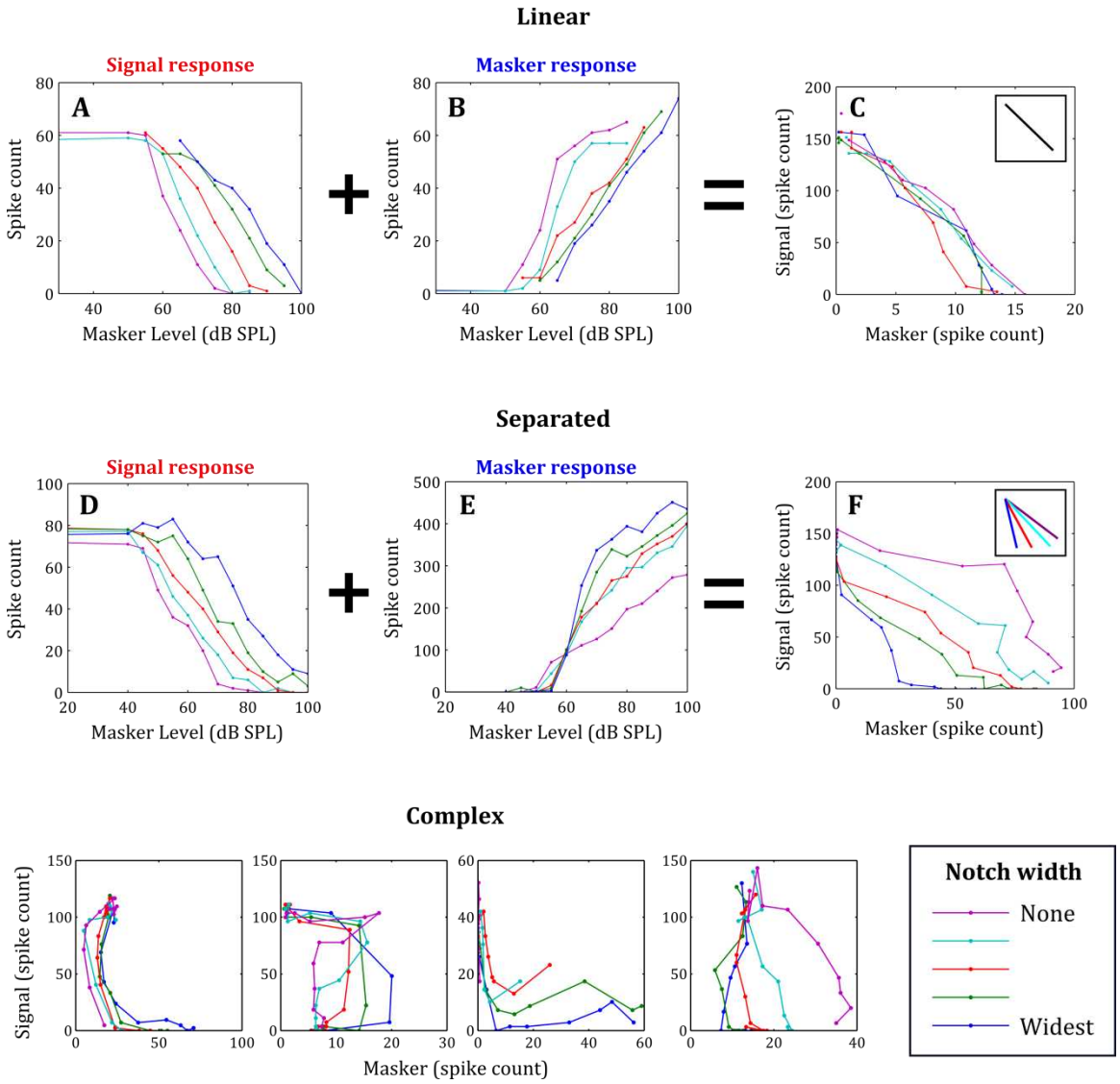


Figure 2.14

RLF and signal vs masker firing rate plots. ***A*** *RLF of response to the signal for different notch widths.* ***B*** *Same as A but with response to the masker.* ***C*** *Firing rate in response to the signal plotted against that of the masker. All are from an SU with 'linear' type firing rate plot.* ***D****,* ***E*** *&* ***F*** *are the same but for a unit with 'separated' type firing rate plot. Lower panels depict units with 'complex' type firing rate plots.*

2.3.7 FRA based bandwidth

An FRA was recorded for each unit to find its CF in order to tailor the notched noise stimuli to it. However, as mentioned previously, it is possible to obtain an independent bandwidth estimate using this information. In total ERB estimates were obtained in such a way from 139 SU (in 21 animals), 124 MU (in 17 animals) and 107 LFP sites (in 12 animals) and these can be seen in Figure 2.15 panels A, C and D. To aid comparison to the notched noise results, the lines of best fit for the respective forward and simultaneously masked ERBs have also been plotted for each data type. Just like the NN data, the FRA derived estimates have a clear linearly increasing relationship, in log-log space, between CF and ERB, and the rate of increase is similar across the board. The most striking difference between the two measurement types is that the FRA ERBs are much broader than the NN ones; on average FRA ERBs are 2.7 times broader than NN for SU, 3.3 for MU and 3.5 for LFP. The variability is also higher with R^2 values of 0.39, 0.46, and 0.51 for SU, MU and LFP respectively. This result is in contrast with previous studies in the IC, VCN and ANF that use the FRA method (shown in panel A of Figure 2.9, Figure 2.10 and Figure 2.11), which show much more consistent results with one another and the NN data. One slight difference in method between the Palmer et al. (2013) and this one, is that Palmer et al. did not fit a roex function to the FRA. Instead they treated the FRA as the AF and calculated ERB directly by integrating the total amount of energy passed by the FRA (i.e. calculated its area) and scaling this by the threshold of the unit. To see what effect this methodological difference has on the estimates of bandwidth, ERBs for my data were recalculated using this 'no roex' approach. These can be seen plotted against the individual data points from Palmer et al. (2013) in panel B of Figure 2.15, as well as the 95 % confidence regions of the lines of best fit plotted inset. The lines of best fit for both data sets have an identical slope of 0.74, but are slightly offset. This difference, although significant for 300 - 10,000 Hz, is very small, and is completely overshadowed by the large variance in each data set.

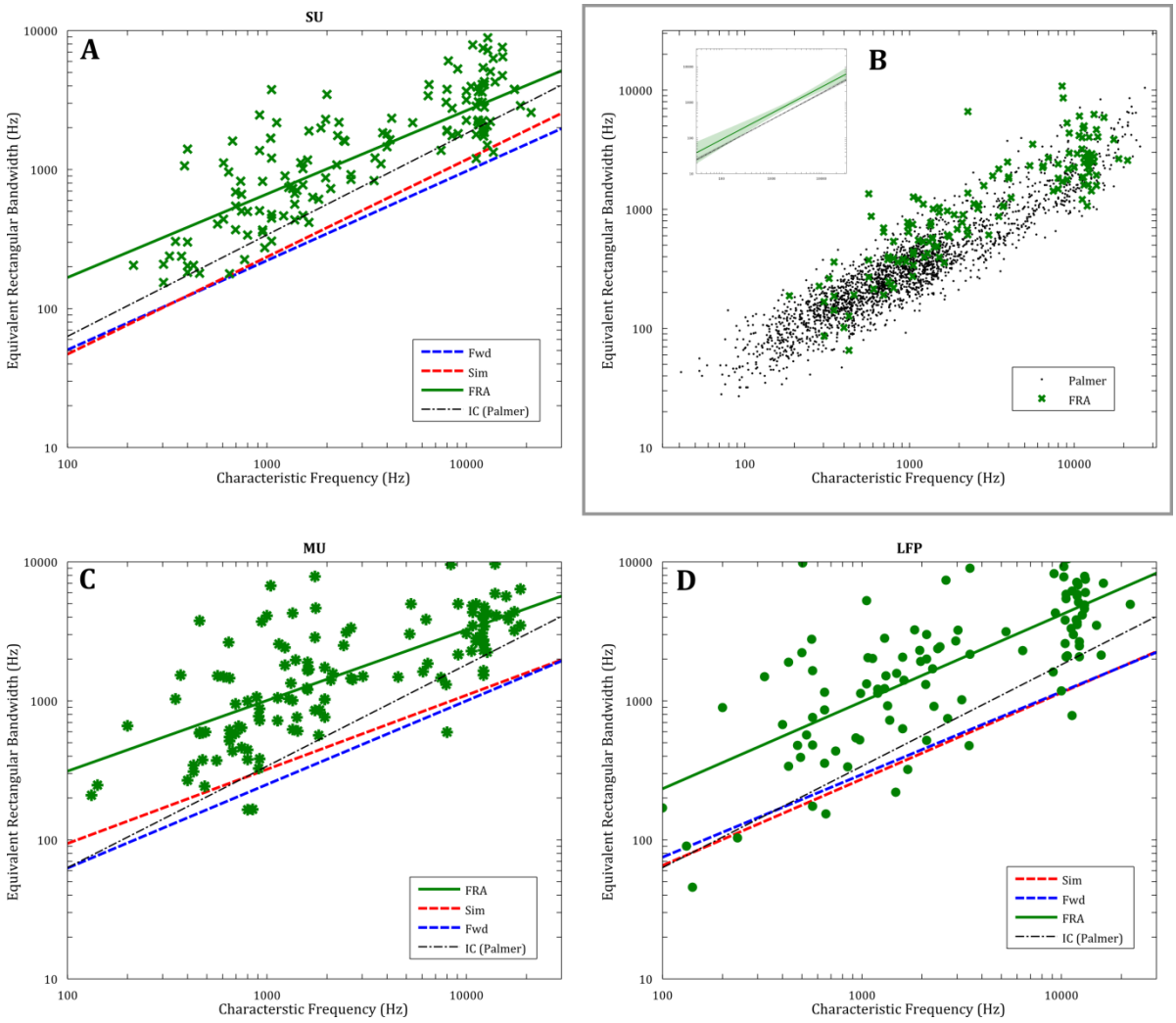


Figure 2.15

*ERB estimates from FRAs of **A** single-units, **C** multi units, **D** LFPs (green points) with line of best fit (green line; R^2 = 0.39, 0.46 and 0.51 respectively), and lines of best fit for forward (blue) and simultaneous (red) notched noise data, for each data type. Also included is the fit for the results from Palmer et al. (2013). **B** Comparison of individual SU ERBs and data from Palmer et al. (2013) using the exact method in the Palmer paper. Inset shows the lines of best fit and 95 % confidence region for each.*

Another difference between the two methods may be in the exact parameters used to determine the edge of the FRA, i.e. the criteria which define a significant increase in firing rate of the unit. As described in *2.2.7* Analysis*:* FRA analysis above this threshold is determined by parameters of various spiking statistics such as 1) minimum spike count, 2) mean, and 3) standard deviation of the firing rate. Differences in these parameters might have an effect on the average relationship between CF and ERB. To test this, ERBs were recalculated from the same FRA using a number of different threshold parameters; in total 80 separate sets of parameters were used, yielding 80 different lines of best fit. The mean line of best fit, and the 95 % confidence region, can be seen plotted in Figure 2.16. Different threshold parameters clearly give a wide range of relationships between CF and ERB. Since the line from Palmer et al. (2013) sits within the 95 % confidence region, they cannot be said to be significantly different. Also, remarkably, roex functions were fitted to the FRA to obtain these bandwidth estimates suggesting that threshold parameter choice plays a larger role.

One final check to see what effect stimulus parameters and guinea pig breed might have on the FRA bandwidth results, was made by collecting FRA from IC units recorded by other researchers at the Institute (Steadman and Mill, unpublished). These were recorded from the central IC of tricolour guinea pigs (as opposed to albinos used in my experiment) with glass-coated tungsten and Neuronexus electrodes. Stimulus parameters, however, were not the same, with pure tone signals being played for 50 ms (as opposed to 10 ms used in my experiment). The same bandwidth analysis, with identical parameters, was applied to these FRA, and the results can be seen in Figure 2.17, separated into SU and MU. (Unfortunately the raw voltage traces were not recorded for these units so no LFP measures could be made). Quite clearly there is no difference between these bandwidth estimates, suggesting parameters determining threshold have a much larger impact than these factors.

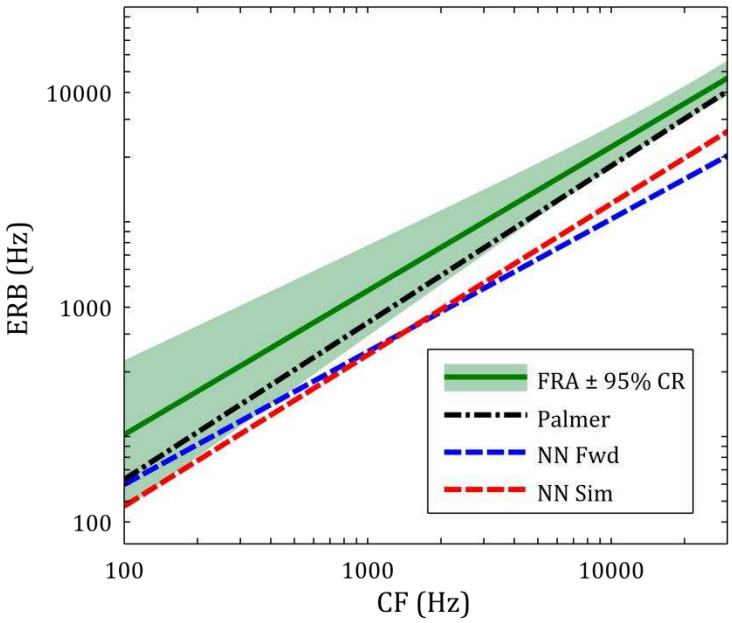


Figure 2.16

Effect of threshold parameters on average FRA calculated bandwidths. The green line represents the average line of best fits for ERBs estimated from the same FRAs but with 80 different combinations of threshold parameters. Shaded area shows 95 % confidence region. Over plotted are the lines of best fit for the results from Palmer et al. (2013) and the NN ERBs shown in Figure 2.9.

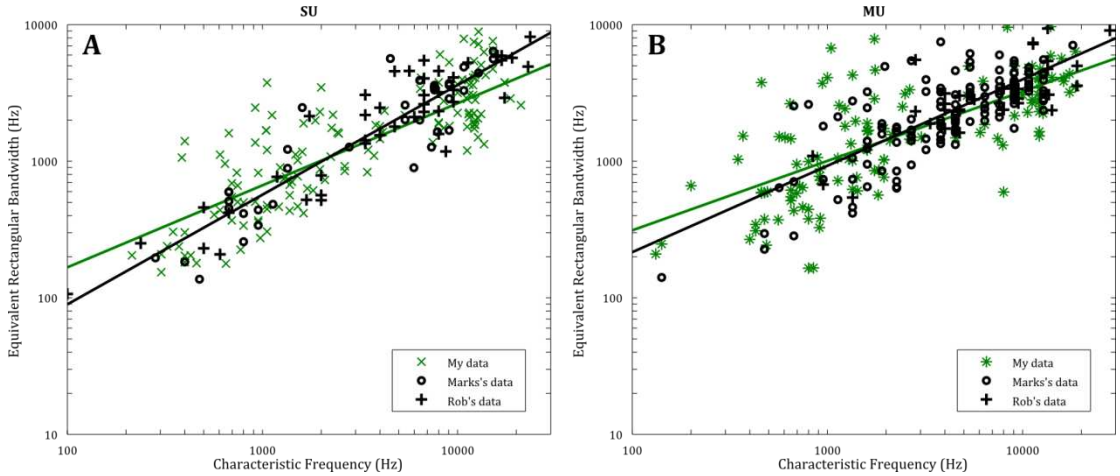


Figure 2.17

SU and MU plots from Figure 2.15 compared with ERBs made from FRAs recorded by other experimenters in the lab. The method for calculating the ERB is consistent across all points, however the stimuli used to derive the FRA differs between experimenters.

2.4 Results - A1

In total successful recordings were made from 77 single-units and 120 multi-units (61 of which were extracted from SU recordings) in A1.

2.4.1 Unit characteristics

The response characteristics of all the units that yielded a notched noise ERB estimate were investigated and units were split into three categories depending on whether they fired only to the onset of the masker or showed a sustained response. There were also a number of units that showed onset behaviour for high masker levels which gradually changed into a sustained response closer to the threshold of the signal. These results are summarised in Table 2.3. The vast majority of units only responded to the onset of the masker for SU, with the proportion falling slightly for MU.

Threshold was calculated for each unit, defined as the lowest level at which the unit had a firing rate more than 3 standard deviations above its mean spontaneous firing rate. Assuming a normal distribution this equates to a 99.73 % confidence of elevated firing. Figure 2.18 A and C show these thresholds plotted against the CF of the unit for SU and MU respectively. The thresholds and CF of units are spread fairly evenly across the range of hearing of the guinea pig, defined by the superimposed audiograms. Perhaps the SU units tend to have higher thresholds than the MU, although this may be due to the larger number of MU points. In addition, panels C and D show the level of the signal in the notched noise experiments against the CF of the unit. These levels are on average 10 dB above the thresholds of SU (mean 10.04 dB, standard deviation 10.88 dB) and 11 dB for MU (mean 10.75 dB, standard deviation 11.16 dB) . The majority of units gave forward masked ERBs but those that also gave simultaneous estimates are distributed evenly, meaning there is no bias of threshold or CF in the notched noise results.

	SU		MU	
Onset	27	(87 %)	29	(63 %)
Sustained	2	(6.5 %)	11	(24 %)
Both	2	(6.5 %)	6	(13 %)

Table 2.3

Number of units that yielded a notched noise ERB with particular response characteristics. Units characterised according to their response to the 150 ms long masker. See text for further detail.

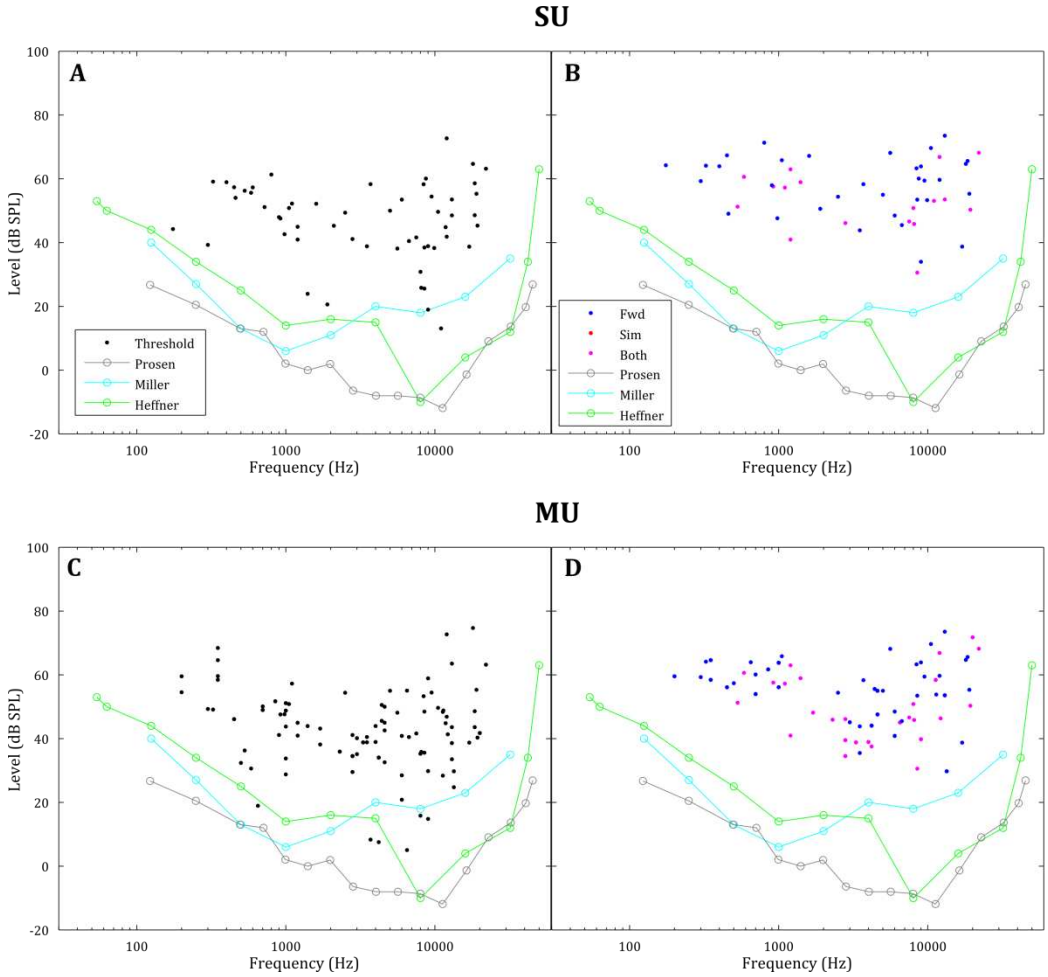


Figure 2.18

A *&* ***C*** *Threshold of units derived from rate level functions plotted against CF, for SU and MU respectively.* ***B*** *&* ***D*** *Fixed level of signal for each notched noise experiment; blue = units with just a forward masked ERB, red = just simultaneously masked, magenta = both. Audiograms of the guinea pig taken from Miller and Murray (1966), Heffner et al. (1971) and Prosen et al. (1978).*

2.4.2 SU results

The full set of results for the SU data can be seen in Figure 2.19. In total 30 forward and 12 simultaneously masked notched noise and 77 FRA calculated bandwidth estimates were made from SU activity in 25 animals. Just like in the IC results, there is no obvious difference between the forward and simultaneous notched noise estimates in panel A. Unlike the IC, however, this is supported by a non-significant difference between the two measures within the same unit (panel B). Variability is also greater, reflected in the much lower R^2 values of 0.27 and 0.09, for Fwd and Sim respectively. This suggests that units in cortex are not as homogeneous as in the IC, which is unsurprising given the

functional differences between the two brain regions. The surprisingly poor fit for the simultaneous data could, at least in part, be down to the low number of data points. Unfortunately since the surgical preparation is much less stable in cortex, it is much harder to record from a unit for any length of time. Couple this with the almost doubling in experiment duration (roughly 3 hours), and it became very difficult to record both NN methods before losing the unit. As a result the forward masked arrangement was favoured and only a small number of successful simultaneous estimates were obtained. Not many strong conclusions can therefore be made from the simultaneous data.

The most notable feature of the NN data is ERBs that are much smaller for units with CF above ~2 kHz than expected or seen elsewhere. The IC NN data reflected closely the bandwidth estimates made in the studies of Palmer et al. (2013), Evans (2001) and Sayles and Winter (2010). However, here units around 10 kHz have ERB estimates roughly a third the size of all other measurements seen so far. What's more, this feature is not an aberration, since these data points actually have the smallest overall confidence regions (panel C). It almost seems like a linear fit may not be appropriate since low CF and high CF units are displaying quite different characteristics; large ERBs for units around 1 kHz deviate substantially from the line of best fit, although these points have very large confidence regions so should be largely ignored. Nonetheless the group of finely tuned high frequency units give the line of best fit a much shallower slope than for IC. Table 2.4 shows that the slopes are 0.32 and 0.24 for the forward and simultaneous NN data respectively. These are far shallower than for the studies mentioned above; Sayles and Winter (2010) has the shallowest slope with 0.50.

However, the ERBs estimated using FRAs are once again broader than the NN estimates. Moreover they display almost identical lines of best fit to the FRA estimates made using IC units; the slope of the A1 is 0.55 (Table 2.4) and that of IC data is 0.60 (Table 2.2). This gives them a steeper slope than the NN estimates for the cortical data (panel A), meaning the different methods of calculating bandwidth in the same population of units give qualitatively different results.

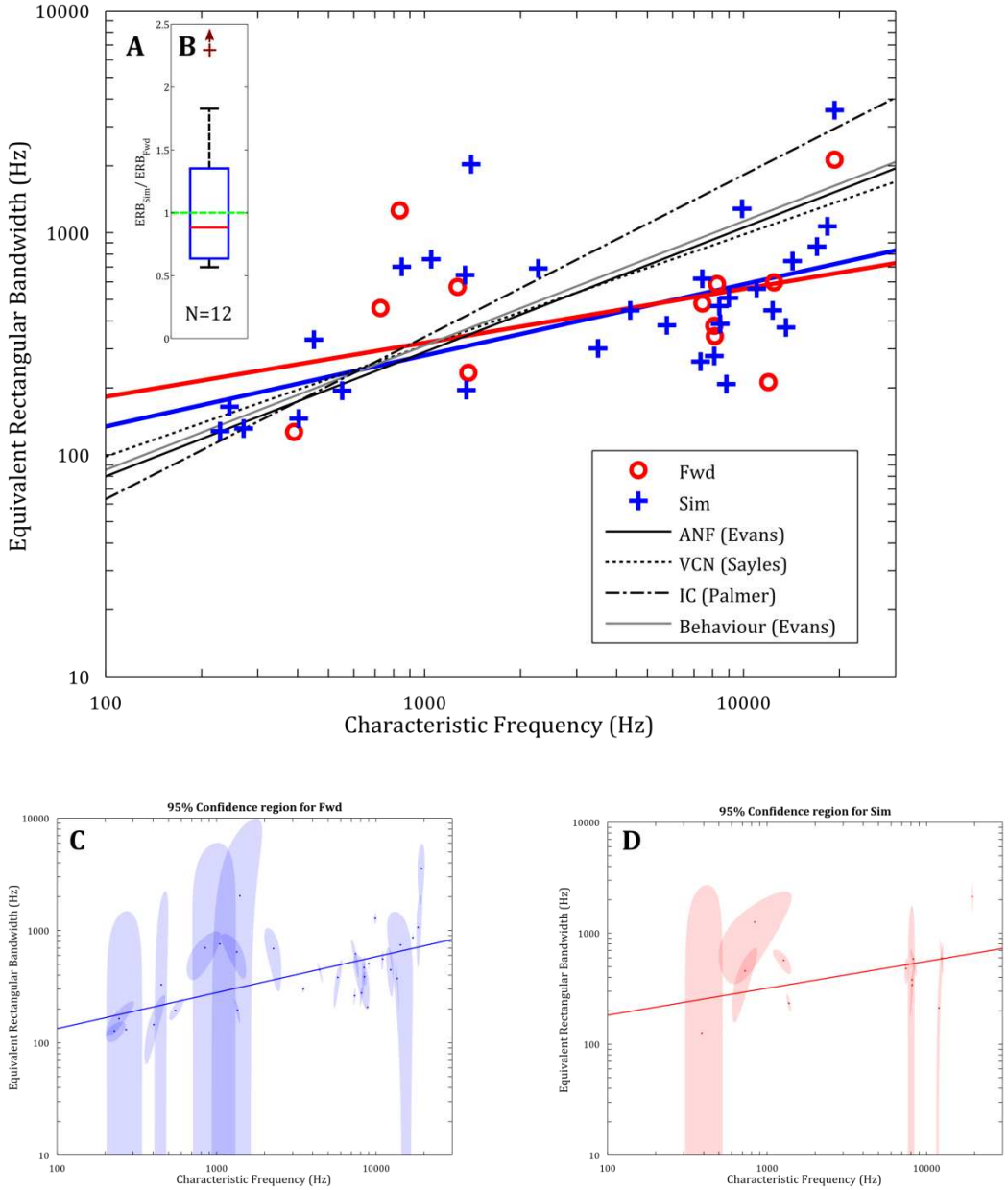


Figure 2.19

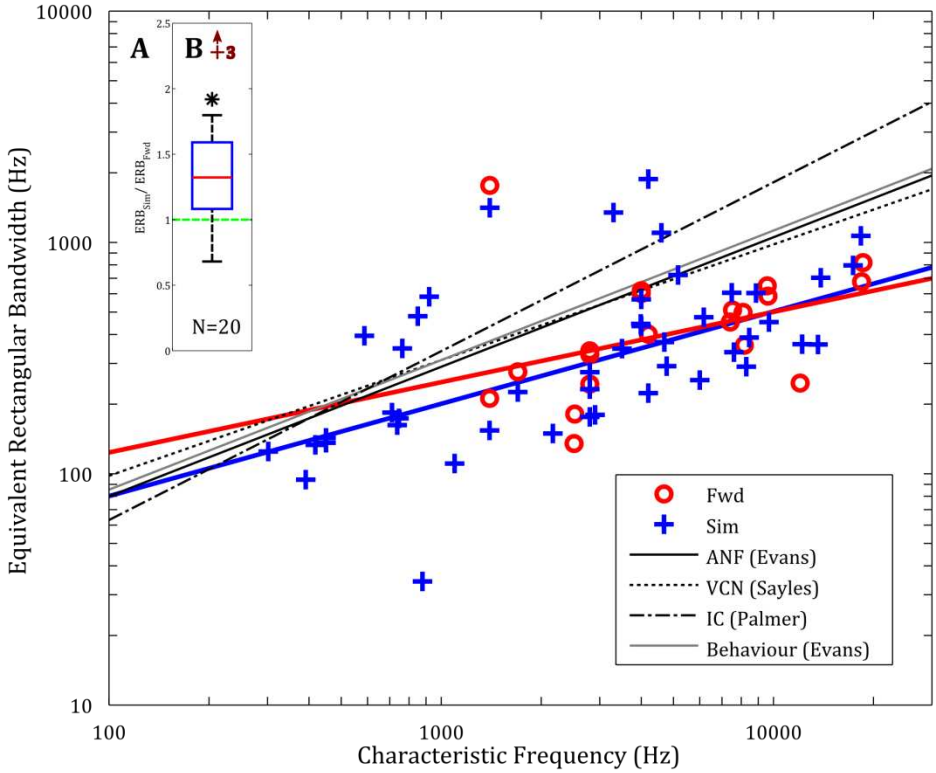
A *Forward and simultaneously masked notched noise ERB estimates plotted against CF of the SU (blue +, red o respectively) with lines of best fit (R^2 = 0.27 and 0.09, χ^2_{red} = 20.3 and 13.0 respectively). CF vs ERB plots for SU in the guinea pig from (Evans, 2001), (Sayles and Winter, 2010), (Palmer et al., 2013). Also for guinea pig psychophysics (Evans, 2001).* ***B*** *Box plot showing the median (red line) and quartiles (blue rectangle) of the ratio between Fwd and Sim ERBs in the same unit. No significant difference (Wilcoxon signed-rank, p-value = 1). Burgundy cross represents outliers.* ***C*** *and* ***D*** *Data points in panel A with respective 95 % confidence regions calculated through bootstrapping.* ***E*** *ERB estimates from FRAs, with line of best fit (green line; R^2 = 0.65), and lines of best fit for forward (blue) and simultaneous (red) notched noise data of the respective type.*

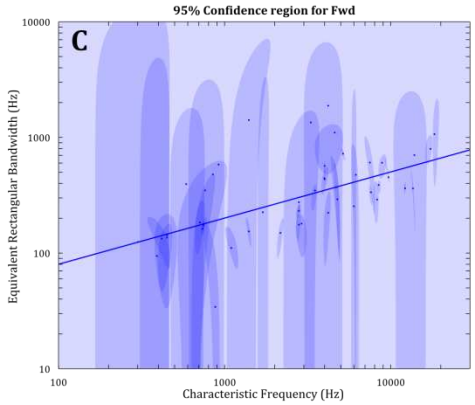
2.4.3 MU results

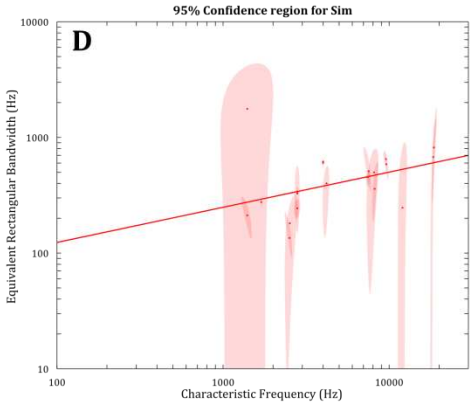
In total 46 forward and 20 simultaneously masked notched noise, and 120 FRA calculated estimates were made using MU from 28 animals. These can be seen in Figure 2.20. Panel A seems to suggest there is no difference between the two temporal arrangements of the NN methods, since they have very similar mean trends, variability and confidence regions around each point. However, the within unit comparison in panel B shows that simultaneous estimates are significantly broader.

The spread of data points in panel A is fairly similar to that of the SU, reflected in the similar R^2 values for the lines of best fit; both are ~0.3. The confidence regions of the data points are also much larger than in the SU, reflected in lower χ^2_{red} values, so we can be less sure of these results. Most notably, however, the smaller than expected ERB estimates seen with SU seem to not only be limited to high frequency units, but seems to be true for units of all frequencies; the large majority of forward masked data points fall below the lines of best fit for all the previous studies. This can be seen in the steeper slopes and larger intercepts; the mean NN data for A1 have slope 0.35 and intercept 21.5, while for IC slope is 0.28 and intercept 45.6.

Once again the FRA results are qualitatively different to the NN ones. In fact the results are very similar to those of the SU, with R^2 of 0.64 and 0.65, and slope of 0.59 and 0.55 for MU and SU respectively.







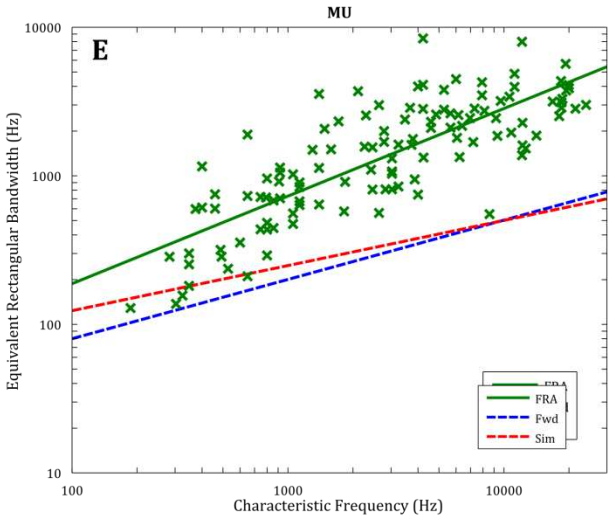


Figure 2.20

Similar to Figure 2.19 but for MU activity. ***A*** *Forward and simultaneously masked lines of best fit,* R^2 *= 0.31 and 0.09,* χ^2_{red} *= 8.1 and 23.9 respectively.* ***B*** *Significant difference (Wilcoxon signed rank p-value < 0.01).* ***C*** *and* ***D*** *Data points in panel A with respective 95 % confidence regions.* ***E*** *FRA estimated ERBs,* R^2 *= 0.64.*

2.4.4 LFP results

In total, 39 forward and 15 simultaneously masked notched noise, and 134 FRA calculated estimates were made using LFPs from 25 animals; Figure 2.21 shows results for these. Once again panel A shows no difference between forward and simultaneously masked estimates, and like the SU this is reflected in no significant difference between the two measures within the same unit (panel B). The trend of high CF units having the same fine tuning and low CF units' ERBs getting smaller, that we saw when moving from SU to MU, continues here as we move to LFP. Even fewer of the forward masked notched noise estimates sit above the data of Palmer et al. (2013), Evans (2001) and Sayles and Winter (2010), and the slope of the line of best fits grows steeper still, to 0.49 and 0.64 for forward and simultaneously masked ERBs, respectively (Table 2.4). A large contributing factor to this steepening of slope, could be the small population of high frequency units (~20 kHz) that have very broad bandwidth estimates. Without these the line of best fit would be shallower and lower. The data points are also more tightly clustered together, resulting in a stronger linear fit (R^2 value of 0.42) than either the SU or MU. Confidence intervals for each of these points are also smaller than MU so we can be surer of the result.

The FRA calculated ERBs display the now familiar trend of being much broader than the NN ones. Here this difference is highlighted even further with FRA estimates being on average 4 times greater than NN, with a larger spread in values leading to a poorer goodness of fit (R^2 of 0.51, as opposed to 0.65 and 0.64 for SU and MU). Additionally, the FRA measured bandwidths are also broader than those for SU and MU, by ~50 %. The rate of increase of ERB with CF is similar across all 3, reflected in the very similar slope values shown in Table 2.4 (0.55, 0.59 and 0.57 for SU, MU and LFP respectively), but the actual size is larger, reflected by the different intercept values in the same table (15.85 and 12.30 for SU and MU, respectively, against 24.55 for LFP).

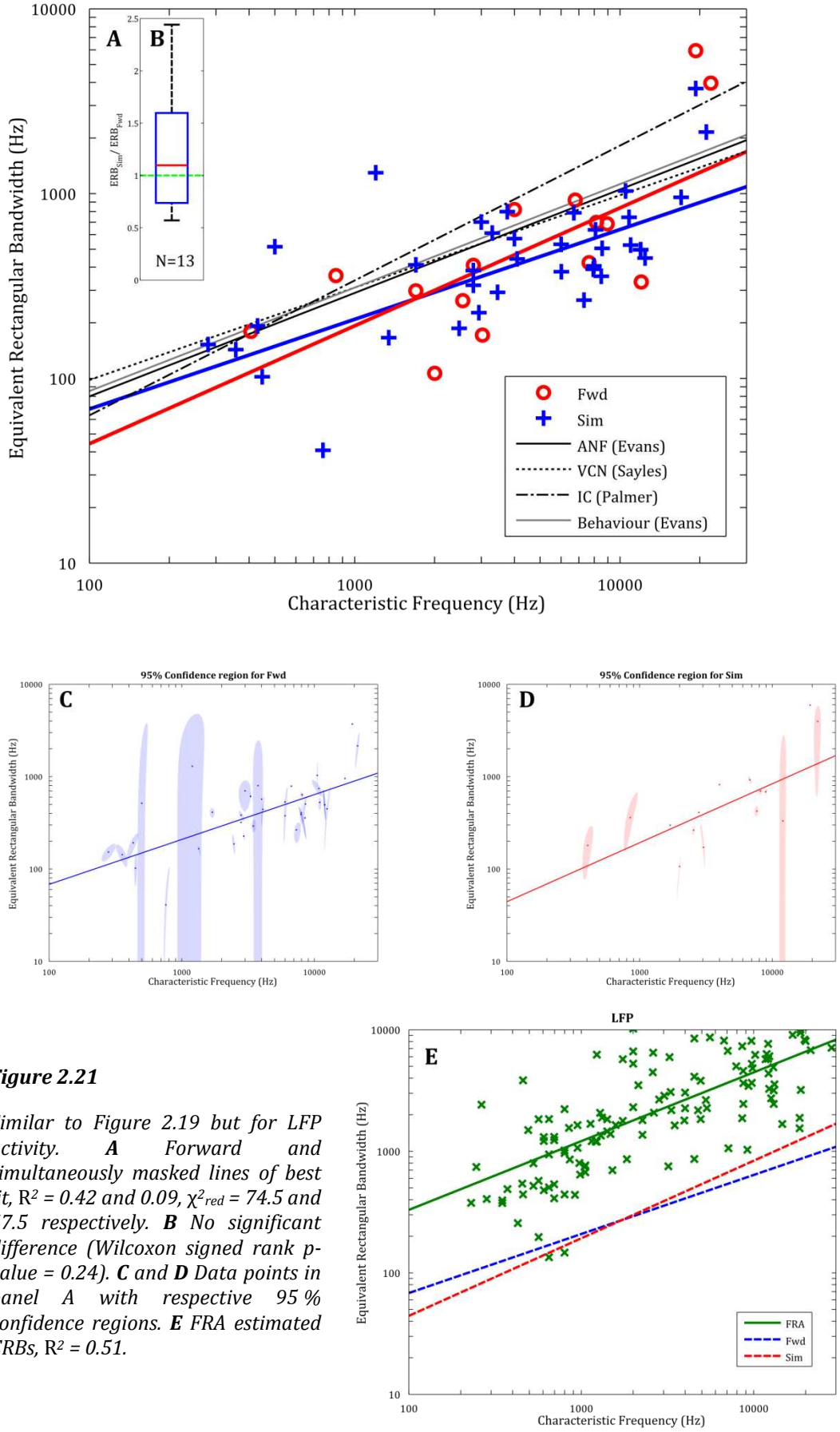


Figure 2.21

Similar to Figure 2.19 but for LFP activity. ***A*** *Forward and simultaneously masked lines of best fit,* R^2 *= 0.42 and 0.09,* χ^2_{red} *= 74.5 and 67.5 respectively.* ***B*** *No significant difference (Wilcoxon signed rank p-value = 0.24).* ***C*** *and* ***D*** *Data points in panel A with respective 95 % confidence regions.* ***E*** *FRA estimated ERBs,* R^2 *= 0.51.*

		$ERB = a \times CF^b$	
		a	**b**
Evans (Psychophysics)		6.45	0.56
Evans (ANF)		6.03	0.56
Sayles (VCN)		9.77	0.50
Palmer (IC)		2.19	0.73
SU	**Fwd**	30.90	0.32
	Sim	60.26	0.24
	FRA	15.85	0.55
MU	**Fwd**	12.88	0.40
	Sim	30.20	0.30
	FRA	12.30	0.59
LFP	**Fwd**	7.24	0.49
	Sim	2.34	0.64
	FRA	24.55	0.57

Table 2.4

Parameters of line describing the mean relationship between CF and ERB for the different data sets: Evans (2001), Sayles and Winter (2010), Palmer et al. (2013), and data from this chapter. CF and ERB are in Hz (hence the difference between equations described in original papers). Equation described in table is equivalent to $log_{10}(ERB) = a \times log_{10}(CF) + log_{10}(b)$.

2.4.5 FRA bandwidth corroboration

As with the IC data, the difference between FRA and NN estimated ERBs is large enough to warrant a more detailed look at what might be the cause. To do so I took FRAs recorded from A1 units in the guinea pig, using glass-coated tungsten electrodes, by another researcher at the Institute (Sollini, unpublished). I then calculated their bandwidth, using the same method as described in 2.2.7 Analysis*:* FRA analysis. This comparison can be seen in Figure 2.22. Once again, due to the difficulties of recording from auditory cortex, there are far fewer SU than MU. Also their CFs are higher than the majority of mine, since the researcher was only interested in high frequency units, which means unit characteristics must be extrapolated. Nevertheless, the mean trend and spread of these data are identical to those of the FRA estimated bandwidths from my own units. Even more compellingly, the stimulus used to create the FRAs in this new

data set were not the same as mine; my FRAs were made using 10 ms tone pips and his with 100 ms. One might therefore expect differences between bandwidth measures, but there are none. Once again we can conclude that there is nothing unusual about the units I have recorded that might explain this unusual finding.

2.4.6 Mechanisms underlying masking in A1

Masking is the main difference between the FRA and NN methods, but how is this masking manifested in the unit, and what causes it? To investigate this, a look at the firing properties of the unit is required.

As with the IC data, firing rates in response to the signal were plotted against proportion correct for each unit. Once again there was a very close relationship between the two (if slightly messier than in IC). An example can be seen in Figure 2.23 panel A. This close relationship was similar for all SU, as can be seen in the signal bars in panels B and C of Figure 2.23; panel B describes the average least squared difference between the data points and their LOWESS regression line (see 2.4.6 Mechanisms underlying masking in IC), giving a measure of variability of the entire curve; panel C shows the coefficient of variation of firing rates at threshold, across notch widths (i.e. the standard deviation of the 5 firing rates at threshold for the different notch widths, is divided by their mean), giving a value of variability at threshold. All the variability values are small for the response to the signal, very similar to those in the IC, with 29/31 having an r^2 value of less than 0.005 and a slightly higher proportion of 19/31 having a coefficient of variation value less than 0.15. The variability at threshold is slightly higher here than in IC, which reflects the slightly messier plots. This close relationship between signal response and proportion correct once again shows that mean firing rate fully determines detection thresholds (and therefore bandwidth estimates) and that variability and response in the absence of a signal have no effect.

Performing the same analysis but on the response to the masker instead of signal, shows a much broader range of activity. The r^2 plot is very similar to that of the IC, but as with the response to the signal, there is higher variability at threshold; in the IC the highest coefficient of variation was 1.1, whereas in A1 it is 1.9. This shows that a unit's response to the masker has even less in common with detection thresholds of the signal, and therefore masking, than in the IC.

Plotting the response of a unit to the masker against that of the signal can tell us about the nature of masking in that unit. In the IC such plots could be categorised by

stereotypical shapes, which described the unit's masking activity as linear or varying amounts of non-linear. Unfortunately, similar plots for A1 units are all so complex and messy that such classification is impossible. Not one of the 31 forward masked and 12 simultaneously masked units gave a 'linear' type plot as described in 2.4.6 Mechanisms underlying masking in IC, and only one could be said to display a 'separated' type plot. The lower panels of Figure 2.23 give examples of these. Such bizarre behaviour is almost certainly the result of a complex arrangement of excitatory and inhibitory units all the way up the auditory system and throughout auditory cortex. This is unsurprising since there are at least 5 synapses between the BM and cortical neurons; plenty of opportunity for complicated interactions. However, as with the complex neurons in IC, this behaviour is driven mainly by the response to the masker; the RLFs of the response to the signal are clean and show clear systematic separation between successive notch widths. It is this feature that determines the units' frequency selectivity and gives an estimate of bandwidth. The question of whether this activity is linear enough for the underlying assumptions of the power spectrum model remains open.

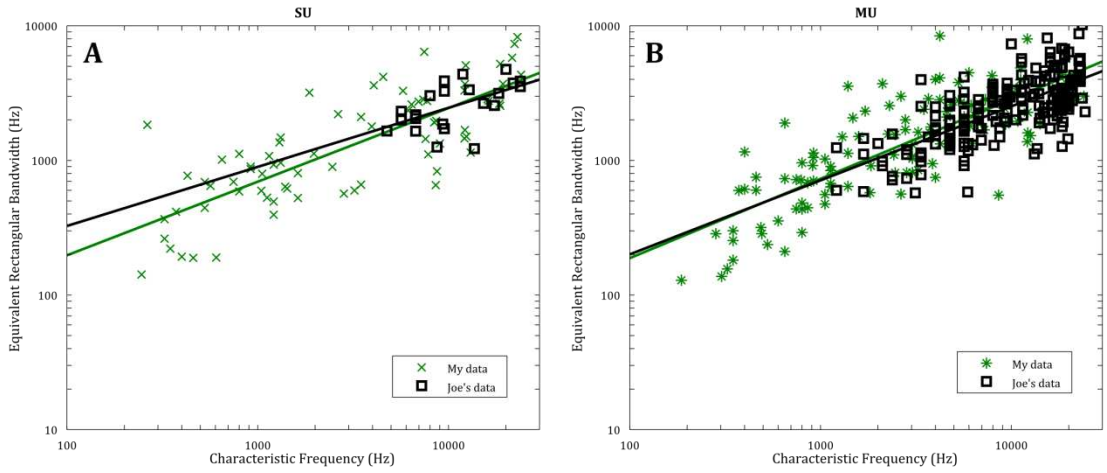


Figure 2.22

*FRA estimated ERBs from Figure 2.19 (**A**) Figure 2.20 (**B**) panel **E**, compared with ERBs made from FRAs recorded by other experimenters in the lab. The method for calculating the ERB is consistent across all points, however the stimuli used to derive the FRA differs between experimenters.*

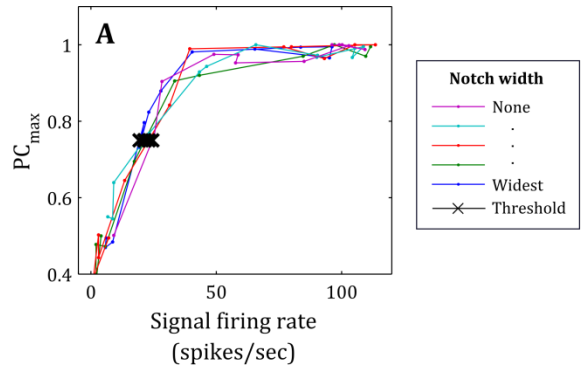
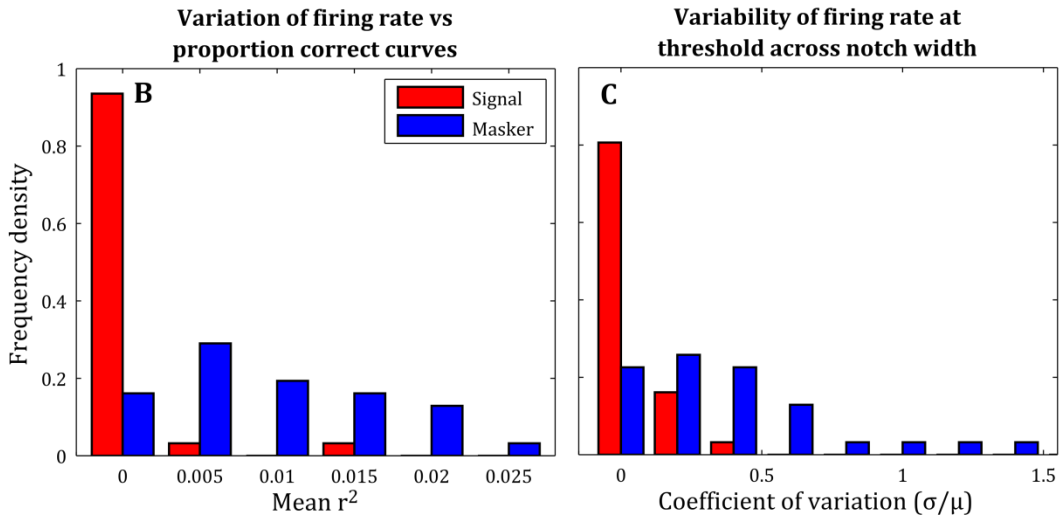
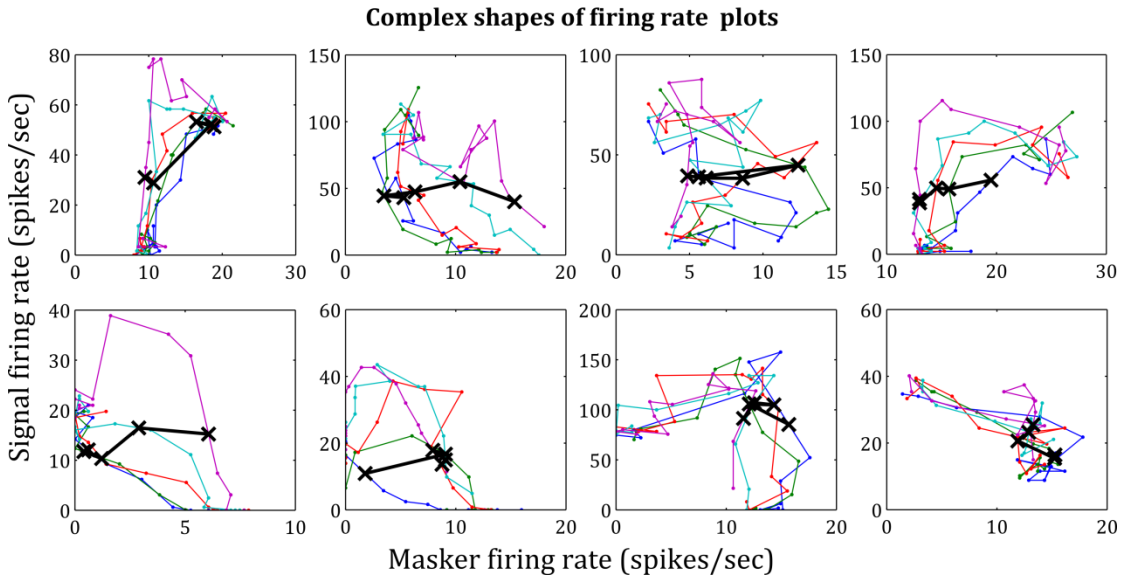


Figure 2.23

A *Example of the firing rate in response to the signal against proportion correct for an SU. × denote threshold.* ***B*** *Average difference between LOWESS fit and data points for the response to the signal (red) and masker (blue), for all SU.* ***C*** *Coefficient of variance for firing rates at threshold across notch width, for all SU. Lower panels: examples of complex nature of firing rate plots.*





2.5 Discussion - IC

Bandwidth estimates were obtained in three different ways from a number of units in the IC of the guinea pig. These units had CFs and thresholds that fairly evenly spanned the audible region of the guinea pig (as quantified by behavioural audiograms). Roughly 20 % of SU and MU showed only an onset response to a long narrowband noise signal, which is consistent with other studies in the IC (Astl et al., 1996; Le Beau et al., 1996). Together these suggest that the units in this study give a balanced view of IC activity.

One of the greatest challenges when trying to interpret the results of this chapter, is in working out how they fit in with other studies of frequency selectivity. The biggest problem behind this, and the driving force behind this study, is the lack of consistency in the method used to obtain the results. Because the idea of AF is fairly abstract, and there is not even a consensus on something as simple as their shape, there is not one established and "correct" way of obtaining an AF bandwidth. Add to this the number of steps required to convert neural activity into a bandwidth estimate, and the assumptions made along the way (what feature of neural activity to use, how to define detectability of a sound, how to define the AF from this), plus deciding the nature and specific parameters of the stimuli to elicit the neural activity in the first place, and there is enormous scope for diversity. Any subtle variation in any of these points makes comparison of results difficult, since one cannot be sure if differences or similarities between the two are caused by method. On top of this differences in species, surgical technique, anaesthesia, recording equipment, and data extraction technique, have an effect on results. With that in mind, in this discussion I have focused mainly on a few studies that are the most comparable to the one in this chapter, and point out the caveats in making comparisons between them.

Let us start, however, with the most general and global feature of my results. For all data types and bandwidth calculation methods in this chapter, the clearest feature is a strong linearly increasing relationship between the CF and ERB of a unit, in log-log space. This means that as CF increases so does AF bandwidth. This is a trend seen in all psychophysical and almost all physiological studies of frequency selectivity in humans (Fletcher, 1940; Zwicker, 1974; Houtgast, 1977; Glasberg and Moore, 2000; Oxenham and Shera, 2003) and animals alike (Pickles, 1975; Ehret and Merzenich, 1985; Ehret and Merzenich, 1988; Syka et al., 2000; Evans, 2001), regardless of any differences described in the previous paragraph. The metric used to describe bandwidth and the parameters of the line describing the relationship between frequency and bandwidth

(i.e. slope and intercept of the straight line in log-log frequency space, or the scale and power of the line in linear frequency space) are the important features that vary between studies. It is this relationship that tells us most about frequency selectivity and is the feature of my results that I will be comparing most to other studies. All such plots in this thesis are in log-log space so I may refer to the 'rate of increase in bandwidth' as the slope of these CF vs ERB lines, and the 'absolute bandwidth' as the intercept (or vertical position on the plot).

Three studies looking at bandwidth in the urethane anaesthetised guinea pig by calculating the ERB are probably the most comparable studies to my own (in terms of species and experimental preparation). These are Evans (2001), looking at individual auditory nerve fibres, Sayles and Winter (2010), SU in the ventro-cochlear nucleus, and Palmer et al. (2013), SU in the inferior colliculus. (Palmer et al. (2013) also used other types of anaesthetic but the difference in bandwidth between them was negligible so all anaesthetic types were combined). When looking at the bandwidth estimates from notched noise data in this chapter, not only is there a general linear increase in CF with ERB (positive slope), but the relationship between the two is in very close agreement with results from these three studies, in particular the results of Evans and Sayles. When the mean CF vs ERB lines are superimposed on the results for my MU and LFP data, they are indistinguishable (see Figure 2.10 and Figure 2.11). For the SU data there is an exact match for units above ~3 kHz with a slight tendency for my NN estimates to be narrower for frequencies below, but this difference is negligible (see Figure 2.9). Palmer et al. (2013) results are also similar, with a slight tendency towards broader filters for higher frequency (> 2 kHz) units, for all data types.

Even more remarkably, all three studies just mentioned used the FRA (firing rate in response to pure tone stimuli) of a unit to calculate the ERB, which is very different to the SDT with notched noise masking approach used in my study. The mechanisms driving masking are still not fully known but they involve a complex relationship between cochlear mechanics and neural connectivity (see 1.8.9 Mechanisms behind masking for more details). In addition the NN masker contains bands of frequencies activating a wider region of the BM than a simple pure tone. Add to this the fact that the IC is further along the auditory pathway and receives input from almost every known auditory brain region (brain stem (Huffman and Henson Jr, 1990), CN (Oliver, 1984, 1987), superior olivary complex (Glendenning et al., 1992), lateral lemniscus (Saint Marie et al., 1997), auditory thalamus (Anderse et al., 1980) and all auditory cortical regions (Winer et al., 1998)), and it becomes astonishing that there is such close

agreement between such vastly different bandwidth calculating methods and brain regions.

As mentioned previously, FRAs were collected for each of the units that gave an NN estimate (and many others that did not), and ERBs were calculated from these FRA in order to see the relationship between notched noise and FRA derived measurements (see Figure 2.15). By using the exact same units, all manner of confounding differences are avoided and any relationship between the two is much more certain. Surprisingly when the NN estimates were compared to the FRA ones made in the same units, there was a large difference; although the rate of bandwidth increase was comparable across the board (for all data types), FRA bandwidth estimates were roughly 3 times broader. This is highlighted in Table 2.2 where the slope of increase (b) varies between 0.5 and 0.7 for all data types and bandwidth method, but the intercept (a) is markedly larger for the FRA method.

What may be causing this systematic shift? Although the parameters used to define threshold for the FRA have been shown to have a significant effect on estimates of bandwidth, it is not enough to explain the difference. This can be seen in Figure 2.16, with the lines of best fit for the NN data sitting outside the 95 % confidence region around the average line of best fit for the FRA. Bandwidth measurements made using data collected by other researchers (who used the same experimental protocol) using the identical analysis method (as described in 2.2.7 Analysis*:* FRA analysis), show the same CF vs ERB relationship. This rules out any potential differences caused by guinea pig strain, stimulus parameters and experimental protocol, e.g. surgical preparation, anaesthetic, data collection method (see Figure 2.17). In addition, the FRA bandwidth measurements are not unusually broad for similar estimates from the IC literature. When the identical method is used as in Palmer et al. (2013), the bandwidth estimates match up very closely, any differences easily being covered by potential differences in threshold parameters. So it would seem that my FRA measured bandwidths are not unusual when compared to other similar data sets, yet show a clear difference to notched noise estimated bandwidths in the same unit. This suggests that bandwidth as defined by responses to pure tones are not the same as bandwidth calculated using NN masking, at least for IC neurons. This has been seen before in other similar studies using pure tones and a traditionally psychophysical method on physiological data. In Pickles (1979), Ehret and Merzenich (1988), Ehret and Schreiner (1997), and Egorova and Ehret (2008) comparisons between FRA and band-widening estimates of bandwidth for neurons in the cat AN, IC and A1 showed that they had very little in common with one

another, although there was no systematic difference between them. This may, however, be due to well-known problems with band-widening (see 1.3.3 Band-widening and the critical ratio).

So how do the results of these different physiological bandwidth measuring methods compare to psychophysical ones, and therefore the performance of the whole system? The only psychophysical measure of frequency selectivity in the guinea pig that I know of was made in Evans et al. (1992). Evans trained 10 guinea pigs to perform a lever pressing task with notched noise and rippled noise stimuli to determine psychophysical measures of bandwidth. Although differences have been found between rippled noise and notched noise methods in humans (Glasberg et al., 1984a), no such difference has been found in animals (Pickles, 1975) and so it is probably safe to combine results from both. The NN results in this chapter match up very closely to these behavioural results; more closely, in fact, than those from the FRA. This notably close relationship would suggest that whatever quality the NN method is capturing, is consistent between behavioural and neural activity in the IC, but is not fully captured when looking at pure tone responses only.

The parity between psychophysical and physiology using the same method contrasts the findings of Ehret and Merzenich (1985). In this study Ehret and Merzenich used the band-widening method (see 1.3.3 Band-widening and the critical ratio) to obtain critical band measurements from IC neurons and from the behaviour of the cat. They found that the rate of increase in bandwidth for both data types was roughly equal (slope of 0.702 for neural and 0.768 for behavioural), but that the neural bandwidths were about 3 times broader than the behavioural ones. Many problems have been found with the band-widening method, including off frequency listening, in which the recruitment of neighbouring AFs that have a better SNR confounds bandwidth estimates (see 1.3.3 Band-widening and the critical ratio for more details). The impact of effectively sampling from a range of AF is to give a much broader estimate of bandwidth than actually exists. Off frequency listening does not exist for single cells, so this may be the cause of the disparity between results from the two data sets. A notched noise approach for both psychophysical and physiological data would not suffer from this problem.

Another implication of the similarity between the NN IC results and those from the studies of Evans and Sayles, is that frequency selectivity may already be established at the periphery, and is maintained (at least) to the level of the IC. This theory is not new and studies comparing behavioural and neural estimates of frequency selectivity in

animals have come to this conclusion after observing the similarity between the bandwidth estimates of each method. The Evans study, which has been mentioned several times already, showed that ANF bandwidths were almost identical to behavioural ones. Bilsen et al. (1975) used rippled noise to calculate the bandwidth of SU in the CN of the cat, and compared these to psychophysical results (also using rippled noise in the cat) from Houtgast (1974) and Wilson and Evans (1971). The similarity between the two led them to hint at a causal relationship. Both of these studies are concerned with more peripheral auditory brain regions than the IC, however there is also evidence that this frequency selectivity is maintained up to the IC. Salvi et al. (1982) calculated psychophysical and evoked potential tuning curves (implanted chronic electrodes in the vicinity of the IC) in the chinchilla, using tone-on-tone masking for both. The close agreement between the two led them to conclude that "tuning takes place primarily at the periphery". My results fit in nicely with the conclusions of these studies. Furthermore, if masking experiments reflect cochlear and psychophysical tuning (at least up to the level of the IC) then if pure tone-measured bandwidths diverge from this, they must no longer reflect cochlear tuning. Perhaps excitatory inputs from the cochlea are pooled as one moves up the auditory system, giving progressively broader bandwidth estimates when using pure tone stimuli. Peripheral tuning is maintained only through mixed signal/masking mechanisms, and therefore it becomes important to use NN, or other masking methods, when measuring physiological bandwidths.

Another trend that is evident in the results of this study is that simultaneously masked estimates of bandwidth are broader than forward ones (at least for SU and MU). This is a well-established phenomenon in psychophysics, seen widely in human studies of all type (Houtgast, 1973, 1977; Wightman et al., 1977; Moore, 1978; Vogten, 1978; Moore and Glasberg, 1981; O'Loughlin and Moore, 1981; Glasberg et al., 1984a; Moore et al., 1984; Weber and Patterson, 1984; Moore et al., 1987; Oxenham and Shera, 2003), and has been observed in animal studies, both psychophysical (Kuhn and Saunders, 1980; Serafin et al., 1982) and physiological (Harris et al., 1976). In humans the difference between the two temporal arrangements is substantial (Houtgast, 1977; Moore and Glasberg, 1981; Oxenham and Shera, 2003), with forward masked bandwidths being reported to be as much as half that of simultaneous masking. However, the difference between the two in this study is slight and only weakly significant. One reason for this could be that the effect of suppression in the guinea pig is thought to be much less pronounced than in humans (Prijs, 1989). Since the difference between forward and

simultaneously masked bandwidth measures is thought to be mainly due to suppression (Sachs and Kiang, 1968; Houtgast, 1974; Moore, 1978; Heinz et al., 2002), this result is unsurprising. So it seems that this difference, seen most prolifically in psychophysical studies, can be identified at the neural level in the IC. This adds weight to the argument that it is caused by a non-linearity on the BM, namely suppression.

An investigation into the mechanisms of masking in the SU shows variations in the masking characteristic between units. A small proportion of units showed 'linear' type behaviour when comparing their response to the masker and the signal. These were characterised by a simple relationship between the two: as masking increases, masker firing rate decreases, and signal firing rate increases at the same rate (see 2.3.6 Mechanisms underlying masking in IC). Such units fit nicely with the assumptions of the power spectrum model described in 1.3.2 Power spectrum model. The majority of units, however, showed varying degrees of non-linear masking behaviour. One feature that appeared in roughly a third of SU was a spread in firing rate in response to masker vs signal curves with notch width. For these units the summation of energy at the output of the AF does not explain everything and there is an additional frequency dependence on firing rates; although RLFs of the response to the signal (which fully drives detection thresholds) displays frequency selectivity, this is not the case for the response to the masker, whose RLFs are invariant across notch width. This implies that these 'separated' units (see Figure 2.14) do not display the expected frequency selectivity in the excitatory response to the notched noise maskers. Again this could suggest that excitatory inputs are pooled together across many units further down the auditory system, losing the fine frequency selectivity seen through masking. For half the units the firing rate in response to the masker vs signal plots showed much more complex, non-linear behaviour, suggesting very complicated masking interactions. However, for all of the units investigated, the RLF of the firing rate in response to the signal are clear and systematically separated across notch width, seemingly showing no link to this complex behaviour. This is reflected in the fact that no pattern is found across CF, ERB or their confidences, across these different unit classifications; the only feature that matters for bandwidth measurements is the response to the signal at threshold. Perhaps the assumptions of the power spectrum model are not violated by such seemingly non-linear behaviour.

One final noteworthy feature of the data is how similar the results of the different data types are (SU, MU and LFP). MU is an amalgamation of electrophysiological activity in a small region, and LFP is the same but over a much larger area (Destexhe et al., 1999;

Katzner et al., 2009). Although LFPs contain post-synaptic potentials and other non-spiking activity (Manning et al., 2009; Kajikawa and Schroeder, 2011), they contain a high proportion of combined spiking activity and can therefore be treated as a 'proxy' for the assessment of neuronal outputs (Buzsáki et al., 2012). The only trend across data types seen in my data is that the confidence of each bandwidth estimate goes down, and the variability between them goes up, as one moves from SU to MU to LFP. This suggests that the activity of neurons in the IC is linear and homogenous. MU and LFP are simply the summation of the activity of a larger number of neurons than SU. The tonotopic representation of frequency in the IC means that as the area of neural summation is increased, the extra neurons have CFs progressively further away. The activity would still be dominated by the characteristics of the closest neurons, so the mean activity would still be the same as with SU. However, the variance would increase since a wider range of responses is being summed. This fits in with our understanding that the IC is concerned with representing low level acoustic features that are used by brain regions higher up the chain (Nelken et al., 2005).

2.6 Discussion - A1

Bandwidth estimates were obtained from single- and multi-unit, and LFP activity in primary auditory cortical neurons of the anaesthetised guinea pig. Three different methods were used to make these measurements: forward and simultaneous notched noise masking, and the firing response of a unit to pure tone stimuli (FRA). The idea behind this study is to compare the different methods in the same populations of units, and see what the resulting measurements teach us about frequency selectivity in A1.

Both SU and MU had CFs and thresholds spread evenly across the range of hearing of the guinea pig, as defined by the guinea pig audiogram (see Figure 2.18). In addition the majority responded only to the onset of stimuli (87 % in SU, 63 % in MU) and did not have a sustained response. This is a well-established property of auditory cortical neurons (Phillips, 1985; deCharms and Merzenich, 1996; Eggermont, 1997; DeWeese et al., 2003). Together these suggest that the units in this study give an unbiased view of A1 activity.

When comparing the results of this study to others in the field, it important to be aware that there are far fewer studies in A1 than other auditory brain regions, and a scarce number in the guinea pig. The most popular animal is the cat and so many of the comparisons made here will have the caveat of this species difference. Also cortex is far less homogeneous than lower brain regions and so results are not as consistent across the literature. Nevertheless, here I have tried to compare my data to the most relevant studies and to include those with conflicting findings.

Firstly, the most prominent feature of my results is a linear increase in bandwidth with frequency, across all data types and method. This is seen in my IC data, all psychophysical studies in humans and animals, and every physiological study of frequency selectivity mentioned so far. In Phillips and Irvine (1981) the bandwidth of cat A1 neurons was investigated by calculating Q_{10} from FRAs. They found a general increase with CF similar to that seen in other species and brain areas (Hernandez et al., 2005; Taberner and Liberman, 2005). However, Schreiner and Mendelson (1990) is a similar study in cat A1 and no such trend was found. Q_{10} did not seem to increase with BF and bandwidth varied non-uniformly across A1. They concluded that there are functional maps in A1 that run orthogonal to the cochleotopic organisation of frequency. This was backed up by a further study Schreiner (1995). Recanzone et al. (2000) also found no increase in bandwidth with BF, although this study was conducted in awake behaving macaque. They also found no systematic organisation of bandwidth in A1. So

there are mixed results for even this fundamental feature of frequency selectivity seen elsewhere.

Another clear trend present in my data is that of sharper high frequency tuning, when estimated using NN, than seen in the IC or in any of the other auditory brain regions described in the previous sections. For all the three data types collected, units with CF higher than ~2 kHz have NN bandwidths about 2-3 times narrower than the IC and periphery. Also the bandwidth estimates of this population of neurons have very small confidence regions meaning we can be fairly certain that this result is not just a fluke of statistics. In addition the bandwidth estimates for low frequency (< 2 kHz) MU and LFP are progressively narrower than in SU.

A concern when looking at notched noise experiments in cortex is that many of the assumptions behind the power spectrum model may be violated here. Mainly the assumption of a linear summation of sound energy. Neurons in auditory cortex have been shown to display highly non-linear characteristics. In Sutter (2000), the FRAs of A1 neurons of the cat were investigated in an attempt to categorise their shape, as has been done in many other brain regions. They found a vast range of shapes and non-monotonic behaviour that was not as prevalent in neurons in lower brain regions. As a result they thought of FRA shape in the A1 as more of a continuum that cannot be categorised clearly and easily. In addition they concluded that FRA shape is established by 'complex patterns of inhibition and excitation accumulating along the auditory pathways, implying that central rather than peripheral filtering properties are responsible for certain psychophysical phenomena'. In Ehret and Schreiner (1997), a similar study in A1 cat neurons was conducted, but in addition bandwidth using the band-widening technique was estimated (see 1.3.3 Band-widening and the critical ratio). They found that FRA shape is so different from the AF calculated using band-widening, that the two aren't measuring the same thing. They concluded that 'spectral integration in the way proposed to be the basis for the perception of tones in noise is not present at the level of the A1'. In Bar-Yosef et al. (2002) the response of cat A1 neurons to modified bird song was investigated to see the effect of subtle manipulations to natural sounds. They found that the slightest change in stimulus property had a vast effect on the response of the neuron and that therefore A1 units respond to a very specific combination of stimulus characteristics. They concluded that this 'pose[s] a challenge to models based on linear summation of energy' such as the power spectrum model used in NN experiments. So the activity of neurons in A1 might be so non-linear that obtaining a realistic bandwidth estimate is impossible. The relationship between the response of my A1 units to masker

and signal (see 2.4.6 Mechanisms underlying masking in *A1*) adds some credence to this point; all units showed very non-linear behaviour when compared to the more homogeneous results in IC. Does this cast doubt on any bandwidth estimates made in A1 neurons?

Evidently further investigation of whether this non-linear behaviour is affecting my results is necessary. There are relatively few NN studies looking at frequency selectivity of A1 neurons, let alone in the guinea pig, so in many cases it is best to compare my A1 results with my IC data. They use the same method and experimental preparation so the only difference should be between neural activity in the respective brain areas. With that in mind I compared the neurometric functions of the units in the A1 study with those if the IC. One of the consequences of linearity required by the power spectrum model is that the only difference between the neurometric functions of successive notch widths is a shift in threshold; all other features should remain unchanged. For example, if the slopes or maxima of the functions show a systematic shift across notch width in a unit, or are significantly different for different populations of units, then we might conclude that the assumption of linearity is violated. Figure 2.24 shows a comparison of the mean neurometric shape (and ± 1 standard deviation) across all notch widths of SU in the IC and A1 data. The plots show that there is no difference between the two brain areas, for both forward and simultaneously masked NN experiments. Moreover the effect of notch width on neurometric shape was investigated by calculating the mean neurometric shape for each size of notch width across all SU; there was no difference here either. Finally, the A1 units were split into 2 groups: those with CF above, and those with CF below 3 kHz. There was no difference in neurometric function shape between these populations as a whole, or when segregated by notch width. This means that the A1 neurometric functions are no different to those of the IC neurons and so cannot explain the sharper tuning we see for neurons with higher CF.

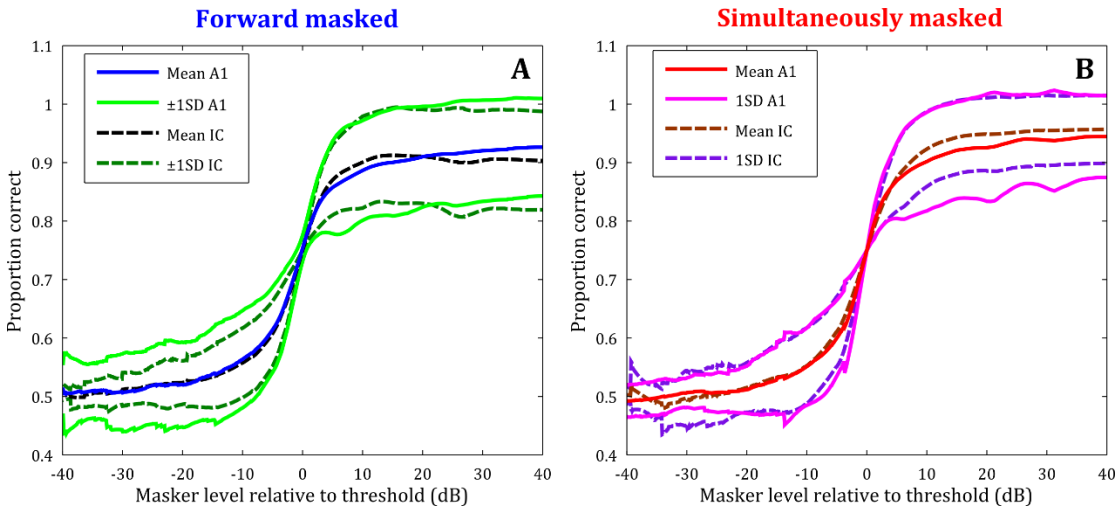


Figure 2.24

*Mean neurometric function, with ± 1 standard deviation, for all notch widths for A1 and IC units, for forward **A** and simultaneously **B** masked NN experiments.*

There is also evidence in other studies of sharper tuning in A1, which would agree with my findings. Bartlett et al. (2011) found finer tuning in SU of both A1 and the thalamus of the awake marmoset than typically seen in AN, and much finer than other anaesthetised animal studies (Miller et al., 2002; Moshitch et al., 2006). They concluded that anaesthesia probably broadens bandwidth estimates. However, the fine tuning seen in my study may refute this as the cause. Also Bitterman et al. (2008) found very fine tuning in human auditory cortex, recorded using intracranial depth electrodes in response to natural stimuli. Their bandwidth estimates were far narrower than typically described in other mammals and substantially exceeded the tuning attributed to human auditory periphery. Although there are vast differences between species, anaesthesia and method, the qualitative findings of these studies are similar to my notched noise results.

These findings were not seen, however, in Fishman and Steinschneider (2006), one of the more similar studies to my own. Fishman used a notched noise type experiment, using roex function fits, to estimate bandwidth of MU and current source density recordings (similar in nature to LFP) in A1. This study was in awake macaques, using noise bands of only 50 Hz and presenting the masker and signal simultaneously, so does have differences to my own, but is more similar than many other studies described above. They found that psychophysical bandwidths matched up very closely to behavioural estimates in the same animals, across all data types. Also they matched up nicely with results of another psychophysical bandwidth study in the macaque

(Gourevitch, 1970), although this used band-widening. The only psychophysical bandwidth estimates in the guinea pig that I could compare my results to, are from Evans et al. (1992), where notched and rippled noise was used. Results from this paper do not match up exactly with my A1 bandwidths, especially at high frequencies; a finding in opposition to Fishman's. What would especially fine tuning mean for A1 units? If SUs show markedly finer tuning than the ability of the system as a whole, then this would imply some sort of sub-optimal coding. One, perhaps, where the average activity of a population of neurons is taken, only a small proportion of which are finely tuned for the stimulus. Alternatively more complex interactions between units across various auditory cortical regions may be inefficient, making the overall perception slightly poorer. However, the evolutionary basis for this may be slightly dubious.

So far I have been mainly talking about my NN results, but what of the bandwidth estimates calculated using the FRA? Once again these were much broader than the NN results in A1 and broader than the NN results in the IC and peripheral guinea pig auditory brain regions. In fact, these bandwidth estimates are very similar to the FRA bandwidths in the IC. If superimposed the two are almost indistinguishable. Notably the difference between FRA and NN bandwidths is no longer a simple case of scale, as it is in the IC, but the rate at which bandwidth increases with CF is actually different as well. This can be seen in Table 2.4 where the slopes of the lines of best fit are substantially different. (This is most prominent for SU). The difference in A1, therefore, is a qualitative difference as well as quantitative.

As mentioned previously in this chapter, the edge of the FRA is used to represent the AF, and the firing characteristics that determine the edge of the FRA are controlled by a set of parameters. The values of these parameters are not predefined and are established experimentally. Is there a chance that this difference between the two methods is due to the choice of these parameters? To investigate this, I refit the FRA 80 times using different combinations of values for these parameters. I then calculated a line of best fit for each resulting bandwidth vs CF plot. (A similar procedure was undertaken with the IC data in 2.3.7 FRA based bandwidth). The mean (± 1 standard deviation) of these lines of best fit can be seen in Figure 2.25. Looking at this plot it is clear that the difference between methods is not caused by the choice of threshold parameters and there truly is a fundamental difference between the two measures. Also, again I calculated bandwidth using the FRAs of A1 units collected by another researcher in the lab (Sollini, unpublished) and compared them to my own; the results were identical. This difference in method is therefore not a fluke of the specific cells or animals used in this study. As

mentioned above, this is similar to the findings of Ehret and Schreiner (1997), where it was deemed that excitatory pure tone bandwidths, and those measured through masking, are not comparable. This is also seen in my IC data, but not in studies looking at more peripheral brain regions, namely VCN and AN (Evans, 2001; Sayles and Winter, 2010). Perhaps then cochlear tuning is maintained throughout the auditory system, right up to cortex, only within inhibitory or masking activity. As one moves up the auditory system, purely excitatory inputs are summed across ever growing populations of neurons, expressing activity over a growing region of the cochlea. This would lead to a large bandwidth estimate when using a purely excitatory input, and a much narrower one when using a masking paradigm. Or perhaps there is some other more complex reason behind the difference. Either way, it is clear that great care must be taken when using different methods for measuring frequency selectivity.

Forward and simultaneously masked NN estimates were also compared with one another in the same units, and were found to be non-significantly different for SU and LFP, but slightly significantly different for MU. This, however, is probably a result of a very low number of units. Because of the difficulties of recording in A1 only 12 SU yielded both forward and simultaneous estimates. This number was 13 for LFP. MU, however, had 20 which probably resulted in the just significant difference.

Another feature of note is the similarity between SU, MU and LFP results. For the NN results the low frequency tuning is slightly different across the data types (sharpening for MU and LFP), but for high frequencies they are very consistent. This has not been found in many studies, which instead report bandwidths that are systematically broader than for SU and MU (Eggermont, 1996; Noreña and Eggermont, 2002; Noreña et al., 2008; Eggermont et al., 2011; Kajikawa and Schroeder, 2011; Gaucher et al., 2012). One of their explanations is that LFP reflects mainly thalamic input and that tuning is sharper in cortex than the thalamus. This may be true for the FRA measured bandwidths, which seem to be slightly broader for LFP than SU and MU. The difference between the two, of ~50 %, in A1 matches with the findings of Noreña and Eggermont (2002), in which bandwidth was measured using FRA. No similar studies using NN exist, so perhaps this is just a phenomenon of pure tone excitatory tuning curves.

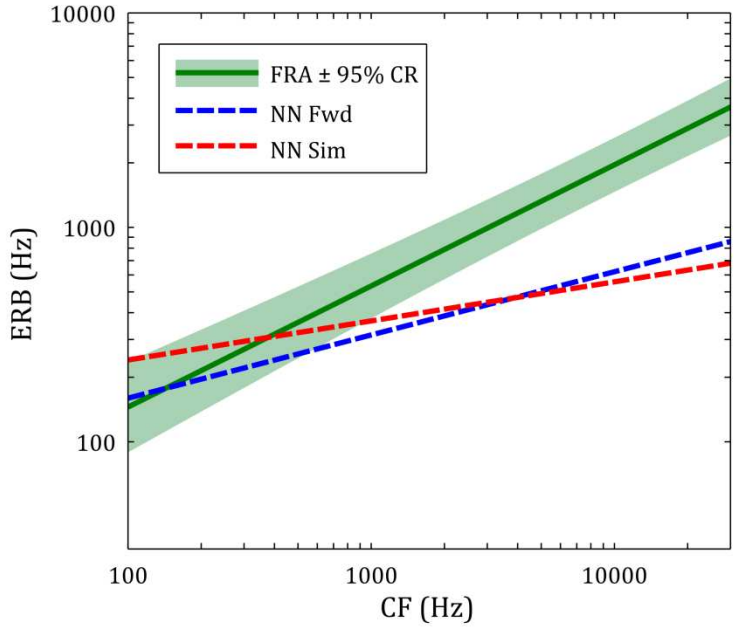


Figure 2.25

Mean line of best fit for FRA calculated bandwidth of SU using a range of different fitting parameters (green line), with 95 % confidence region (light green area). Lines of best fit for NN estimated SU for comparison; forward (blue) and simultaneously (red) masked arrangements.

CHAPTER 3: Ferret psychophysics

3.1 Introduction

Frequency selectivity describes the ability of a person or animal to resolve the individual frequency components in a complex sound, a fundamental feature of hearing vital for communication and survival. From the earliest days of its study, the auditory system has been modelled as a series of band-pass filters, overlapping one another and running along the full frequency range of hearing (Fletcher, 1940). In this model frequency selectivity is described by the width of these auditory filters. The goal of psychophysical research into frequency selectivity has been to calculate the size and shape of these filters and thereby characterise the properties of the auditory system. The main tool for doing so has been masking; establishing the detection threshold of one sound in the presence of another masking sound. Many different approaches have been used such as band-widening (Zwicker et al., 1957), tone masking (Chistovich, 1957; Houtgast, 1973) and rippled noise (Houtgast, 1974; Pick et al., 1977). Band-widening has been found to be very inaccurate for measuring bandwidth (Bos and de Boer, 1966; Patterson and Henning, 1977), tone masking is susceptible to interactions between signal and masker producing audible 'beats' which confound results (Plomp and Steeneken, 1968; Rodenburg et al., 1974; Terhardt, 1974; Wightman et al., 1977; Patterson, 1986), and each of these suffer from problems caused by off-frequency listening (Johnson-Davies and Patterson, 1979; Patterson and Nimmo Smith, 1980; O'Loughlin and Moore, 1981). As a result notched noise masking has emerged as the favoured method (Patterson, 1976; Patterson and Nimmo Smith, 1980; Moore et al., 1984; Moore et al., 1987; Glasberg and Moore, 1990) since it suffers from none of these problems. It has been developed over the years in human psychophysics and a number of features adapted and adopted, such as forward masking. This has superseded simultaneous masking since it does not suffer from problems caused by the non-linearity of the cochlea, such as

suppression (Sachs and Kiang, 1968; Heinz et al., 2002), which leads to broader bandwidth estimates made with simultaneous masking than forward (Houtgast, 1974; Leshowitz and Lindstrom, 1977; Moore and Glasberg, 1981; Moore, 1987; Oxenham and Shera, 2003). Also keeping the level of the signal fixed and varying the level of the masker when determining detection threshold has emerged as the favoured method (Rosen et al., 1992; Rosen and Baker, 1994; Baker et al., 1998; Rosen et al., 1998), although this is still somewhat controversial (Eustaquio-Martin and Lopez-Poveda, 2011; Lopez-Poveda and Eustaquio-Martin, 2013).

These developments have mainly been made in human studies, although animal behavioural studies in a wide range of animals have also contributed. Most of these studies have used band-widening (Watson, 1963; Miller, 1964; Gourevitch, 1970; Pickles, 1975) and tone masking (Saunders et al., 1978b; Dooling and Searcy, 1985; Burkey and Gans, 1991), but a few have used notched noise (Halpern and Dallos, 1986; Evans et al., 1992; Niemiec et al., 1992). Almost all these studies used simultaneous masking with only a handful using forward (McGee et al., 1976; Kuhn and Saunders, 1980; Serafin et al., 1982; Smith et al., 1987), and only one comparing forward and simultaneous notched noise masking (Halpern and Dallos, 1986). This study, however, did not successfully manage to measure the AF with forward masking. Also not one animal behavioural study that I have found has kept the signal level fixed and varied the masker level. It is therefore of interest to estimate AF widths in animals using methods from modern human psychophysics.

It is the aim of this experiment to develop a method for measuring both forward and simultaneously masked bandwidth measurements in the same animal, and use this to calculate bandwidth across frequency. This will be done by using a similar approach to that of Halpern and Dallos (1986), in which the masker is played in short repeating bursts with short gaps between, and short signal bursts played in these gaps to achieve forward masking. Unlike the aforementioned study, simultaneous masking will be achieved using the same stimuli but with the signal bursts appearing during the masker bursts. By using the same method it is hoped that the same animal can easily switch between the two tasks without needing retraining, and a direct comparison of the two masking types can be made. The ferret will be used, since it is a good behavioural mammal (Kelly et al., 1986), with a one-interval two-alternative forced choice (1I2AFC) lateralisation task to obtain individual trial data. This method gives lower and more reliable thresholds (Sollini, 2013) than the pure detection task used in Kelly et al. (1986), and would result in less bias in simultaneous cortical recordings. If this method

can be made to work, then animal behavioural techniques can be brought more in line with human studies and we could learn more about the effect that methodological differences have on bandwidth estimates.

3.2 Methods

A number of ferrets were trained to perform a behavioural task to be used in conjunction with notched noise masking stimuli to measure their frequency selectivity. A novel method was introduced to allow stimulus paradigms that are comparable with recent human psychophysical methods. This Methods section does not only describe the experiment and how the data was collected and analysed, but also the development and validation of the novel approach.

3.2.1 Animals

In total 4 male and 4 female pigmented ferrets (*Mustela Putorius*), all of whom were bred in house, were trained to perform the behavioural task. Table 3.1 below shows the ages of the ferrets at the start and end of their training. All ferrets were housed independently bar 2 females (Lola and Lacey) who were housed together. 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. Certain physiological processes of the ferret, such as breeding cycle, weight gain/loss, hair gain/loss, are seasonal, triggered by hours of daylight and temperature. To keep the ferrets in good health the animal environment was changed according to the season, with winter in November-February, spring in March, summer in April-September, and autumn in October. Temperature was set at 17 and 19°C ± 2˚C for winter and summer respectively, and at 18°C ± 2˚C for spring and autumn. Also light was varied seasonally with 8 hrs. light/16 hrs. dark in winter, 12 hrs. light/12 hrs. dark in spring and autumn, and 16 hrs. light/8 hrs. in summer. For the time that the ferrets were performing the behavioural task, they were kept on water regulation. This involved removing water from their home environment meaning all the water they received was during the behavioural sessions. This fluid intake was supplemented by 'wet mash' which consisted of water mixed with ground protein pellets (their normal diet) and a food supplement (Cimicat, Petlife International Ltd., UK). While on water regulation ferrets performed two behavioural sessions a day, for up to 14 (but usually 10-11) consecutive days. After such a period animals were given water ad *libitum* for a minimum of 3 days. After 3 consecutive regulation periods (with intervening rests) the animals were given a longer rest. These respites were intended to allow the animals to regain any weight lost during water regulation. Since such weight loss is common, the weight of each animal was monitored daily and not allowed to fall 20 % below normal.

Ferret	Age (in months) at...	
	Start of training	**End of training**
Hendrix	38	43
Lily	5	10
Leeroy	6	10
Lex	12	13
Lola	18	23
Lacey	18	23
Manuel	10	11
Naomi	21	23

Table 3.1

Ferret ages at start and end of training.

3.2.2 Ferret arena

Figure 3.1 contains a schematic of the ferret arena. All behavioural experiments were performed in a raised cage (60 cm between room floor and cage floor) with wire mesh top and sides, and hard PVC floor, within a sound attenuated room. The cage is T-shaped with a rectangular central platform (26 × 20 cm) extending from the centre of a longer platform (30 × 100 cm). The central platform has a water spout, surrounded by a small brass cone, three small fences and two posts containing an LED beam breaker. The small fences and cone are used to align the ferret's body and head, minimising the variability in head position between behavioural trials. When the ferret aligns itself on the platform the beam between the posts is broken; this ensures a trial only occurs when the ferret is aligned correctly. Two water dispensing modules flank the cage, one on each end of the longer platform, with water spouts protruding into the cage. Each water spout contains an infrared sensor which is used to detect ferret licks. Four speakers (ADAM, A3X), with frequency response range 60 Hz - 50 kHz, stand on raised platforms (centre of tweeter 75 cm, subwoofer 65 cm above floor), facing the central water spout at a distance of 88 cm. If the point directly in front of the central platform (also the direction the ferret faces at the start of each trial) is taken as 0° on a compass, then the 4 speakers are arranged with their centres at 75°, 90°, 270° and 285°. To ensure that all speakers and room furniture do not adversely affect the delivery of stimuli, a maximum length

sequence (MLS) measurement was taken regularly to determine the frequency response of the setup (Rife and Vanderkooy, 1989). An example of such a recording, with frequency responses from each speaker, can be seen in Figure 3.2. Such recordings were also used to ascertain the level of the stimuli in dB SPL by referencing to a 1 kHz pure tone calibrator (B&K; 4231).

All speakers, water spout infrared sensors, and central LED beam breaker are connected via an audio interface system (MOTU 24 I/O) to a computer. Here an in-house software program (sitafc, written in VBA) takes all the apparatus input and uses this to automatically control the experiment i.e. determine correct/incorrect behaviour, dispense water as appropriate, play the appropriate stimulus at the appropriate time, record all events, etc. The system is designed to be fully automated so that the ferret triggers trials, which are controlled by the software.

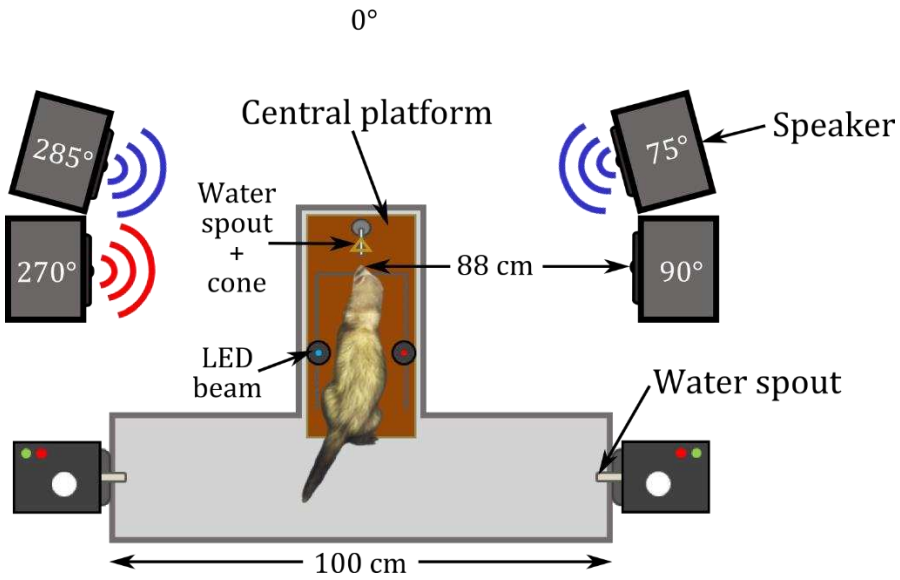


Figure 3.1

Behavioural arena. Four speakers placed at 75°, 90°, 270° and 285° relative to directly in front of the ferret's head at a distance of 88 cm. The central platform extends from a longer cage area, flanked by water spouts. The masker is played from the 75° and 285° speakers (blue lines) and signal played from 90° or 270° (red lines). Water spouts, speakers and the LED beam breaker are all connected to a computer which controls the experiment.

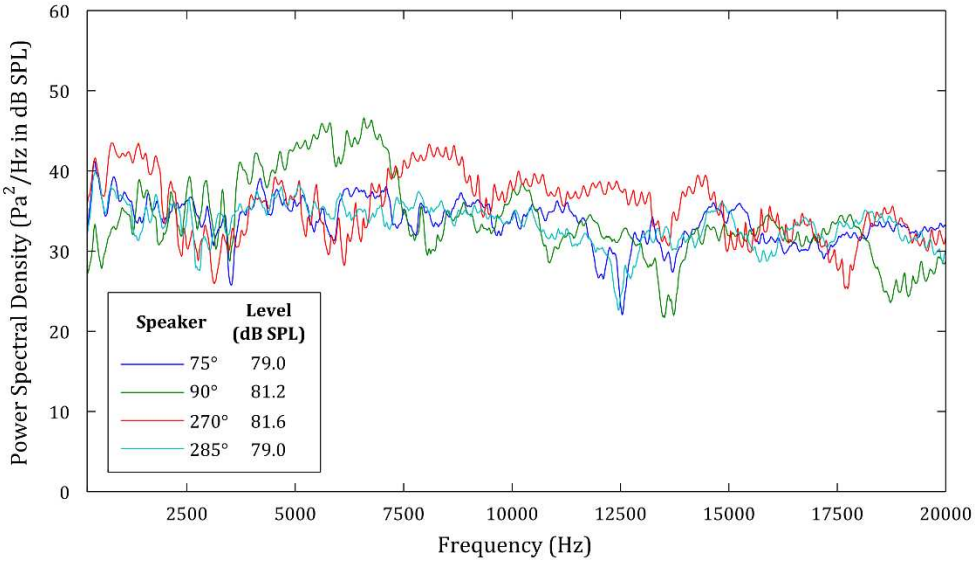


Figure 3.2

Frequency response functions in the behavioural arena from each speaker. MLS, with n *= 17 and sample rate 96 kHz, were played from each speaker and recorded by a ¼" free-field microphone (GRAS; Type 40BF) placed in the location of the ferret's head. The power spectral density was calculated in dB referenced to 20 µPa. Level was calculated by recording a 1 kHz tone played from each speaker and comparing the RMS with that of a 1 kHz calibrator (B&K; 4231).*

3.2.3 Behavioural task

In human psychophysics one of the most common tasks is a 2I2AFC. A masker is played twice in quick succession with the signal following one of the masker presentations, and the participant is asked to select which interval the signal appeared in. A psychometric function from this has range 50 % to 100 % correct, and can usually be fit by a logistic (or similar) function. Detection threshold is then taken at 75 % correct. This technique is very successful and is therefore used extensively in humans (Macmillan and Creelman, 2004). A variant is used in animal psychophysical studies (Walker et al., 2009; Bizley et al., 2013a; Bizley et al., 2013b), however its cognitive complexity makes it difficult to implement and maintain high performance. Another commonly used method is the go/no-go task where the animal is trained to respond to a certain stimulus arrangement and do nothing for another (McGee et al., 1976; Koay et al., 1998). Another method, and often the two are combined, is a shock-avoidance task where the animal performs a discrimination task and is given a small electric shock for incorrect behaviour (Green, 1975; Willott et al., 1984).

One problem with negative reinforcement, such as shock avoidance, is that it encourages the animal to err on the side of caution causing the under-sampling of false positives, introducing an inherent bias (Willott, 2001). Also it causes a lot more stress in the animal, compared to water/food regulation, which has been shown to weaken the conditioning (Bechtholt et al., 2004). A disadvantage of the go/no-go method arises when making physiological recordings during a behavioural task (Brosch et al., 2005); movement is coordinated by motor cortex, which may have related activity in auditory cortex (Brosch et al., 2005). This means that some of the neural activity recorded during a sound trial will actually be in response to motor activity (or the planning of motor activity). Additionally neural activity in primary auditory cortical regions have been shown to reflect decision activity about which way the animal should respond (Bizley et al., 2013b). In a go/no-go task one response requires movement and the other does not, so one will contain motor signals as well, and the other will not. This creates an inherent bias in any physiological recordings. Therefore, despite behavioural benefits of the go/no-go task (Burdick, 1979), a 1I2AFC type task requiring a response to each stimulus presentation would be preferable when physiological as well as behavioural data is being collected. There would still be unavoidable interference from motor cortex, and related decision activity, but at least it would affect each response equally.

In Kelly et al. (1986), Hine et al. (1994) and Alves-Pinto et al. (2012) such a task was used in ferrets. In a single trial a stimulus was either presented or not. The ferret was trained to approach one location to answer 'stimulus detected' and another for 'stimulus not detected'. Such an association is still quite abstract and has a high cognitive load. This means that confusion due to the task, and not psychophysical detection, is more likely, which would interfere with results. Also it takes a long time to train the animals and the association is more likely to be lost. As a result I used a slightly adapted version of this 1I2AFC task that incorporates lateralisation, which makes the task easier and more likely to succeed. Instead of randomly presenting or not presenting the stimulus and training the ferret to approach one of two separate locations to answer which occurred, the target stimulus appears randomly at one of the two locations and the animal must approach the source. This removes the extra cognitive load since ferrets are predatory animals and are evolutionarily shaped to localise and approach sounds, making the task far more natural (Harrison, 1992; Iversen, 1994). As a result this approach has been shown to produce lower thresholds in the ferret when compared to Kelly's (Sollini, 2013).

The ferrets were trained to perform such a task with the signal coming from either the 90° or 270° speaker, with the closest water spout as the associated answer point. (Although the two are separated in location, the ferrets had no problem associating the speaker with the water spout). This was done gradually in stages; first copious amounts of water were given upon licking any water spout to gain an association between the spouts and water, then gradually water reward was reduced until the ferret was only given water when it aligned itself properly on the central platform. Next a high level, long duration, broadband test stimulus was introduced, appearing randomly from either side speaker with equal probability. The stimulus was very easy to detect and localise and ferrets very quickly (usually on the first trial) went to investigate the sound and lick the water spout for a water reward, associating the water spout with the stimulus from the nearby speaker. This process is known as 'shaping' and took between 3-10 sessions depending on the ferret (failing completely with one ferret). From then on the water reward at the centre was reduced gradually to only a few drops, every 10 trials with almost all water reward coming from a correct answer. This is done because head movement and noise caused by licking at the centre spout can affect the attention and hearing of the ferret when the stimulus is being played, which could adversely affect the experiment. The small proportion of trials for which a centre reward is given are discarded. Next the signal was made progressively more narrowband and shorter in duration until it was appropriate for the task. The final step was to gradually introduce a very wide notch (ΔF of 0.5 × SF) masker from the two speakers at 75° and 285°, initially very low level; effectively providing a very low amount of masking (if any). The aim is to accustom the ferret to the masker without making the task too difficult. They needed to learn to ignore the masker, and all of them did. With this done the ferret was ready to perform the behavioural task.

As mentioned above, a trial was triggered by the ferret and the procedure went as follows:

1) The masker plays from the 75° and 285° speakers continuously. The notch width of the masker is fixed for a session, but the level is randomly selected from a list of predetermined values. Randomly selecting the level is known as the 'method of constant stimuli' and prevents errors of habituation and expectation which result from the subject from being able to predict stimulus level (Laming and Laming, 1992).
2) The ferret aligns itself on the central platform, breaking the LED beam, which allows signals from the central water spout to be relayed to the computer.

3) It licks the central water spout, which is picked up by the infrared sensor and relayed to the computer. If the central water spout is licked for a certain short duration and the LED beam is broken, a trial begins.
4) After a delay, random in length between 0.5 and 2 seconds, the signal is played for a short duration from either the 90° or the 270° speaker, randomly chosen with equal probability. The delay is there to stop the ferrets from gaining a detection advantage by anticipating the start of the signal, or leaving the spout early. Also the level of the signal is randomly chosen from a predetermined list.
5) The ferret then moves to a water spout on either side of the cage. If this is the side the signal came from it receives a water reward of ~0.3 ml and this trial is deemed a 'correct' trial. If it goes to the other side, it is deemed an 'incorrect' trial, no water is given, and the same trial is repeated. This is done to reduce bias caused by ferrets only answering on one side since doing so would result in no water reward ever. Results from repeat trials are not included in further analysis. If the ferret doesn't lick a side spout within 30 seconds, the trial times out and no response is expected or accepted.
6) The masker level is adjusted and the process repeats until either an hour has passed or the ferret has had enough (usually going to sleep).

The responses of individual trials are taken and proportion correct calculated across all trials for each stimulus condition to make psychometric functions. This process is described below.

3.2.4 Psychometric functions

In a given behavioural session the frequency properties of the signal and masker are fixed, i.e. single notch width, and the levels change randomly between trials. In every case one stimulus is fixed in level and the other varies, depending on the experiment. This way a psychometric function can be built comparing proportion correct across SNR. One way of calculating proportion correct using SDT is to use the following formula:

$$PC = \frac{1}{2}(H - F) + \frac{1}{2} \tag{3.1}$$

where H is the hit rate and F is the false alarm rate of one of the distributions in Figure 1.13, defined as:

$$H = \frac{\#trials\ correct\ on\ right}{\#\ trials\ total\ on\ right} \tag{3.2}$$

$$F = 1 - \frac{\#trials\ correct\ on\ left}{\#trials\ total\ on\ left}. \tag{3.3}$$

This method, and the example described in 1.8.11 Signal detection theory, is usually used in a detection task where one distribution represents 'signal present', and the other 'signal absent'. Also it assumes that there is no bias towards answering one way or the other. In the 1I2AFC task the ferrets perform, in each trial the signal appears on one side and doesn't on the other, which makes it slightly different to (easier than) a pure detection task. Also, slight differences in the amount of water given at each answer spout can affect the likelihood of a ferret answering at a certain side. Some ferrets simply develop a tendency to preferably answer at a certain spout for unknown reasons, and so bias is an issue. Calculating the proportion correct for the other side, i.e. reversing the sides in the formulae so that *equation* (3.2) uses trials from the left and *equation* (3.3) from the right, gives a different answer if there is bias. Which one therefore is correct? To combat this I used a slight variation of *equation* (3.1) known as $PC_{\max}$ defined by the following formula:

$$PC_{\max} = \Phi\left(\sqrt{2}\frac{\Phi^{-1}(H) - \Phi^{-1}(F)}{2}\right). \tag{3.4}$$

where Φ is the standard normal cumulative distribution function, and H and F are defined as above. $PC_{\max}$ is equivalent to taking the average d' of both sides with a $\sqrt{2}$ scale factor to account for the 2AFC aspect making the task easier. For more details see Macmillan and Creelman (2004).

Levels of stimuli are chosen so that the resulting psychometric function crosses 75 % correct. Ideally it would be best to sample the full range of the psychometric function (from 50 % to 100 %), but with the ferrets there was a tendency to sample from the top to maintain 'stimulus control'. Near detection threshold of the signal, and below, the task becomes very difficult and the ferret will get many trials wrong. Without the water reward acting as positive reinforcement the animal can become disheartened and lose the understanding of the task and start to simply randomly answer even when the task becomes easy and the signal is clearly audible. This is known as losing stimulus control, which adversely affects the results. Therefore making sure there are always trials that are easy, i.e. sampling the top of the psychometric function, maintains good behaviour. In sessions where the signal was fixed in level, in a proportion of trials (roughly 20 %)

the signal was played at a higher level to ensure the task was easy, for the reasons just mentioned. These easier trials were discarded.

Once all the points of a psychometric function were collected for a particular set of stimuli (i.e. signal frequency, notch width and temporal arrangement), a logistic function was fitted to the $PC_{\max}$ vs SNR points using a least squares error method. The formula of the equation used is as follows:

$$logistic(SNR; \alpha, m, s) = \frac{1}{2} + \frac{\frac{1}{2} + \alpha}{1 + e^{-\frac{SNR-m}{s}}} \tag{3.5}$$

where s describes the slope of the function, m the position of the centre, and 1 - α the maximum of the function. Some ferrets 'guess' the answer for a proportion of trials no matter how detectable the signal is. This means that the psychometric function will never reach 100 %, and so the α value is included to account for this proportion of guesses. In general this logistic function fit the psychometric data very well.

3.2.5 Bandwidth estimates

When bandwidth was being calculated, the procedure described above would be repeated for a number of different notch widths for a single signal frequency, so that there were a number of psychometric functions. The logistic function in *equation* (3.5) would be fitted to all these psychometric functions, using a least squares error approach, in parallel such that all had the same α and s. This is because one of the consequences of the assumption of linearity, required for the power spectrum model, is that only m should change between notch widths. Using this method also gave much more stable fits. Threshold for each notch width was then taken at the 75 % correct point of the corresponding logistic function (although in the parallel fit case this is arbitrary since the difference between each logistic function is the same at all proportions correct). An example can be seen in Figure 3.3. An AF function was then fitted to the threshold SNRs using the following procedure.

1) The spectra of the stimuli were built by taking their spectral shape and scaling them to the level at the detection threshold of the signal. For example, for a single stimulus condition a vector 20,000 elements long of 0s was created, each point representing a 1 Hz bin in frequency space. Ones were inserted at every frequency bin in which the masker had energy. These bins were then scaled to have the correct height in µPa corresponding to the level of the masker at

detection threshold. This was repeated for the signal, and the two summed. The resulting vector represents the sound energy (in frequency space) for that stimulus condition. This was repeated for each notch width using the levels of each stimulus at detection threshold.

2) The AF function was evaluated, with initial parameters (established through experimentation), over the range 1-20 kHz (using 1 Hz bins). Several initial parameters were tried to ensure a stable fit.
3) This AF vector was then multiplied with the stimulus sound energy vectors described in step 1) and the results summed. This gives the total power of the stimuli passing through the AF at threshold for each notch width.
4) The resulting power values were converted into dB (by $\times\ 10\log_{10}(values)$) and the amount each value deviated from the mean was calculated.
5) A least squares method was then used to optimise the AF parameters to minimise the difference between these values across notch widths.

This method is slightly different to that described in 2.2.7 Analysis*:* Notched noise analysis but has the advantage of not needing to fit a gain parameter, which makes the fits more stable, and is more in line with the assumption that detection threshold is equivalent to a constant SNR at the output of the AF (see 1.3.2 Power spectrum model for more details).

The final step is to calculate the ERB, as described in 2.2.7 Analysis*:* Notched noise analysis. Because this method of fitting the AF shape doesn't involve a gain parameter, gain is arbitrarily set at 1, meaning the area under the curve is equivalent to the ERB.

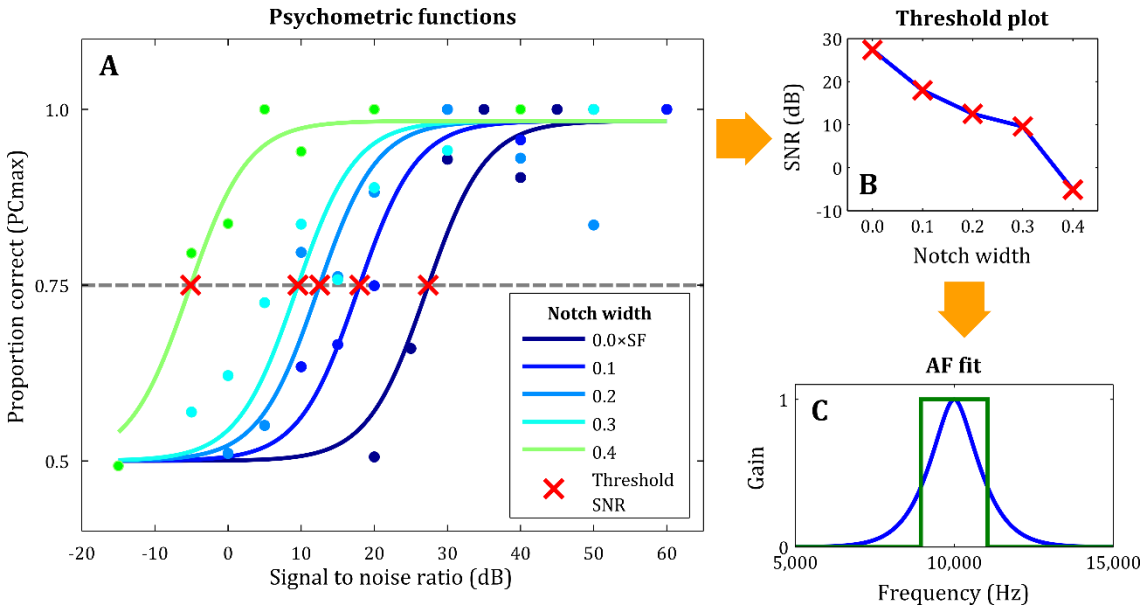


Figure 3.3

Example of stages in analysis. ***A*** *Psychometric functions for each notch width, with logistic functions fitted in parallel. Threshold SNR taken at 75 % correct.* ***B*** *Same threshold SNR values from panel A plotted against notch width. This threshold plot represents the drop in threshold with increasing notch width.* ***C*** *AF is fitted to the threshold SNR by constructing spectral shapes of stimuli (scaled to the appropriate threshold SNR), integrating over the AF, and minimising the difference in SNR across notches. ERB of the AF is calculated as the width of the rectangular filter with equal height and power (green line).*

3.2.6 Bootstrapping

In some cases, when it was desirable to obtain a confidence region on a result, a bootstrapping method was used. Figure 3.5, Figure 3.7 and Figure 3.8 show examples of psychometric functions with confidence intervals for each point, and Figure 3.9 shows bandwidth estimates with confidence regions. To obtain these, the original trial results for a given stimulus condition were taken and recreated as a vector of 0s and 1s; 0 representing an incorrect answer, and 1 a correct one. A simulated result was then created by selecting from these trials with replacement. For example, imagine a ferret gets 20 presentations of a certain signal with a certain level, combined with a certain notch width masker of a certain level, in a certain temporal arrangement, and gets 13 trials correct. A vector would be created with 13 ones and 7 zeroes. A random element of this vector would be selected 20 times to create a simulated proportion correct for that specific stimulus condition. (This is equivalent to selecting a sample from a binomial distribution with $p = 13/20$ and $n = 20$). This is done for every stimulus condition to obtain a simulated data set. Whatever analysis that is applied to the real data is then applied to this simulated data set and a simulated result is produced. In the case of the psychometric function this means a simulated proportion correct for an SNR, and a simulated bandwidth estimate in the bandwidth case. If this is repeated a number of times (usually at least 500), a distribution of points around the true calculated value is produced. Taking the 5 % and 95 % quantiles of this distribution gives the approximate 95 % confidence interval which is then applied to the original data points.

3.2.7 Stimulus arrangement in novel method

As mentioned previously, a novel method was used to obtain bandwidth measurements in the ferret with the behavioural task described above. It involves notched noise masking but unlike almost all other animal behaviour studies, allows forward as well as simultaneous masking in the same animal without retraining. Essentially it involves playing the masker in repeated bursts with silent gaps between, and interleaving short signal bursts in these gaps. By ensuring that both stimuli have the same cycle length (duration of stimulus burst plus silence), it is easy to induce forward, partially forward and simultaneous masking by simply varying the time delay between the two. Figure 3.4 shows an example of simultaneous and forward masking with 4 signal bursts. The standard approach used in almost all animal behaviour studies is to use simultaneous masking with the masker playing constantly and the signal briefly appearing within it. A similar stimulus arrangement to the new method was used to induce forward masking in the chinchilla in McGee et al. (1976), but was not used for the simultaneous masking.

The benefit of using the same 'bursting' approach for both cases is that the ferret need not be retrained for the separate tasks. Instead the temporal relationship between signal and masker can be switched between trials in the same sessions with little notice.

The benefits described above are all theoretical and the method could be rendered useless by several possible problems. It therefore had to be tested before it could be deemed appropriate for measuring frequency selectivity. Also stimulus parameters had to be refined to give the best possible measures of bandwidth. This validation and refinements are described below. (Some parameters of the stimuli, e.g. duration and envelope, vary seemingly inconsistently between consecutive sections below. This is because the measurements were not necessarily made in the order they are presented, so refinements made in one section may not be present at the beginning of the next).

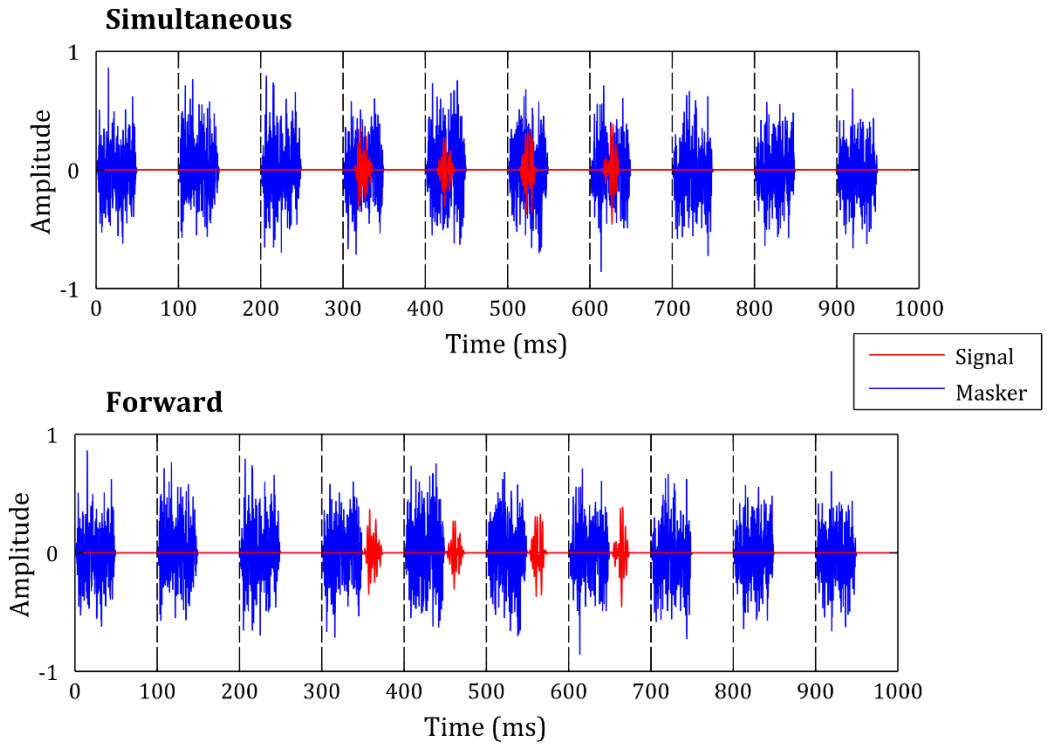


Figure 3.4

Temporal arrangement of stimuli. Ten masker bursts with 4 signal bursts in the simultaneous and forward masking arrangements. Masker bursts are 50 ms in duration followed by 50 ms silence between. Signal bursts are 25 ms in duration with 75 ms silence between. A cycle is therefore 100 ms.

3.2.8 Testing on a human subject

There was a chance that the constant bursting of the stimuli would create confusion, resulting in informational masking that would make it impossible to produce a reliable psychometric function based on sound masking. Also the brief duration of the stimuli might mean very little masking occurs, making any meaningful notched noise experiment inaccurate and unusable. The first set of tests were therefore conducted on myself. I used 5 signal bursts, 50 ms in duration (inc. 25 ms cosine squared on/off ramps), embedded in 50 ms masker bursts (inc. 25 ms cosine squared on/off ramps) with 50 ms of silence between bursts. The signal was 1/32 octave wide narrowband noise centred on 10 kHz. The masker was 1 kHz wide narrowband noise linearly centred on 10 kHz played, appearing at the same time as (simultaneous) and immediately preceding (forward) the signal. Masker level was kept fixed at 30 dB SPL and signal level varied. Stimuli were generated in Matlab and delivered through a pair of headphones (Sennheiser HD 600) in a sound attenuated booth (Industrial Acoustics Company, Winchester, UK). In a single trial the signal would be played randomly in the left or right ear with equal probability and I had to state which side it came from. Figure 3.5 shows the psychometric functions produced from 50 repetitions of each signal level for each temporal arrangement. A logistic function has been fit to each set of results.

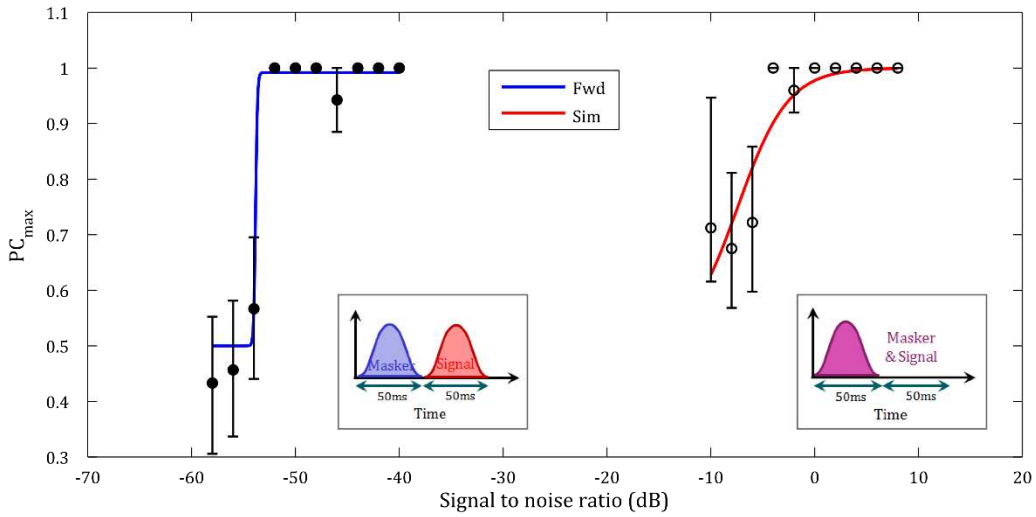


Figure 3.5

Human test of the behavioural task. Four 50 ms narrowband noise signals (25 ms cosine squared on/off ramps, 1/32 octave around 10 kHz) embedded in 50 ms masker bursts (25 ms cosine squared on/off ramps, 1 kHz wide narrowband noise centred on 10 kHz). In the forward masked condition the signal and masker immediately alternated, in the simultaneous arrangement the two appeared at the same time with 50 ms silences between. Error bars made with bootstrapping.

The first notable feature is that both psychometric functions are reasonably monotonic; proportion correct increasing with SNR. This is as expected for such a task. Also, whilst performing the task I did not experience any confusion between the stimuli, and both temporal arrangements sounded very similar meaning switching between the two was not problematic. A second notable feature of the figure is the almost 40 dB difference in threshold between the forward and simultaneous masking arrangements. Simultaneous masking is known to be much stronger than forward masking so this fits in with what would be expected. The difference between the two, however, is surprisingly large, which suggests that the amount of forward masking is very small. Three ways in which forward masking can be increased is to a) reduce the distance between signal and masker, b) increase the duration of the masker, or c) reduce the duration of the signal. Since there is no gap between signal and masker so this could not be reduced, the signal was reduced to 25 ms, and the masker made longer and its envelope changed.

3.2.9 Room acoustics

Another potential problem is the acoustics of the ferret arena. Although every effort is taken to reduce echoes in the behavioural room, i.e. sound proof booth, minimal reflective surfaces, cage suspended above the ground, it is not anechoic. Therefore there is potential for echoes to 'smear' the stimuli temporally meaning that the short gaps between stimuli are actually being 'filled in'. This would mean that when the stimuli are in the forward masking arrangement, true forward masking is not actually occurring since masker energy is still present. To test whether this could be a problem, recordings of the masker and signal bursts were made in the ferret arena, using the same setup for recording MLS described in 3.2.2 Ferret arena. Figure 3.6 shows the average envelope of 2500 signal and 4200 masker cycles, in dB. The envelope of the original sound bursts are overlaid for comparison. The signal was the same narrowband noise used in 3.2.8 Testing on a human subject above, but reduced to 25 ms (inc. 12.5 ms cosine squared on/off ramps). The masker was the same 1 kHz wide narrowband noise used in 3.2.8 Testing on a human subject, but the envelope was changed to 150 ms in duration with 5 ms on/off ramps, since the longer duration and steeper ramps produce more masking. (Steeper slopes also produce more spectral splatter, a spread of power across frequency, but the difference between the two ramp lengths is small enough to not adversely affect the experiment).

Panels A and B of Figure 3.6 show that the signal and masker do indeed continue as echoes after they have stopped being played (since the blue lines deviate from the red

lines). These echoes start roughly 20 dB down from the peak and decay at ~1 dB/ms for both stimuli. Since the detection threshold of the signal is being measured in this experiment, the levels of stimuli are very low and in many cases these echoes will be below the level of audibility for the ferret, so their impact should be small. However, this is an issue to be aware of; as with all free-field experiments there is the inevitable possibility of non-pure forward masking.

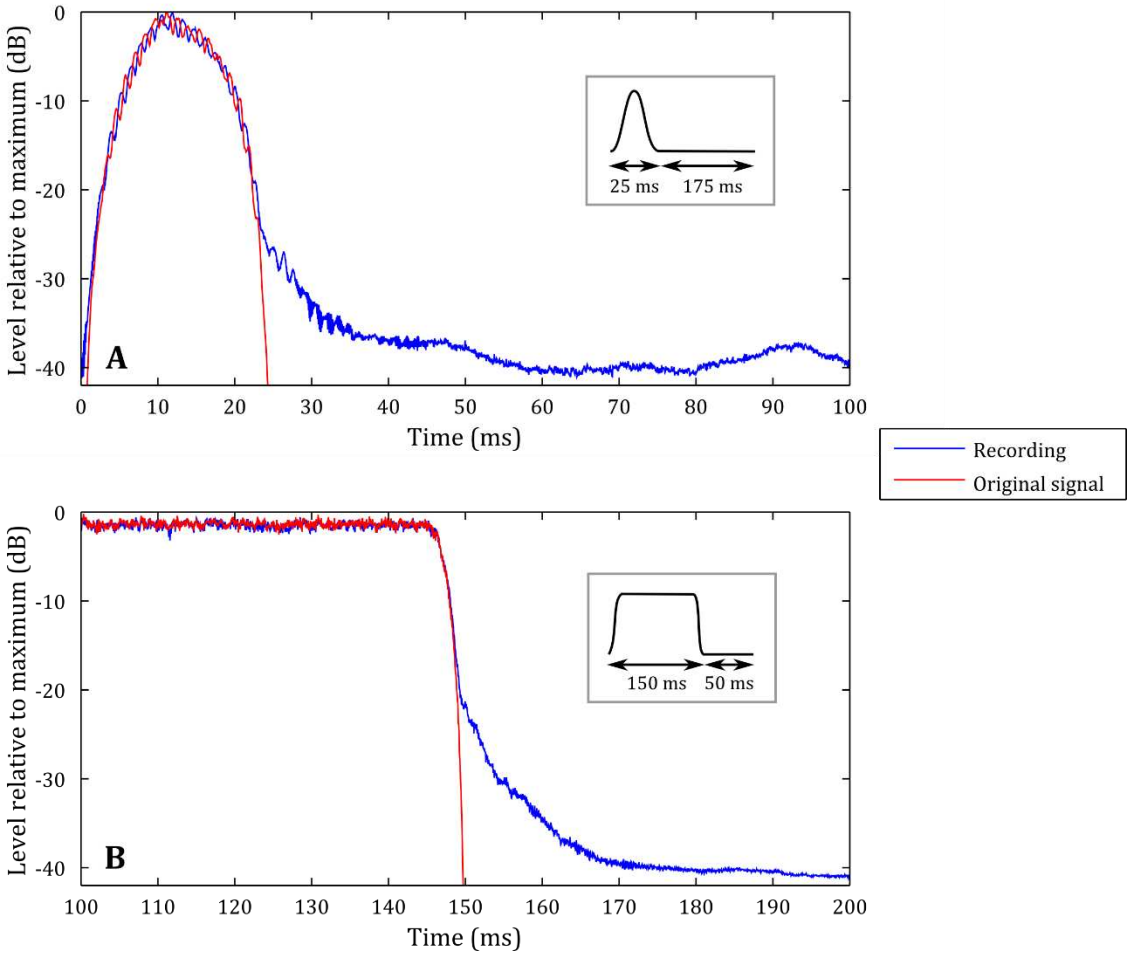


Figure 3.6

Effect of room reverberations. ***A*** *Recording of 2500 cycles of the 25 ms cosine-squared signal averaged together and the envelope taken (blue line). The envelope for a single cycle of the original signal is overlaid for comparison (red line). Both are presented in dB relative to the maximum to allow easy comparison.* ***B*** *Same but for the 150 ms masker. Only the region around the end of the masker is shown since the rest does not deviate from the original. The amount the recording deviates from the original shows the temporal effect of reverberations.*

3.2.10 Testing on a ferret

The final hurdle was to see whether the ferrets took to the task or not. I didn't find the bursting of the stimuli confusing, but would the ferrets? Also, would this method provide enough masking to effectively explore the edges of the AF? The only way to answer this question was to try it out on a ferret.

A ferret was shaped and trained to perform the behaviour task described in 3.2.3 Behavioural task with the bursting stimuli. Firstly the experiment described in 3.2.8 Testing on a human subject was carried out in the ferret; forward and simultaneous masking with 50 ms narrowband signal and 50 ms masker. The signal was the same 1/32 octave wide narrow band noise centred on 10 kHz used before. Pure tones were not used due to their propensity to form standing waves; the reflection of a pure tone combines with the original pure tone, producing areas in the room where the level is higher than the original, and areas where the signal is obliterated (Auld, 1973). Because a pure tone only has one phase value, this event is likely. Noise contains many frequencies with random phase so this ordered summation is far less likely. Therefore narrow band noise gives much more control over stimulus level across the arena. Twenty signal bursts were played over the space of 2 second in each trial; much shorter and the task would be too hard for the ferret, much longer and the signal would still be audible once the ferret had left the central platform, compromising control over the stimulus level.

Figure 3.7 panel A shows the psychometric functions produced for both temporal conditions. Very encouragingly both conditions yielded clear, monotonically increasing psychometric functions, transitioning smoothly from 50 % to just under 100 % correct. The maximum of both functions was less than 100 % due to the 'guess' rate described earlier. Also the ferret psychometric functions are shallower showing more of a variability in detection threshold. Shallower slopes suggest more confusion in the task (Green, 1995) but these features are commonly seen in animal psychophysics (Tolhurst et al., 1983; Gleich et al., 2006) and do not suggest the task is too difficult. Also again the simultaneous masking arrangement has a higher detection threshold than the forward one. Remarkably, the two are very similar to the human ones, being only 5 dB higher in both cases.

So the ferrets took to the task and gave usable psychometric functions. But what of the amount of masking? To test whether there is enough masking to explore the filter shape, a wide notched noise masker was compared to the no notch condition, in the

simultaneous masking arrangement. The signal was the same 1/32 octave narrowband noise centred on 10 kHz as before, except the duration was 25 ms (inc. 12.5 ms cosine squared on/off ramp) for the reasons described in 3.2.8 Testing on a human subject. The no notch masker was the same 1 kHz wide narrowband noise as before with 50 ms duration, and the wide notch was made up of two 500 Hz wide bands of noise separated by 10 kHz, centred linearly on 10 kHz, also 50 ms in duration. (This can also be referred to as 0.0 × SF and 1.0 × SF, where SF is signal frequency). All notches that would be used to measure the bandwidth fall in this range, so if the difference in threshold between the two is very small, then the dynamic range is not enough to obtain any accurate results. The psychometric functions produced using both these maskers can be seen in Figure 3.7 panel B. Again both functions are monotonic and of similar steepness to those in panel A. The no notch function may be slightly shallower but this could be due to the relatively low number of trials. Importantly there is a 40 dB difference between the two which should provide enough space for further psychometric functions to slot in between, allowing for accurate AF measurement.

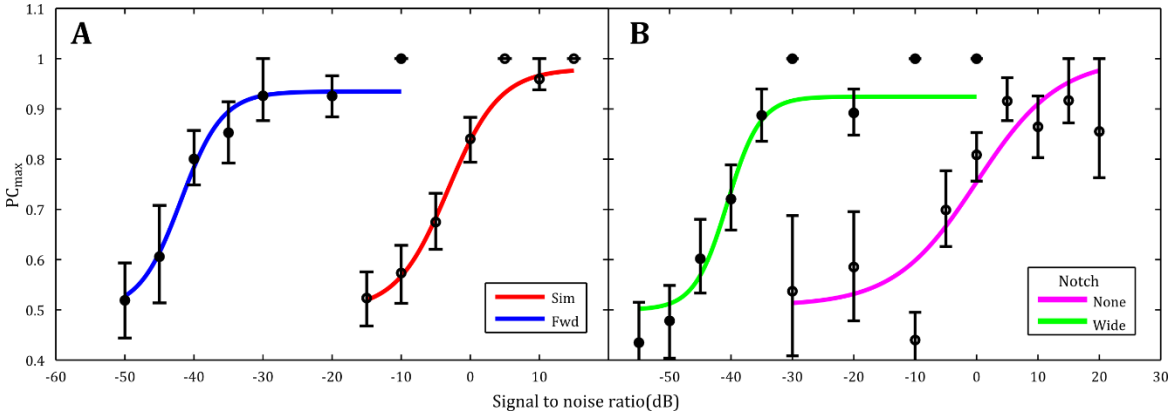


Figure 3.7

Psychometric functions from a single ferret using the novel stimulus paradigm. ***A*** *Simultaneous vs forward masked arrangements with logistic function fit to each (red and blue lines). Signal is 10 consecutive 1/32 octave narrowband noise bursts centred on 10 kHz, 50 ms in duration (25 ms cosine squared on/off ramps). Masker is continuous 50 ms 1 kHz wide narrowband noise centred on 10 kHz bursts separated by 50 ms. In the simultaneous case both stimulus bursts start together, in the forward case they alternate without gap. Confidence intervals obtained through bootstrapping method.* ***B*** *No notch vs wide notch (10 kHz wide centred on 10 kHz) with 500 Hz wide noise bands (pink and green lines). Masker duration same as in panel A. Signal has the same frequency properties as in panel A but is shorter (25 ms with 12.5 ms cosine squared on/off ramps). Both stimulus bursts have peak intensity at the same point in a cycle (25 ms) causing simultaneous masking.*

One final test that was done was to see how masking changed for intermediate masker and signal arrangements, between fully simultaneous and fully forward. We know that simultaneous masking is stronger than forward, so if the masker level is kept fixed and the signal position slowly moved further away from the masker, then the detection threshold of the signal should decrease since less masking occurs. This transition should be gradual if we are truly seeing the effects of sound masking. Figure 3.8 shows the results of finding proportion correct for various masker and signal positions. In panel A the 50 ms with 25 ms cosine squared on/off ramps was used, with the 25 ms signal. The envelope of each burst can be seen superimposed onto the plot; masker in blue, signal in red. Each proportion correct point (black circle) corresponds to the position of the signal it sits in the centre of, each of which are relative to the masker burst. For example, the $PC_{\max}$ point at 25 ms corresponds to the simultaneous masking position where the signal and masker peak simultaneously. The point at 62.5 ms corresponds to the signal immediately following the masker. Masker level was fixed midway between the two psychometric functions in Figure 3.8 panel A so that the signal was fully masked in the simultaneous position, and not masked in the forward position. And indeed the proportion correct increases gradually as the signal moves further from the masker, plateauing at ~90 % correct. The increase is rather rapid though, so the same duration masker with steeper slopes was tried, and the results can be seen in panel B. Here the slope is much shallower between the 25 and 37.5 ms positions, resulting in proportion correct reaching a maximum at 62.5 ms instead of at 50 ms in panel A. This is due to more masker energy being present and the sharper drop. Both these plots show that the type of masking we are seeing is consistent with true sound energy masking and that the novel stimuli are appropriate for AF measurements.

So the testing seems to suggest that there are no major hurdles to using this method and it should yield suitable bandwidth estimates. These bandwidths were measured at different SF and using slightly different methods, all of which is described in the following section.

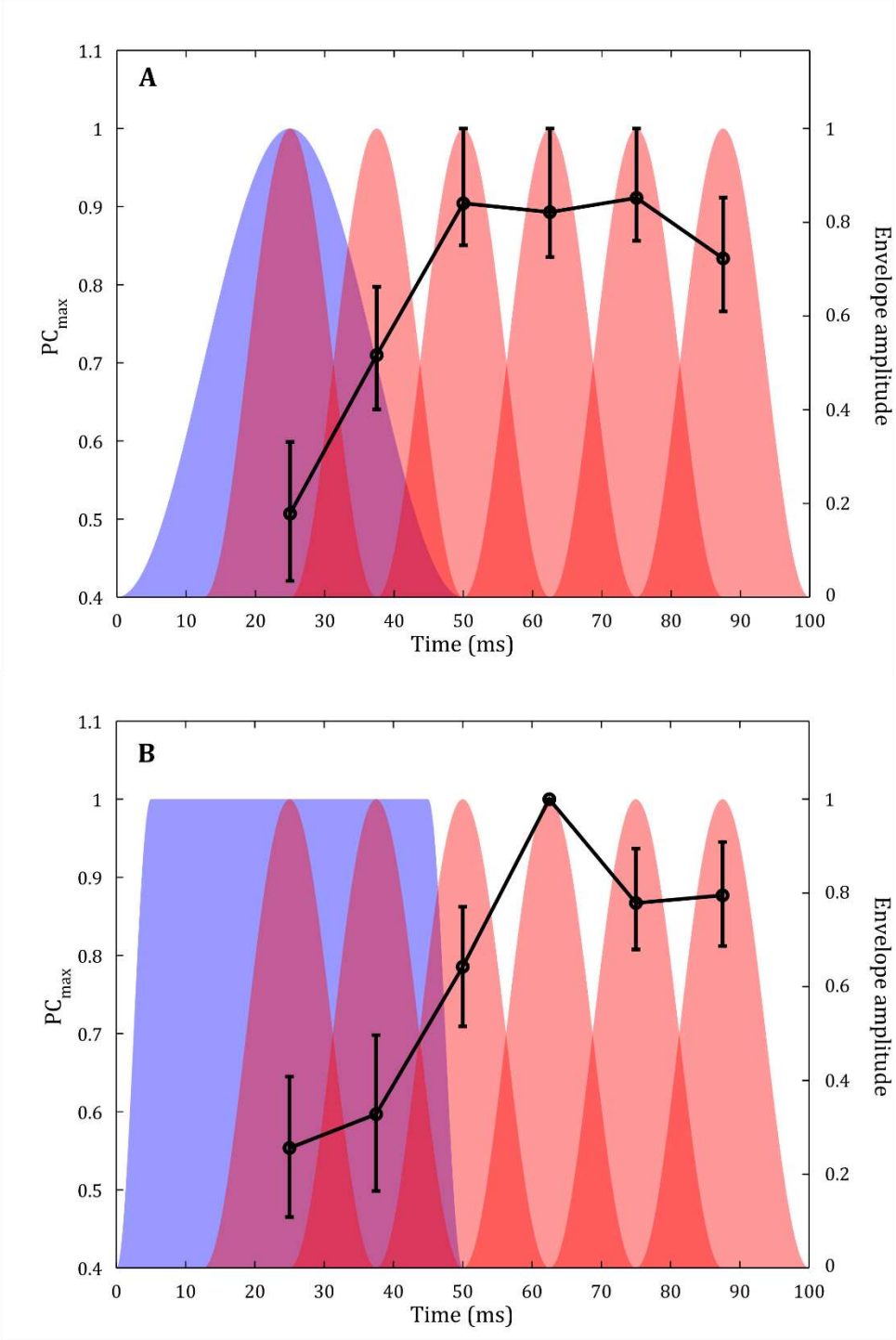


Figure 3.8

Proportion correct versus stimulus temporal arrangement. Blue shaded area represents the shape and position of the envelope of the masker in a single cycle. Red shaded areas represent the shape of the envelope of the signal, and its various positions relative to the masker. Signal has a duration of 25 ms (inc. cosine-squared on/off ramps of 12.5 ms). Scale on right axis. Black points and lines show PC_{max} *for each signal position, i.e. 87.5 ms point represents the* PC_{max} *for the trials where the signal starts 75 ms after the masker starts. Confidence intervals calculated using bootstrapping method.* ***A*** *represents results for the 50 ms masker, with 25 ms cosine-squared on/off ramps.* ***B*** *represents results for the 50 ms masker, with 5 ms cosine-squared on/off ramps.*

3.3 Results

After the novel method had been validated, the task of calculating AF size and shape began. The overall aim was to measure AF widths using fixed signal level, forward masking, but it would also be interesting to see how the different methods affect results. With that in mind, both forward and simultaneous masking, with fixed signal level and fixed masker level bandwidth estimates were made. The results can be seen in Table 3.2 and Figure 3.9.

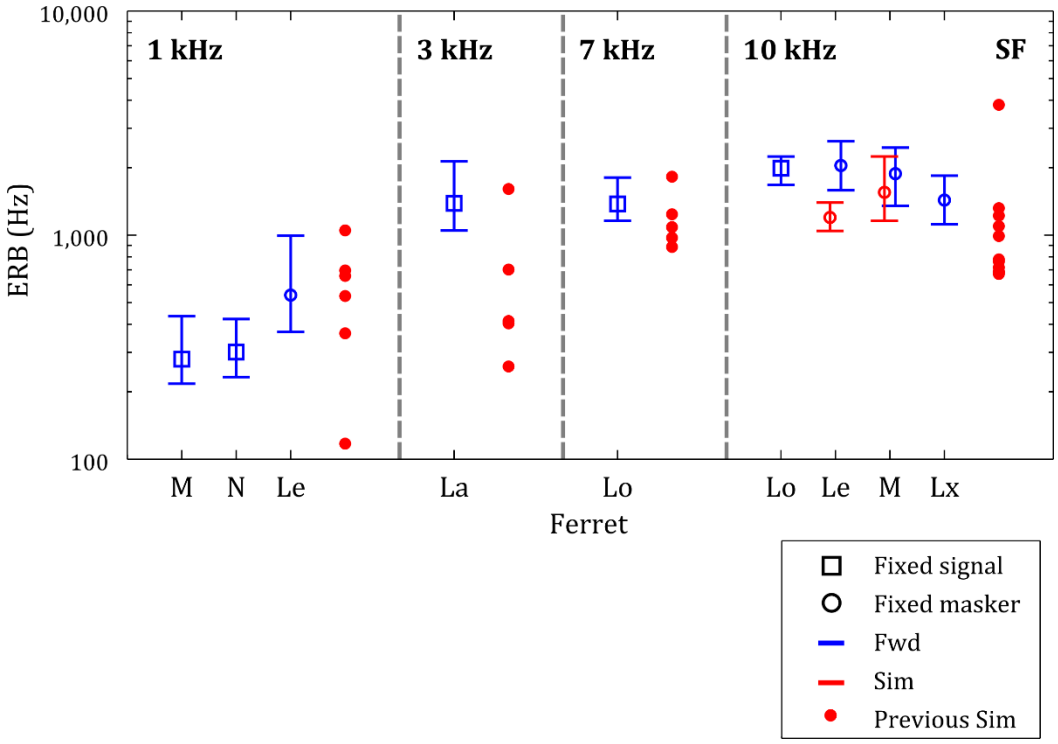


Figure 3.9

Bandwidth estimates of individual ferrets at different signal frequencies using different methods. Signal frequencies are 1, 3, 7 and 10 kHz. Square points show fixed signal level (10 dB above threshold), and circular points show fixed masker level (60 dB SPL) measures. Blue points are in the forward masked arrangement (no gap between signal and masker), and red points show the simultaneous arrangement. All error bars produced using bootstrapping method, n *= 500. Red filled circles are bandwidth estimates made in the ferret using fixed masker, simultaneous masking by previous researchers.*

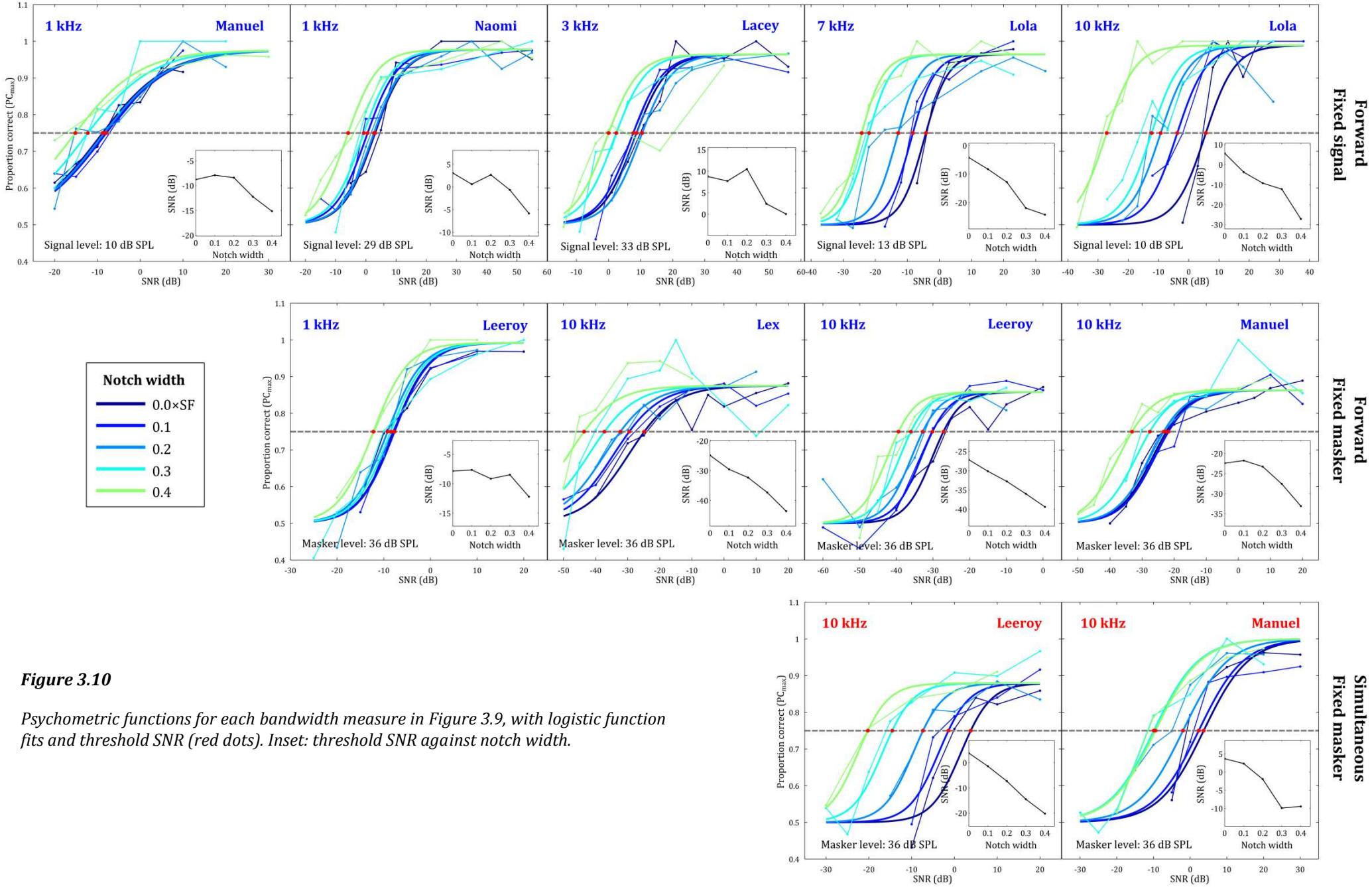


Figure 3.10

Psychometric functions for each bandwidth measure in Figure 3.9, with logistic function fits and threshold SNR (red dots). Inset: threshold SNR against notch width.

SF (kHz)	Fixed	Fwd/ Sim	Ferret	ERB (Hz)	95 % CI bound	
					Lower	Upper
1	Signal	Fwd	Manuel	230	217	434
1	Signal	Fwd	Naomi	301	233	422
1	Masker	Fwd	Leeroy	540	370	995
3	Signal	Fwd	Lacey	1387	1049	2136
7	Signal	Fwd	Lola	1378	1160	1806
10	Signal	Fwd	Lola	1989	1676	2242
10	Masker	Sim	Leeroy	1199	1044	1397
10	Masker	Fwd	Leeroy	2045	1586	2626
10	Masker	Sim	Manuel	1553	1158	2244
10	Masker	Fwd	Manuel	1880	1350	2457
10	Masker	Fwd	Lex	1433	1117	1842

Table 3.2

Bandwidth results for different ferrets, methods and signal frequency.

The same analysis method was used for each of the ferrets and signal frequencies shown in the figure; $PC_{\max}$ calculated, logistic functions fitted in parallel to the resulting psychometric functions, SNR at threshold taken for each notch width, and a roex(p,r) function fitted using an optimisation procedure (see 3.2.5 Bandwidth estimates for more details). In some cases (empty circles), the masker level was kept fixed at 60 dB SPL, and in others (squares) the signal level was fixed at 10 dB above the absolute detection threshold of the signal at that frequency. Absolute threshold was calculated by building a psychometric function for the signal without a masker and taking the signal level at 75 % correct. Figure 3.10 shows the psychometric functions and their logistic fits for each of the bandwidth measurements in Figure 3.9. Inset for each panel is the corresponding threshold against notch width plot, which is used to fit the AF function. In each case masker bursts were 150 ms in duration (inc. 5 ms cosine squared onset/offset ramps) with 10 consecutive 25 ms signal bursts (inc. 12.5 ms cosine squared onset/offset ramps), both repeating in 200 ms cycles. In some cases simultaneous

masking was used (red points in Figure 3.9), with the signal appearing 100 ms after the start of a masker burst (ending 20 ms before the masker starts to ramp off), and in others forward masking was used (blue points), with the signal appearing immediately after the masker. The signal had frequency components 1/32 of an octave around the frequency stated. Frequency parameters of the masker were similar to those used in Oxenham and Shera (2003); two narrow bands of noise 0.25 × SF wide, with 5 different notch widths, 0, 0.1, 0.2, 0.3 and 0.4 × SF spaced symmetrically (in linear frequency space) about SF. Error bars were made using the bootstrapping method described in 3.2.6 Bootstrapping with n = 500. Bandwidth estimates made by other researchers (red filled circles) have been included for comparison. These were made at the Institute in the same behavioural arena, with ferrets from the same colony, using the same equipment, and methods, i.e. PC_{max}, roex fitting, etc. (Sollini and Alves-Pinto, unpublished). The stimuli were different however: masker was constantly played at a fixed level in a behavioural session, signal was presented in a single burst longer than 25 ms (simultaneous masking).

Overall, the results at each SF are fairly consistent across method and ferret, at least at 10 kHz where the majority of bandwidths were measured. Certainty of bandwidth (i.e. error bar size) also appear to be fairly evenly spread, without a bias towards method, frequency, or ferret. The only comparisons between forward and simultaneous masking are at 10 kHz, with fixed masker. Both forward and simultaneous bandwidths were estimated in two males; Manuel and Leeroy. Surprisingly for both ferrets the simultaneous masked bandwidths are actually narrower than the forward ones (1199 against 2045 Hz, and 1553 against 1880 Hz for simultaneous and forward masking respectively); when taking into account the confidence intervals of the estimates this difference is only significant for Leeroy. A comparison of fixed signal and fixed masker levels can be made at 1 kHz and 10 kHz. At 1 kHz the fixed signal level estimates of Manuel and Naomi (female) are remarkably similar (230 and 301 Hz respectively) as are the 95 % confidence intervals (217-434 and 233-422 Hz respectively). The fixed masker level, by comparison, gives a broader bandwidth (540 Hz) and larger confidence interval (370 - 995 Hz). This broader fixed masker level is, however, not seen in the 10 kHz ERBs which all match up fairly closely, with the Lex (male) measure being slightly (and significantly) narrower. Finally, a comparison between the novel method and the traditional method (filled red circles) reveals not a great difference, with every new measure falling within the range of the old. At 1 kHz the forward fixed masker measure of Leeroy matches very closely the bandwidth and spread seen in the

traditional, simultaneous, fixed masker condition; whereas the fixed signal measures of Manuel and Naomi are on the narrower end of this range. At 3 kHz the new method gives a quite broad measure which is still within the range of the old method, and the same is true at 7 kHz. At 10 kHz the new method tends to give slightly broader bandwidth estimates, although they are all still within the range of results for the traditional method.

Measuring bandwidth at several signal frequencies allows us to see how bandwidth changes with frequency, an important feature often discussed in the previous chapter. Figure 3.11 shows the forward masked, fixed signal psychometric bandwidth measures plotted beside physiological bandwidth measures in the ferret. One (grey filled circles) is from SU in the ferret AN made using the FRA (with roex function fit), a method described in more detail in 2.2.7 Analysis*:* FRA analysis (Sumner and Palmer, 2012). The other is from SFOAEs of the ferret, some of which are the same animals used in this thesis. These measurements were carried out at the Institute by Andy Oxenham and Chris's Sumner, Bergevin, and Shera, using the method described in Shera et al. (2010b) and are yet to be published. The brown line shows the mean and the yellow area the 95 % confidence region. Both sets of peripheral data match up very closely with one another, and correspond perfectly with the two behavioural measures at 1 kHz in both bandwidth and spread. The higher psychophysical frequencies do not match up quite as closely, with the largest difference being seen at 3 kHz, with the psychophysical bandwidth estimate being roughly twice as broad. (This point, however, had the messiest psychometric functions and non-monotonic threshold plot, which is reflected in the large CI). At 7 and 10 kHz the bandwidth estimates are only slightly broader than the mean AN and OAE data. However, there is a general trend for bandwidth to increase with SF, which is a feature of frequency selectivity seen throughout auditory science.

As well as different stimulus arrangements, different methods were also tried in the analysis stage, to make results comparable to a wider range of studies and to see the effect decisions in analysis have on the results. The first was to fit gammatone functions to the psychophysical data (another popular AF function described in more detail in 1.4 Auditory filter shape) instead of the roex. The main differences between the two are that the GTF has one fewer parameter and a fixed slight asymmetry. The GTF has a fixed order value which controls the steepness of its slopes; in human studies it is usually set at the 4th order (Patterson et al., 1992). Since human bandwidth estimates tend to be narrower than those of animals (Glasberg and Moore, 1990; Oxenham and Shera, 2003), and to my knowledge there is no established order of GTF used in ferrets, I fitted several

orders of GTF. The results can be seen in Figure 3.12, beside the roex fitted results from Figure 3.9 for comparison. The first clear feature is the similarity between the two filter types; for most of the measurements, regardless of frequency, method or ferret, all GTF order ERBs fall within the 95 % CI of the roex. (Perhaps this is why the GTF has not superseded the roex, and both are still used to represent AF today). The other notable feature is that for 4 of the data sets the GTF order has a significant impact, with lower orders giving narrower bandwidth estimates; for Lola's 10 kHz data the difference between the 2nd and 8th order GTF is almost a factor of 8. Also the confidence regions for the higher orders tend to be the smallest and match up most closely with that of the roex. (The biggest violation of this is Lacey's 3 kHz measure, which, as previously mentioned, is particularly messy). This highlights the importance of choosing the correct order when fitting GTF.

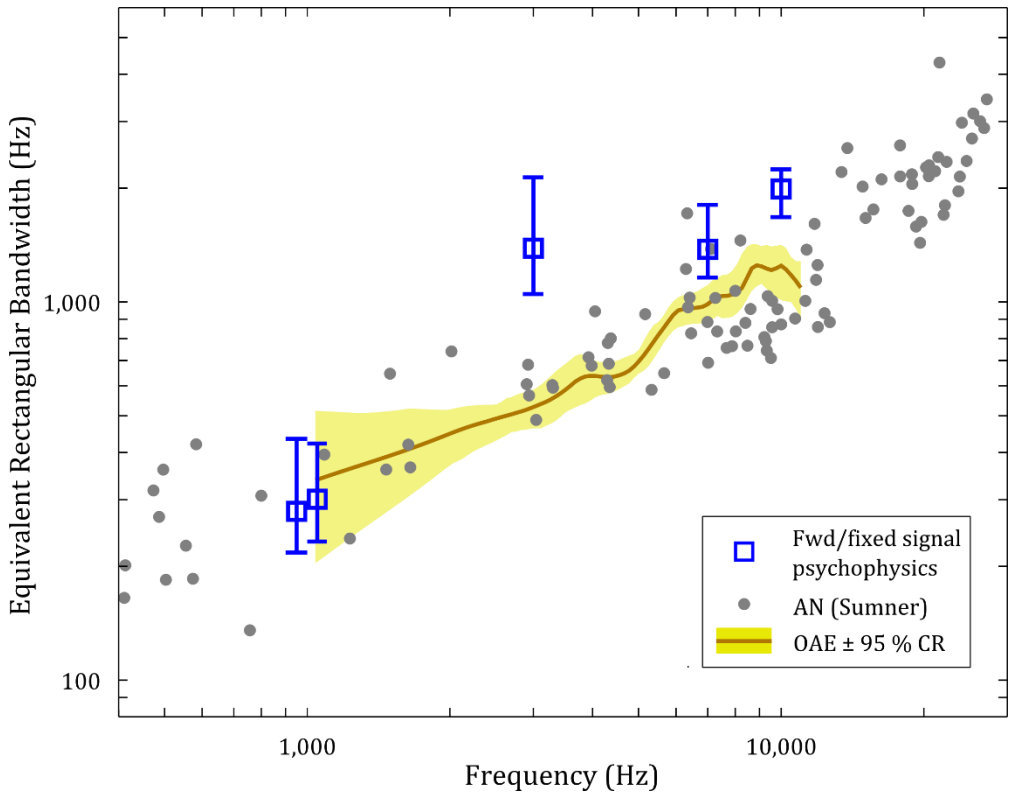


Figure 3.11

Comparison of psychophysics and physiology in the ferret. Blue squares are forward masked, fixed signal, psychophysical data from Figure 3.9 showing bandwidth versus signal frequency. Grey dots show bandwidth estimates made from SU in the ferret AN plotted against CF taken from Sumner and Palmer (2012). Brown line and yellow shaded area show mean OAE measured bandwidth of the ferret with 95 % confidence region (unpublished).

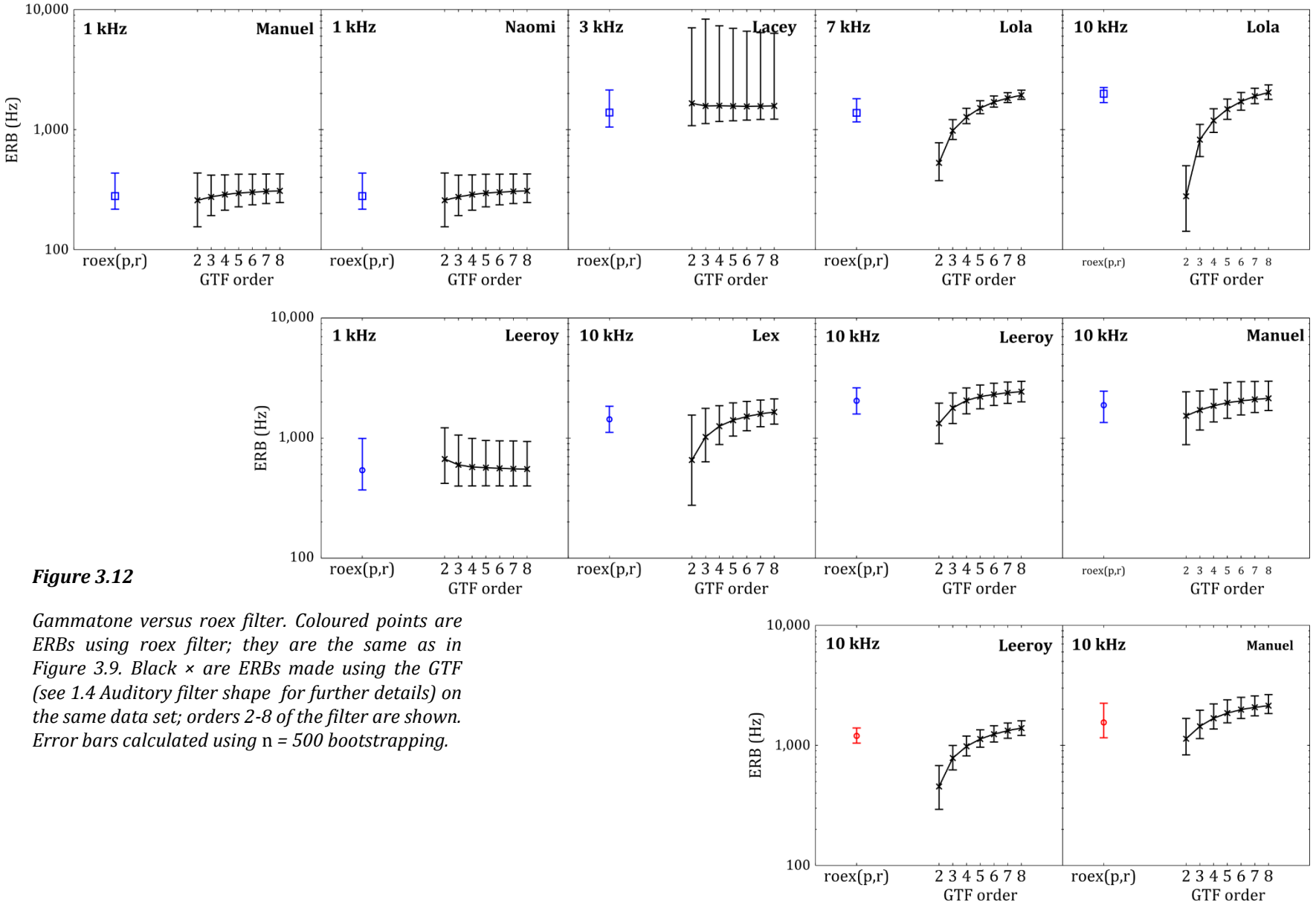


Figure 3.12

Gammatone versus roex filter. Coloured points are ERBs using roex filter; they are the same as in Figure 3.9. Black × are ERBs made using the GTF (see 1.4 Auditory filter shape for further details) on the same data set; orders 2-8 of the filter are shown. Error bars calculated using n = *500 bootstrapping.*

Another difference in analysis that was tried was to do away with the logistic fit of the psychometric functions and simply use the raw data to determine threshold. In this approach threshold SNR is taken as the point at which the straight line connecting two consecutive raw psychometric data points intercepts the 75 % correct line. This method does not take into account any other points on the psychometric function and is susceptible to small fluctuations in data points near threshold. The results of this slightly different method can be seen in Figure 3.13. Each plotted pair of points shows the logistic fitted results in the left (same as Figure 3.9) and raw psychometric function results on the right. Across each of the measures there is no systematic difference between the two methods; in some the raw bandwidth is marginally higher and in others marginally lower. One clear trend, however, is that every one of the 'raw' measures has a larger confidence interval; in some cases this difference is quite extreme (3 kHz Lacey, 10 kHz Fwd Leeroy, 10 kHz Lex). Also the bootstrapping method used to obtain the error bars had a higher failure rate, meaning many more of the simulated results could not give a bandwidth estimate. This is clearest for the 3 kHz Lacey point where the failure rate goes from 8.7 % with the fit to 82.3 % without. These findings are unsurprising since fitting logistic functions is meant to reduce variability and increase stability of the results, and not introduce bias.

One further method tweak was to not only fit a single AF, but to fit a filter bank of AF, allowing for potential off-frequency listening. According to the power spectrum model (see section in 1.3.2 Power spectrum model) the filter with the highest SNR will be used to detect a signal. Since there are continuously overlapping filters along the full frequency range of hearing, AFs with centres slightly away from the signal frequency might give a better fit. Fitting an array of overlapping filters instead of just one AF allows for this selectivity. This was done by following the same procedure described in 3.2.5 Bandwidth estimates but where the AF is fitted, a series of overlapping AFs are fitted instead, with the one with the highest SNR for a given notch width determining the AF used. The centre of each successive AF in the bank was set at 0.01 × SF steps above and below SF out as far as ± 0.2 × SF. In one instance the parameter for bandwidth (*p* for roex, *b* for GTF) was kept fixed for each successive AF, and in another instance it increased with filter centre frequency at a rate consistent with the single AF fits seen so far, i.e. a straight line was fitted to the plot of ERB vs SF of points in Figure 3.9, the slope taken, and this used to determine the relationship between bandwidth in successive AF. However, both methods yielded almost identical results, which showed very little difference to the single AF fits used so far in this study. This suggests either the ferrets

do not use off frequency selectivity, or the method is too sensitive for the data, and the simpler version does the same job.

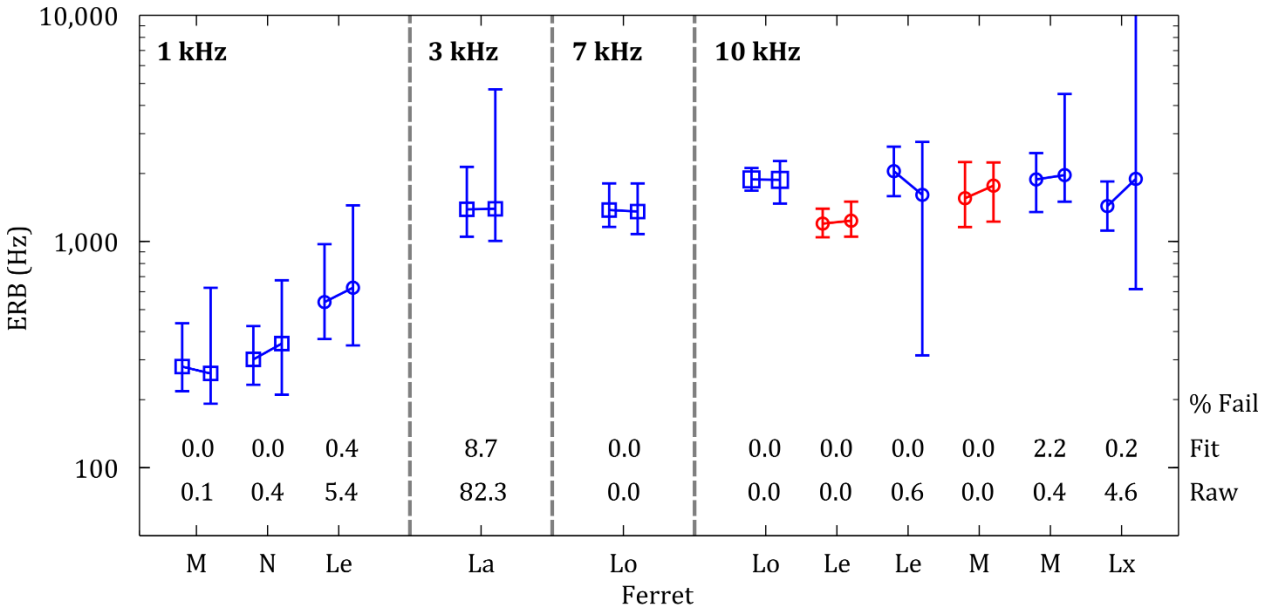


Figure 3.13

Raw psychometric functions versus logistic function fits. In each pair of data points, the left describes ERB calculated by fitting logistic functions (in parallel) to the raw psychometric data (same as in Figure 3.9), and the right point describes ERB calculated by taking threshold from the raw psychometric data, i.e. take the 0.75 PCmax crossing of the raw psychometric functions in Figure 3.10, not the logistic fits. Other parameters are the same as in Figure 3.9 and described in the legend. Bootstrapping with n = 500 used to obtain error bars. Values below each point describe the proportion of bootstrapped simulations that did not yield an ERB for the fit and raw psychometric data.

3.4 Discussion

One of the main aims of this chapter was to test the viability of a novel method for measuring bandwidth behaviourally in the ferret. If proven, the task was to optimise it and measure bandwidth, exploring the effect of slightly different experimental paradigms. The novel method should allow such comparisons in the same animals to remove subject variability. I believe that all these aims were met and the method a success.

Every effort was taken to make the room as echo free as possible to allow enough 'silence' between masker bursts to enable true forward masking. Recordings of the stimuli in the arena showed that the maximum level of the echoes is 20 dB below stimulus level, and that these decay rapidly. This is important since any echoes lingering after the stimulus has stopped could prevent true forward masking. Since the levels of stimuli used around threshold (the area of interest) are low, the effect of these echoes is minimised and should not adversely affect the experiment. The proof of concept was established in a human test, with clear psychometric functions for the forward and simultaneous masking cases separated by 40 dB. This was replicated in the ferret, with the animals seemingly showing no problems with the constantly bursting nature of the stimuli, such as confusion or loss of motivation. (In addition, the similarity in the difference between simultaneous and forward masked threshold estimates in humans and ferrets adds weight to the argument that echoes do not have an adverse effect, since the human study was conducted over headphones without these echoes). Also the dynamic range of an AF measurement was tested, by comparing a no notch with a wide notch masker condition (in the simultaneous arrangement), and the difference between the two psychometric functions (40 dB) was adequate for accurate measurement. Further investigation showed that the masking properties of the different temporal arrangements were consistent with true forward and simultaneous masking; the simultaneous masking position gave a much higher signal detection threshold, meaning masking is stronger, consistent with observations of simultaneous and forward masking elsewhere (Moore, 2003). Also proportion correct changed smoothly from 50 % to 100 % as the temporal arrangement of stimuli moved from fully simultaneous to fully forward positions, consistent with true sound masking of this nature. Finally the envelope properties of the stimuli were optimised to produce maximum forward masking (Kidd Jr and Feth, 1982; Zwicker, 1984) while still maintaining a viable behavioural task.

The new method was then used to calculate bandwidth estimates for different signal frequencies, different temporal arrangements, and with a different stimulus being kept fixed in level. The first clear feature is a general increase in bandwidth with signal frequency, with ERBs increasing from an average 357 Hz at 1 kHz, to an average 1683 Hz at 10 kHz. This is a trend seen in all psychophysical and almost all physiological studies of frequency selectivity in humans (Fletcher, 1940; Zwicker, 1974; Houtgast, 1977; Glasberg and Moore, 2000; Oxenham and Shera, 2003) and animals alike (Pickles, 1975; Ehret and Merzenich, 1985; Ehret and Merzenich, 1988; Syka et al., 2000; Evans, 2001).

A surprising finding is the comparison between forward and simultaneously masked measures. ERBs for each temporal arrangement were measured in two ferrets (Manuel and Leeroy, both male) at 10 kHz and in both cases the simultaneous estimate was narrower. This difference was, however, only significant in one of them (Leeroy). This is in direct contrast to the findings of almost all human psychophysical studies (Houtgast, 1973, 1977; Wightman et al., 1977; Moore, 1978; Vogten, 1978; Moore and Glasberg, 1981; O'Loughlin and Moore, 1981; Glasberg et al., 1984a; Moore et al., 1984; Weber and Patterson, 1984; Moore et al., 1987; Oxenham and Shera, 2003) and some animal studies (Kuhn and Saunders, 1980; Serafin et al., 1982). One animal study, however, found no difference between the two measures (McGee et al., 1976), although this was with a tone masker, not preventing off frequency listening. This strange finding may be a result of not taking into account the growth of forward masking (Moore and Glasberg, 1981). A given increase in masker level does not yield the same increase in masking when using forward masking, e.g. a 1 dB increase in masker level may only raise the detection threshold of a signal by 0.3 dB (Moore, 2003). This is known as the growth of forward masking and is described in more detail in 1.6 Forward versus simultaneous masking. For simultaneous masking this relationship is roughly one-to-one (Hawkins Jr and Stevens, 1950) and such a problem does not arise. However, if the masker level is kept fixed (as in this case) then a 1 dB increase in the amount of masking would not equate to a 1 dB shift in threshold SNR. As a result signal detection thresholds for successive notch widths would appear to be closer together than they truly are, producing broader filter estimates. To make the simultaneous and forward masked estimates comparable it would be necessary to take into account the effect of this growth of forward masking first.

One such growth of forward masking function was calculated for a single ferret (Lex) at 10 kHz. The same signal and no notch masker used for the 10 kHz AF measurement

were used; 10 consecutive 25 ms (12.5 ms on/offset ramps) bursts of 1/32 octave narrowband noise centred on 10 kHz as the signal, each immediately following a 150 ms (5 ms on/offset ramps), 5 kHz wide (centred on 10 kHz) noise burst as masker. Masker level was fixed, and the signal level varied to produce a psychometric function; this process was repeated for several different masker levels. Logistic functions were fit to each psychometric function, and the signal level at 75 % correct taken. This threshold signal level is plotted against masker level in Figure 3.14 panel A. The growth of forward masking is almost perfectly linear, with the line of best fit having a slope of 0.4337. This means that a 1 dB increase in masker level only produces an extra 0.4337 dB of masking. In order to make the forward and simultaneous AF fits comparable to one another, it is first necessary to convert all signal levels to their equivalent masker level (Glasberg and Moore, 1982; Jesteadt et al., 1982; Moore and Glasberg, 1983a; Plack and Oxenham, 1998). When this is done, the difference between the detection threshold of successive notch widths increases and the threshold plot becomes steeper. The resulting narrower bandwidth estimates can be seen in Figure 3.14 panel B and Table 3.3, alongside the other bandwidth measures at the same frequencies. Since the growth of forward masking was measured for Lex at 10 kHz, only this point (black *) has the completely correct conversion factor. However, there is evidence that its slope is frequency invariant (Jesteadt et al., 1982), so the corrected bandwidth measures for other ferrets and frequencies (grey *) are reasonable approximations. Each of these corrected bandwidths are smaller than the uncorrected ones, as expected. Moreover this decrease is noticeable; corrected values are on average 0.51 times the size. In both 10 kHz, forward vs simultaneous masking comparisons this makes the forward estimate narrower than the simultaneous one, agreeing with all the frequency selectivity studies mentioned above.

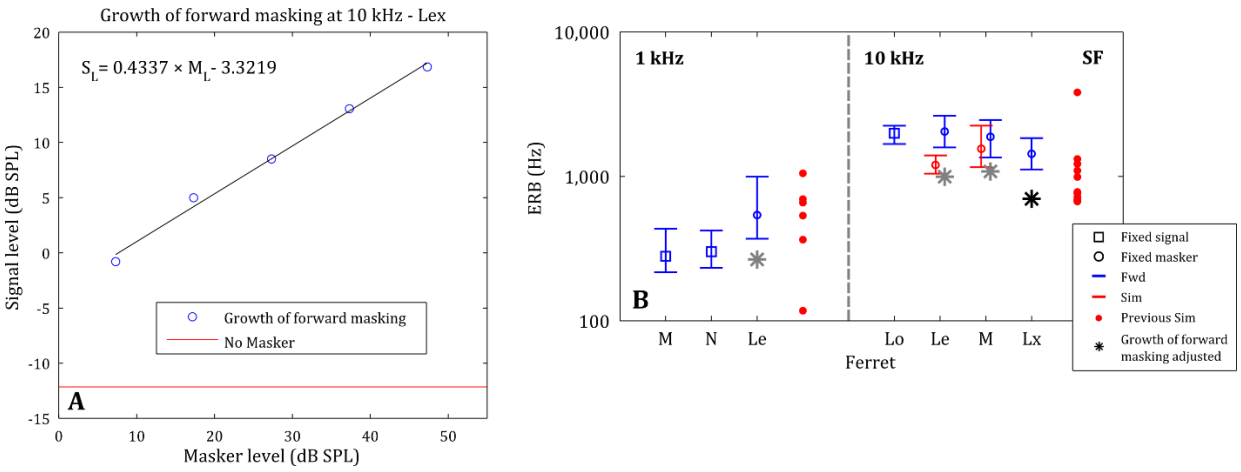


Figure 3.14

A *Growth of forward masking for Lex using a 10 kHz signal and no notch masker of various fixed levels. Each point represents the signal level at threshold against the level of masker used. Inset: equation of line of best fit; S_L = Signal Level (dB SPL), M_L = Masker level (dB SPL).* ***B*** *Bandwidth measures, taken from Figure 3.9, with growth of forward masking correction superimposed (*). The black * uses the growth of forward masking curve for that ferret, at that frequency; the grey ones do not.*

SF (kHz)	Ferret	ERB (Hz)		
		Fwd	Fwd – growth of forward masking corrected	Sim
1	Leeroy	540	266	-
10	Leeroy	2045	995	1199
10	Manuel	1880	1082	1553
10	Lex	1433	699	-

Table 3.3

Bandwidth results for fixed masker level experiments, including correction for growth of forward masking.

A comparison between the signal fixed and masker fixed method, therefore becomes significantly dependent on whether the growth of forward masking correction is used. At 1 kHz the uncorrected, forward masked, fixed masker bandwidth estimate of Leeroy is significantly broader than both fixed signal counterparts of Manuel and Naomi. However, if the correction based on Lex's 10 kHz growth of forward masking is true, then there is no difference between these 1 kHz measurements. The lack of difference matches up closely with the findings of Moore and Glasberg (1981) and Glasberg and Moore (1982), where AF bandwidth was measured at 1 kHz using notched noise, forward masking, compensating for the growth of forward masking. The opposite is true at 10 kHz: with correction, the forward masked fixed masker bandwidths of Leeroy, Manuel and Lex are all smaller than the fixed signal measurement of Lola; without the correction they are indistinguishable. This highlights the importance of taking the growth of forward masking into account. If the 10 kHz, bandwidth estimate of Lola were indicative of all fixed signal level measures at this frequency, then this method gives broader AF estimates than the fixed masker method. This would be the opposite of the findings of Lopez-Poveda and Eustaquio-Martin (2013), which used modelling to investigate the potential difference between the two. They used a model of the auditory periphery, based on the frequency response of the chinchilla BM from Meddis et al. (2001) followed by a temporal window integrator (see 1.8.8 Theories behind psychophysical forward masking) with parameters based on the human study Oxenham (2001). For a signal of ~9.6 kHz (the frequency to which the BM model was tuned), estimates of a known filter size were accurate when using the fixed masker technique. However, measurements made using the fixed signal approach tended to underestimate the true bandwidth by ~ 20 %, deteriorating further at high levels (> 50 dB SPL). How then are the 10 kHz bandwidth estimates in this thesis at least twice as large for the fixed signal as opposed to the fixed masker level? One reason may be that the results of Lopez-Poveda come from a model optimized to a different mammalian species. Also with the ferrets, the signal was fixed ~10 dB above threshold, which is ~13 dB SPL (Kelly et al., 1986), lower levels than were explored in Lopez-Poveda and Eustaquio-Martin (2013). Perhaps bandwidth measurements at this low level (in the linear portion of the BM I/O function (see 1.2 The basilar membrane and its non-linearities)) do not underestimate the true value as much.

Using forward masking and fixed signal level is the current practice in human psychophysics, whereas in animal psychophysics the traditional method is to use continuous simultaneous masking with fixed masker level. Previous estimates of

bandwidth measured behaviourally in the ferret by other researchers at the Institute, using the traditional method, are compared to the results of the novel method used in this thesis in Figure 3.9 and Figure 3.14 (open shapes versus filled circles). Unsurprisingly these results are no different to the simultaneous fixed masker level results of Leeroy and Manuel at 10 kHz (open red circles, vs solid red circles), since the methods are fundamentally identical. More surprisingly though, all of the ERB estimates from the novel method, regardless of which stimulus level is fixed, or temporal arrangement of the signal and masker, fall within the range of previous estimates. At 1 kHz the forward values seem to be on the narrower end of the range covered by previous estimates, whereas the opposite is true at frequencies 3, 7 and 10 kHz. At 10 kHz, in fact, the newer forward masked, fixed signal level estimate is broader than all bar 1 of the previous measurements. Also the relatively large confidence intervals suggest that perhaps the variability in animal behavioural studies, which is much higher than in humans, is so large as to swamp out any benefit of the more up-to-date approach. Although it is very important to compare bandwidth estimates made using the same method, due to systemic non-linearities causing inherent differences in method, perhaps it is not as vital in animal studies.

A comparison was also made between the psychophysical bandwidths of the ferret calculated in this study, and physiological bandwidths measured in the ferret periphery through AN and OAE recordings. The two physiological measures match up exceptionally well with one another, and perfectly with the lowest frequency psychophysical results, 1 kHz. The measurements at this frequency are very similar and their mean and confidence regions sit in the centre of and exactly span the range covered by both physiological measures. This harmony is remarkable since the machinery of the entire auditory system sits between the two measurements; one representing the initial point of neural auditory activity, and the other overall perception. Higher frequencies do not match up quite as closely, with the 3 kHz ERB being ~3 times broader. (As mentioned previously, however, this measure was particularly inaccurate so it should not be given too much credence). At 7 and 10 kHz this difference is smaller, appearing within 50 % of the mean physiological results. One possible interpretation of this, driven mainly by the agreement at 1 kHz, is that frequency selectivity is established at the periphery. This is a contested theory that has studies showing evidence for (Bilsen et al., 1975; Salvi et al., 1982; Evans, 2001) and against (Pickles, 1975; Pickles, 1979; Ehret and Merzenich, 1985). In each of these, physiological recordings have been compared to psychophysical ones and have been

shown to be in agreement or contrast. All the studies just mentioned rejecting this theory, however, used band-widening, which is known to be highly inaccurate. This could easily therefore lead to different bandwidth measurements from physiological and psychophysical data, which would lead to the conclusion that the periphery and perception have different tuning properties. The body of evidence for the peripheral establishment of frequency selectivity is therefore growing.

The comparison between roex and GTFs also showed a reassuring consistency, suggesting that the choice of filter does not have a major effect on results. In some cases this is only true for GTF orders of 4 (the standard order used in human studies) or higher. Also the use of logistic functions fitted in parallel (a restriction I have not seen used before) does not bias the results, but simply reduces variability and makes fits more stable. Both these findings suggest that the precise details of the analysis carried out at these stages does not carry an inherent bias. This would suggest that results from other studies using slightly different analysis methods would still be comparable to those in this study.

CHAPTER 4: Summary and future work

The two main aims of this study were to 1) measure frequency selectivity in the IC and A1 of the guinea pig, using an up-to-date human psychophysical method, and 2) develop a behavioural task that would allow the same method to be used for behavioural data in the ferret. The idea is to unify psychophysical and physiological frequency research, by developing one method for both. Due to the non-linear nature of the auditory system this is an important step in comparing results from both sub-fields, with the aim of finding the neural basis of frequency selectivity. So what has this study achieved?

A current human psychophysical method for measuring frequency selectivity is notched noise forward masking with fixed signal level, as used in Oxenham and Shera (2003). This was applied to guinea pig IC cells and was found to match up very closely with results from previous studies in more peripheral guinea pig auditory brain regions (Evans, 2001; Sayles and Winter, 2010), as well as guinea pig psychophysics (Evans et al., 1992). This suggests that cochlear tuning is maintained up to the level of the IC, if not further. Remarkably the methods used to measure bandwidth in studies in the more peripheral brain regions are markedly different to the NN method; they use excitatory pure tone responses, or FRA. In order to test methodological differences, I used the FRA method to estimate bandwidth in the same units in which NN estimates were made; FRA bandwidths were roughly 3 times broader than NN ones. This clear difference between the two methods in the same neurons was also seen in A1, suggesting that the two methods are not measuring the same feature of frequency selectivity. Perhaps cochlear tuning is maintained throughout the auditory system through inhibitory/masking activity, and FRA show only excitatory input. This sums activity over a wider region of the cochlea as one moves to higher brain regions, giving the impression of broader AFs.

Bandwidth estimates using NN in A1 were found to be much narrower for high frequencies than in the IC, and also narrower than psychophysical estimates in the

guinea pig. This may be due to the highly non-linear nature of A1 neurons violating one of the principal assumptions of the power spectrum model, despite a check of neurometric functions finding no evidence for this. Other studies have found finer tuning in A1 than measured psychophysically, so the finding is not unprecedented (Bitterman et al., 2008; Bartlett et al., 2011).

Also an SDT approach (used extensively in psychophysics) was used with the NN bandwidth measurements, instead of the traditional method of using only mean firing rates in response to the signal. SDT allows other features of neural activity, such as variance and the effect of masking activity in the absence of a signal, to factor into the detection of the signal. An investigation of the relationship between detectability measured using SDT and mean firing rates showed that there was no difference, suggesting that none of the extra available neural information is used. This was true for both IC and A1 neurons, showing that using mean firing rates is comparable to psychophysical methods of signal detection, at least in both these brain regions.

The relationship between CF and ERB for the guinea pig data from this thesis are plotted together with AN results from the ferret, in Figure 4.1 panel A. The FRA bandwidths of the IC and A1 are so similar, that the mean of the two has been plotted. For the same reason the mean of the forward and simultaneously masked NN results for IC and A1 are plotted. The difference between the FRA and NN methods are very clear in this plot. Also note the similarity between bandwidths of the ferret AN and the NN bandwidths in IC (which in turn match up closely with the guinea pig AN bandwidths of Evans (2001)). Despite the two having very different evolutionary pressures, i.e. predator versus prey, cochlear frequency selectivity seems to be similar for both species. Frequency selectivity comparisons between these species are therefore not unwarranted.

A method was also developed to allow both forward and simultaneous notched noise masked bandwidths to be measured behaviourally in the ferret. It relies on the signal and masker being played in bursts, with the temporal relationship between the two determining the type of masking. The method worked well. Ferrets took to the task and easily ignored the masker without losing motivation. Measurements showed that it was possible to achieve forward masking as well as simultaneous, with little noticeable audible difference between the two, meaning a ferret does not need to be retrained to switch between them. Bandwidth estimates were made for a number of different signal frequencies, in both forward and simultaneous arrangements, with some having the masker level fixed and others the level of the signal fixed. Results for these different

methods were similar, in contrast to human studies in which simultaneously masked bandwidths are broader than forward ones (Moore et al., 1984), and in which it has been suggested that varying the masker level leads to overly narrow bandwidth estimates (Lopez-Poveda and Eustaquio-Martin, 2013). Perhaps the variability in animal behavioural data is larger in comparison to similar human data, so that any underlying differences caused by method are hidden.

Also psychophysically measured bandwidths were very similar, across frequency, to physiological bandwidths measured at the periphery in the ferret; in the AN (Sumner and Palmer, 2012) and cochlea, the latter using SFOAEs (unpublished). This suggests that frequency selectivity is established at the periphery and maintained throughout the auditory system.

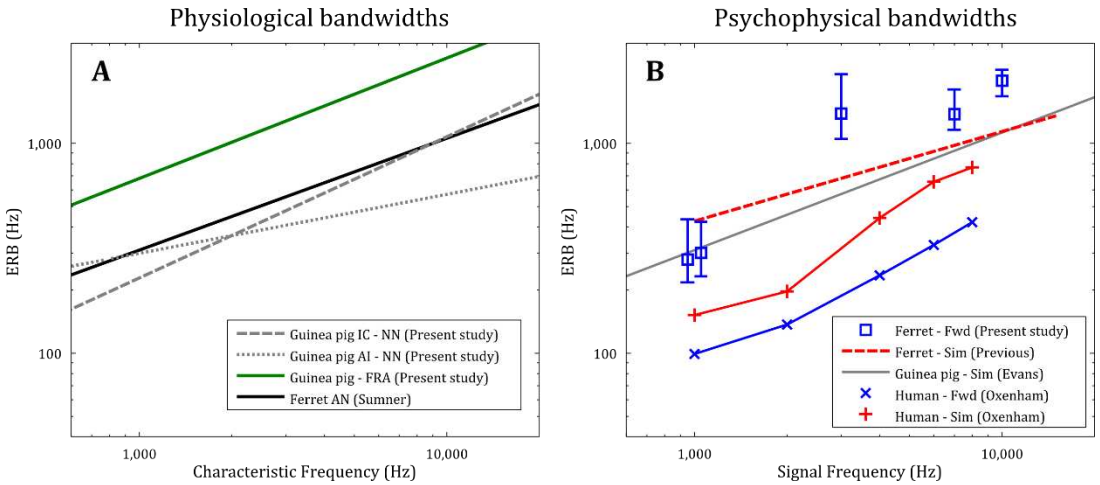


Figure 4.1

A *Relationship between CF and ERB for various physiological studies; guinea pig IC and A1 NN from this thesis, guinea pig IC and A1 average using FRA from this thesis, ferret AN using FRA from Sumner and Palmer (2012).* ***B*** *Relationship between signal frequency and ERB for various notched noise psychophysical studies; ferret forward masked fixed signal level from this thesis (see Figure 3.11), ferret simultaneously masked fixed masker level from other researchers at the Institute Sollini and Alves-Pinto (unpublished), guinea pig simultaneously masked fixed masker level from Evans et al. (1992), human psychophysics from Oxenham and Shera (2003) (both forward and simultaneous NN masking with fixed signal level).*

Figure 4.1 panel B shows how the ferret behavioural results (using forward notched noise masking, with fixed signal level; blue squares) compare to psychophysical results in other species. The dashed red line shows the mean relationship between frequency and bandwidth for notched noise experiments in the ferret, using the traditional continuous masker with fixed level (these were collected by other researchers at the Institute; Sollini and Alves-Pinto, unpublished). At 1 kHz the new bandwidths are narrower, but for the other frequencies they are slightly broader. The spread of 'previous' bandwidths, however, completely encompasses the newer estimates, showing there is actually little difference. The grey line shows the results from Evans et al. (1992) for the guinea pig. This too sits in the region of the ferret psychophysics, matching up precisely at 1 kHz. So not only do guinea pig bandwidths match up with those of ferrets in the AN, but also for the entire auditory system through psychophysics. Both species, however, have broader tuning than humans, as measured in Oxenham and Shera (2003). The method used to measure both forward masked estimates (blue points for human and ferret) are identical, yet the ferret bandwidths are 3 times broader than the human ones. The clear difference is unsurprising since physiological studies suggest humans have narrower tuning at the periphery (Shera et al., 2002; Shera et al., 2010b) and in A1 (Bitterman et al., 2008). This may be a result of the development of language forcing very fine frequency selectivity, an evolutionary pressure unparalleled anywhere else in the animal kingdom. Inferring properties of human frequency selectivity from results of the guinea pig and ferret must therefore be done with great care.

The finding that animals have broader AFs than humans has interesting consequences beyond the study of frequency selectivity, for example in auditory scene analysis (Bregman, 1990; Darwin et al., 1995). This is the ability of the auditory system to separate elements of sound into groups or auditory objects. It relies heavily on inter-aural differences in the detection of a sound, to identify its location in space, as well as its frequency characteristics. Poorer frequency selectivity would result in a poorer ability to segregate auditory objects.

Another interesting consequence is to do with the way in which sound is encoded. ANFs, and neurons higher up the auditory pathway (Langner and Schreiner, 1988), have been shown to fire preferentially to a certain phase of a sound, or “phase lock” (Johnson, 1980; Palmer and Russell, 1986). This is a result of the way in which hair cells release neurotransmitter to elicit action potentials. The frequency of a sound can therefore not only be inferred from the “place code” along the BM, but also from the precise inter-spike time intervals, or a “temporal code”. Young and Sachs (1979) recorded from ANF

of the cat in response to synthesised vowel sounds and found that combining firing rate, place and temporal firing information provided a much more stable representation of the vowel sounds than simply using the firing rate of different neurons along the BM. This was especially true for sounds with higher level and the representation of the vowel provided by firing rate alone degenerated even at levels within the conversational range. Temporal coding therefore plays a key role.

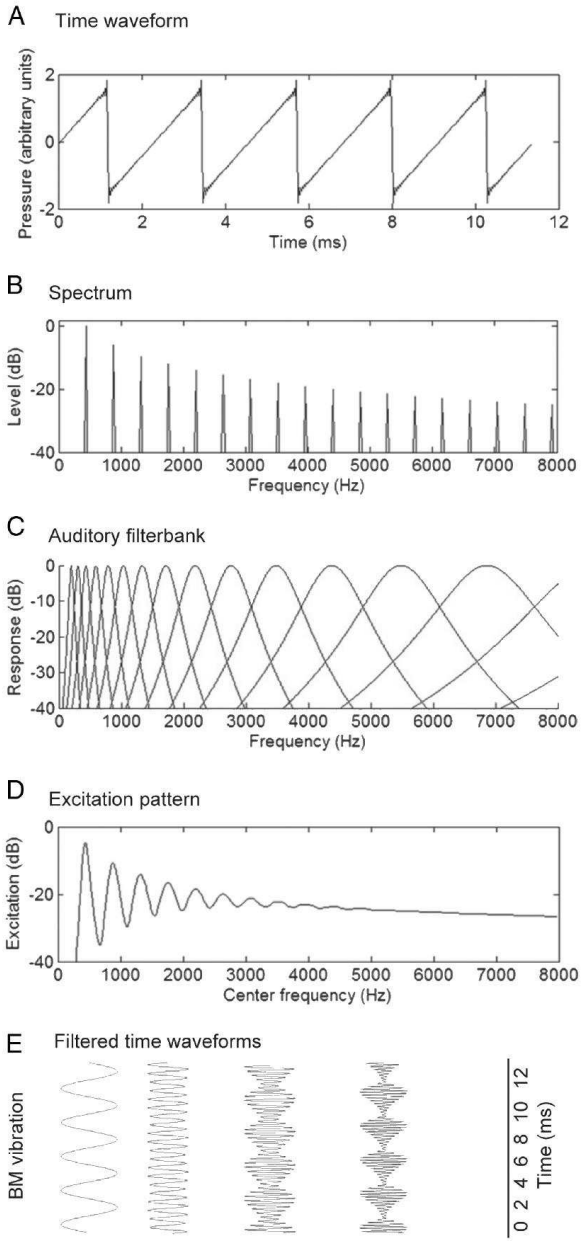


Figure 4.2

Representations of a harmonic complex tone with an f_0 *of 440 Hz (taken from Oxenham (2012)). A Time waveform. B Power spectrum of the same waveform. C AF bank representing cochlear filtering. D Excitation pattern, or the time-averaged output of the AFs. E Sample time waveforms at the output of the filter bank at* f_0 *and the 4th, 8th and 12th harmonics.*

Another auditory phenomenon in which temporal coding is important, and one closely related to frequency selectivity, is the perception of pitch. The ANSI (1994) defined pitch as "that auditory attribute of sound according to which sounds can be ordered on a scale from low-to-high", and is the perceptual correlate of the periodicity of sounds; the way in which their waveform repeats throughout time. In the same way as frequency, the faster the repetition in the waveform the "higher" the perception of pitch. A pure tone will elicit a strong perception of pitch and in such a case, frequency and pitch are directly related. One might therefore expect that poorer frequency selectivity leads to poorer pitch perception. However, complex sounds can also have pitch and the relationship between frequency and pitch becomes much more complicated. A key factor governing the pitch of a sound is its harmonics, and whether they are "resolved" or "unresolved". Most naturally occurring periodic sounds have a fundamental frequency component (which is equivalent to a pure tone of frequency f_0) with several harmonics (flanking pure tones of lower amplitude at integer multiples of f_0). Figure 4.2 panel A shows the waveform of such a sound, with panel B showing its spectrum. If we treat the BM as a sequence of overlapping AFs (Figure 4.2 panel C) we can filter the sound accordingly. Figure 4.2 panel E shows the resulting output of 4 of the AFs (at f_0, the 4^{th}, 8^{th} and 12^{th} harmonics), and panel D shows the "excitation pattern" or the time averaged output of the AFs. For the first two AFs in panel E, only one of the harmonics passes through the filter and the resulting output is a pure tone of fixed amplitude at the centre frequency of the filter; this harmonic is said to be resolved. For the last two AFs multiple harmonics pass through the same filter and the output has a modulated envelope; these harmonics are said to be unresolved. The frequency of these modulations is the same as f_0, and the more harmonics pass through the AF, the more defined this modulation is.

ANFs responding to the output of an AF with unresolved harmonics could phase lock to the modulation rate of the signal, which would be at frequency f_0. This phase locking would be stronger for more pronounced modulations, so the more unresolved the harmonics, the stronger the phase locking. This leads to two possible encoding strategies of pitch; 1) in which spectral cues are used, by observing the peaks of the excitation pattern which will give the location of the resolved harmonics, and 2) in which temporal cues are used by taking the phase locked response of all the filters with unresolved harmonics. Solving the problem of which coding strategy is used has been the source of much research (Loeb et al., 1983; Shackleton and Carlyon, 1994; Shamma

and Klein, 2000; Qin and Oxenham, 2005; Schouten et al., 2005; Schroeder, 2005; Sayles and Winter, 2008; Cedolin and Delgutte, 2010; Carlyon et al., 2012).

Studies looking at the ability of SU in ferret auditory cortical regions show that regardless of decoding strategy, the psychophysical ability of the animal can be completely explained by the combined activity of small ensembles of SU (Walker, 2008; Bizley, 2010). Importantly, temporal information in spike patterns were found to be important for accurate pitch discrimination (Walker, 2008). A recent study by Walker et al. (2014) showed, through a psychophysical pitch discrimination task, that ferrets preferentially use unresolved harmonics in their pitch detection, whereas humans tend to use resolved harmonics. This has also been found in the other species, such as the gerbil (Linge and Klumpe, 2009) and marmoset (Bendor, Osmanski & Wang, 2012).

All this fits together nicely with the findings in this thesis that guinea pigs and ferrets have broader AFs than humans, neurally and behaviourally. Broader filters means more of the harmonics will pass through the same filters making more of them unresolved. Temporal coding of pitch in ferret auditory cortex would therefore be important since most of the information about pitch will be carried by AN phase locking. Broader filters, therefore, do not result in poorer pitch perception, but a shift in coding strategy.

So what future work follows on from this thesis? One of the main findings of the guinea pig physiology experiments is a clear difference between the FRA and NN methods for measuring bandwidth, in both IC and A1. FRA measured bandwidths in AN and VCN, however, match up with IC NN measurements. It would be interesting to see if the disparity seen in IC and A1 is present in either of these early auditory brain areas. This would answer the question of whether the two methods diverge as one moves up the auditory system, or whether they measure different features of frequency selectivity throughout. One of the factors that would severely limit performing an NN experiment in the AN, is time. The AN is an unstable place to record from, and holding a unit for tens of minutes is very difficult. Even in IC collecting a forward and simultaneous NN estimate took almost 2 hours, so using the approach used in this thesis is impractical. If instead of measuring full neurometric functions, finding threshold and using this to fit an AF function, the parameters of the AF were fit directly, experiment duration could be cut dramatically. An adaptive procedure, such as that described in Shen and Richards (2013), could be used to pick stimuli in real time that would be maximally informative for fitting AF parameters. Using this method Shen showed that a stable and accurate roex fit could be obtained in a human 2I2AFC notched noise task in under 150 trials.

This is a fraction of the almost 4000 trials required for a single NN estimate in the experiments in this thesis. Such a reduction in experiment duration could make AN NN measurements of bandwidth possible.

Another benefit of such an adaptive method would be that much more comprehensive FRAs and RLFs, and even tone-on-tone masking plots could be made in the same unit, in the time that is saved (at least in brain areas where recording is more stable). More detailed FRA would lead to more precise measurements of bandwidth using this method. Equally more extensive RLFs could allow the unit to be classified as monotonic or non-monotonic, allowing a further exploration of its linearity. Collecting tone masking data in every unit, using both forward and simultaneous masking, would allow another bandwidth measuring method currently used in physiology to be assessed. It would also allow a more detailed look at the mechanisms of masking in these units, e.g. inhibitory side bands implying inhibitory inputs. This would make a broader study of the effect of method on bandwidth, as well as a more detailed investigation into how non-linearities affect bandwidth measurements in different brain areas. This would determine whether the sharp tuning seen in A1 is true or whether it is an artefact of non-linearity.

A similar adaptive or tracking algorithm might be applied in the ferret behavioural task to reduce the multiple month duration, 1 measurement currently takes. This may, however, be unrealistic since it is already very difficult to make the task easy enough to maintain stimulus control over the ferrets; such an algorithm would target threshold, when the signal is barely audible, and therefore the task would be very difficult. It seems that future work in the ferret behaviour must follow the familiar mantra of 'collect more data'. Currently there are very few points at each frequency using the same method, so it is hard to know what affect variability, in time and subject, has on the results. If several measurements are made at each frequency, using both forward and simultaneous masking, and keeping the signal or masker levels fixed, then more concrete assertions can be made about the effect these methodological differences have on bandwidth measurements.

However, the main future work that can be conducted from the research in this thesis (and the main driving force behind it), is to measure frequency selectivity simultaneously from physiological and psychophysical data, in the same animal. The next step is to implant chronic electrodes into the cortex (and maybe other brain regions in the future) of the ferret, and record from neurons while the animal is performing the behavioural task. This allows bandwidth measurements from both data types to be

made using an up-to-date psychophysical method, with comparisons free from the caveat of species differences (or even subject differences), methodological differences, or anaesthesia. Additionally the relationship between neural activity in a trial and its outcome, i.e. correct/incorrect, could be explored, to see whether neural activity driving signal detection can be identified. Work like this would be invaluable in searching for the neural correlates of frequency selectivity, and get us one step closer to understanding the almost inconceivably extraordinary capabilities of the brain.

References

Alves-Pinto A, Baudoux S, Palmer AR, Sumner CJ (2010) Forward masking estimated by signal detection theory analysis of neuronal responses in primary auditory cortex. Journal of the Association for Research in Otolaryngology 11:477-494.

Alves-Pinto A, Sollini J, Sumner CJ (2012) Signal detection in animal psychoacoustics: Analysis and simulation of sensory and decision-related influences. Neuroscience 220:215-227.

Anderse RA, Roth GL, Aitkin LM, Merzenich MM (1980) The efferent projections of the central nucleus and the pericentral nucleus of the inferior collculus in the cat. Journal of Comparative Neurology 194:649-662.

ANSI (1994) American National Standard Acoustical Terminology, ANSI. vol. S1.1-1994 New York: American National Standards Institute.

Ashmore J (1987) A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. The Journal of physiology 388:323-347.

Astl J, Popelár J, Kvašnák E, Syka J (1996) Comparison of response properties of neurons in the inferior colliculus of guinea pigs under different anesthetics. International Journal of Audiology 35:335-345.

Au WW, Moore PW (1990) Critical ratio and critical bandwidth for the Atlantic bottlenose dolphin. The Journal of the Acoustical Society of America 88:1635.

Auld BA (1973) Acoustic fields and waves in solids: Wiley New York.

Bacon SP (1996) Comments on ‘‘Manipulations of the duration and relative onsets of two-tone forward maskers’’[J. Acoust. Soc. Am.[bold 94], 1269–1274 (1993)]. The Journal of the Acoustical Society of America 99:3246.

Baker RJ, Rosen S (2006) Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. The Journal of the Acoustical Society of America 119:454.

Baker RJ, Rosen S, Darling AM (1998) An efficient characterisation of human auditory filtering across level and frequency that is also physiologically reasonable.

Bar-Yosef O, Rotman Y, Nelken I (2002) Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. The Journal of neuroscience 22:8619-8632.

Bartlett EL, Sadagopan S, Wang X (2011) Fine frequency tuning in monkey auditory cortex and thalamus. Journal of neurophysiology 106:849-859.

Barton EH (1908) Scientific Books: A Text-Book on Sound. Science 28:888-889.

Bechtholt AJ, Gremel CM, Cunningham CL (2004) Handling blocks expression of conditioned place aversion but not conditioned place preference produced by ethanol in mice. Pharmacology Biochemistry and Behavior 79:739-744.

Bendor D, Wang X (2008) Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. Journal of neurophysiology 100:888-906.

Bergevin C, Walsh E, McGee J, Shera C (2012) Probing cochlear tuning and tonotopy in the tiger using otoacoustic emissions. J Comp Physiol A 198:617-624.

Bilsen F, Ten Kate J, Buunen T, Raatgever J (1975) Responses of single units in the cochlear nucleus of the cat to cosine noise. The Journal of the Acoustical Society of America 58:858.

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. Nature 451:197-201.

Bizley JK, Walker KM, King AJ, Schnupp JW (2013a) Spectral timbre perception in ferrets: discrimination of artificial vowels under different listening conditions. The Journal of the Acoustical Society of America 133:365-376.

Bizley JK, Walker KM, Nodal FR, King AJ, Schnupp JW (2013b) Auditory cortex represents both pitch judgments and the corresponding acoustic cues. Current Biology 23:620-625.

Borst A, Theunissen FE (1999) Information theory and neural coding. Nat Neurosci 2:947-957.

Bos CE, de Boer E (1966) Masking and Discrimination. Journal of the Acoustical Society of America 39:708-&.

Bregman A (1990) Auditory scene analysis: the perceptual organization of sound. 1990. MIT Press, Cambridge, MA.

Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. Neuron 26:695-702.

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Visual neuroscience 13:87-100.

Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. Journal of Neurophysiology 77:923-943.

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

Bullock DC, Palmer AR, Rees A (1988) Compact and easy-to-use tungsten-in-glass microelectrode manufacturing workstation. Med Biol Eng Comput 26:669-672.

Burdick CK (1979) The effect of behavioral paradigm on auditory discrimination learning: A literature review. Journal of Auditory Research.

Burkey J, Gans D (1991) Psychophysical tuning curves in gerbils. The Journal of the Acoustical Society of America 89:1822-1823.

Buus S (1985) Release from masking caused by envelope fluctuations. The Journal of the Acoustical Society of America 78:1958.

Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. Nature Reviews Neuroscience 13:407-420.

Carlier E, Pujol R (1982) Sectioning the efferent bundle decreases cochlear frequency selectivity. Neuroscience letters 28:101-106.

Carlyon RP (1988) The development and decline of forward masking. Hearing research 32:65-79.

Carlyon RP, Long CJ, Micheyl C (2012) Across-channel timing differences as a potential code for the frequency of pure tones. Journal of the Association for Research in Otolaryngology 13:159-171.

Casseday J, Covey E (1992) Frequency tuning properties of neurons in the inferior colliculus of an FM bat. Journal of Comparative Neurology 319:34-50.

Cedolin L, Delgutte B (2010) Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve. The Journal of Neuroscience 30:12712-12724.

Chistovich L (1957) Frequency characteristics of masking effect. Biofizika 2:743-755.

Clark W, Stebbins W, Moody D (1975) Behavioral measures of the critical bandwidth in the chinchilla. The Journal of the Acoustical Society of America 58:S124.

Cleveland WS (1981) LOWESS: A program for smoothing scatterplots by robust locally weighted regression. The American Statistician 35:54-54.

Cooper NP, Yates GK (1994) Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibres: Variations with characteristic frequency. Hearing Research 78:221-234.

Costalupes JA, Young ED, Gibson DJ (1984) Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. Journal of neurophysiology 51:1326-1344.

Darwin C, Carlyon R, Moore B (1995) Hearing. Handbook of Perception and Cognition. San Diego, CA: Academic Press Inc.

De Boer E (1967) Correlation studies applied to the frequency resolution of the cochlea. J aud Res 7:209-217.

De Boer E (1969) Encoding of frequency information in the discharge pattern of auditory nerve fibers. International Journal of Audiology 8:547-556.

De Boer E, Jongkees L (1968) On cochlear sharpening and cross-correlation methods. Acta oto-laryngologica 65:97-104.

De Boer E, Kuyper P (1968) Triggered correlation. Biomedical Engineering, IEEE Transactions on 169-179.

De Mare G (1940) Fresh observations as to the so-called masking effect of the ear and its possible diagnostic significance. Acta Oto-Laryngol 28:314-316.

Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. Nature neuroscience 8:1684-1689.

deCharms RC, Merzenich MM (1996) Primary cortical representation of sounds by the coordination of action-potential timing. Nature 381:610-613.

Delgutte B (1989) Physiological mechanisms of masking and intensity discrimination. The Journal of the Acoustical Society of America 85:S14.

Destexhe A, Contreras D, Steriade M (1999) Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. The Journal of neuroscience 19:4595-4608.

DeWeese MR, Wehr M, Zador AM (2003) Binary spiking in auditory cortex. Journal of Neuroscience 23:7940-7949.

Dooling RJ, Saunders JC (1975) Hearing in the parakeet (Melopsittacus undulatus): Absolute thresholds, critical ratios, frequency difference limens, and vocalizations. Journal of Comparative and Physiological Psychology 88:1.

Dooling RJ, Searcy MH (1985) Nonsimultaneous auditory masking in the budgerigar (Melopsittacus undulatus). Journal of Comparative Psychology; Journal of Comparative Psychology 99:226.

Du Verney J (1683) Traité de l'organe de l'ouie, contenant la structure, les usages et les maladies de toutes les parties de l'oreille.

Dubno JR, Dirks DD (1989) Auditory filter characteristics and consonant recognition for hearing-impaired listeners. The Journal of the Acoustical Society of America 85:1666.

Duda RO, Hart PE, Stork DG (2001) Pattern classification: Wiley.

Duifhuis H (1970) Audibility of high harmonics in a periodic pulse. The Journal of the Acoustical Society of America 48:888.

Duifhuis H (1973) Consequences of peripheral frequency selectivity for nonsimultaneous masking. The Journal of the Acoustical Society of America 54:1471.

Eggermont J (1996) How homogeneous is cat primary auditory cortex? Evidence from simultaneous single-unit recordings. Auditory Neuroscience 2:79-+.

Eggermont JJ (1997) Firing rate and firing synchrony distinguish dynamic from steady state sound. Neuroreport 8:2709-2713.

Eggermont JJ, Munguia R, Pienkowski M, Shaw G (2011) Comparison of LFP-based and spike-based spectro-temporal receptive fields and cross-correlation in cat primary auditory cortex. PLoS One 6:e20046.

Egorova M, Akimov A (2013) Spectral coding in auditory midbrain neurons. Journal of integrative neuroscience 12:1-15.

Egorova M, Ehret G (2008) Tonotopy and inhibition in the midbrain inferior colliculus shape spectral resolution of sounds in neural critical bands. European Journal of Neuroscience 28:675-692.

Egorova M, Ehret G, Vartanian I, Esser K-H (2001) Frequency response areas of neurons in the mouse inferior colliculus. I. Threshold and tuning characteristics. Experimental brain research 140:145-161.

Ehret G (1975) Masked auditory thresholds, critical ratios, and scales of the basilar membrane of the housemouse (Mus musculus). Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 103:329-341.

Ehret G, Egorova M, Hage SR, Müller BA (2003) Spatial map of frequency tuning-curve shapes in the mouse inferior colliculus. Neuroreport 14:1365-1369.

Ehret G, Merzenich MM (1985) Auditory midbrain responses parallel spectral integration phenomena. Science 227:1245.

Ehret G, Merzenich MM (1988) Complex sound analysis (frequency resolution, filtering and spectral integration) by single units of the inferior colliculus of the cat. Brain Research Reviews 13:139-163.

Ehret G, Moffat AJ (1985) Inferior colliculus of the house mouse. J Comp Physiol A 156:619-635.

Ehret G, Schreiner CE (1997) Frequency resolution and spectral integration (critical band analysis) in single units of the cat primary auditory cortex. J Comp Physiol A 181:635-650.

Ehret G, Schreiner CE (2000) Regional variations of noise-induced changes in operating range in cat AI. Hearing research 141:107-116.

Elliott L (1971) Backward and forward masking. International Journal of Audiology 10:65-76.

Escabí MA, Read HL (2003) Representation of spectrotemporal sound information in the ascending auditory pathway. Biological cybernetics 89:350-362.

Eustaquio-Martin A, Lopez-Poveda EA (2011) Isoresponse Versus Isoinput Estimates of Cochlear Filter Tuning. JARO-Journal of the Association for Research in Otolaryngology 12:281-299.

Evans E (1972) The frequency response and other properties of single fibres in the guinea-pig cochlear nerve. The Journal of physiology 226:263-287.

Evans E (2001) Latest comparisons between physiological and behavioural frequency selectivity. Physiological and Psychophysical Bases of Auditory Function, edited by J Breebaart, AJM Houtsma, A Kohlrausch, VF Prijs and R Schoonhoven (Shaker, Maastricht).

Evans E, Palmer AR (1980) Relationship between the dynamic range of cochlear nerve fibres and their spontaneous activity. Experimental brain research 40:115-118.

Evans E, Pratt S, Spenner H, Cooper N (1992) Comparisons of physiological and behavioural properties: Auditory frequency selectivity. Auditory Physiology and Perception Pergamon, Oxford 159-170.

Fairhall AL, Lewen GD, Bialek W, van Steveninck RRdR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787-792.

Fastl H (1976) Temporal masking effects: I. Broad band noise masker. Acta Acustica united with Acustica 35:287-302.

Fay RR (1974) Masking of tones by noise for the goldfish (Carassius auratus). Journal of Comparative and Physiological Psychology 87:708.

Fay RR (1988) Hearing in vertebrates: a psychophysics databook: Hill-Fay Associates Winnetka, IL.

Ferrington DG, Downie JW, Willis WD, Jr. (1988) Primate nucleus gracilis neurons: responses to innocuous and noxious stimuli. J Neurophysiol 59:886-907.

Festen J, Plomp R (1983) Relations between auditory functions in impaired hearing. The Journal of the Acoustical Society of America 73:652.

Fidell S, Horonjeff R, Teffeteller S, Green DM (1983) Effective masking bandwidths at low frequencies. The Journal of the Acoustical Society of America 73:628.

Fishman YI, Steinschneider M (2006) Spectral resolution of monkey primary auditory cortex (A1) revealed with two-noise masking. Journal of neurophysiology 96:1105.

Fletcher H (1940) Auditory patterns. Reviews of Modern Physics 12:47.

Fletcher H (1953) Speech and Hearing in Communication: N.Y., Robert E. Krieger Pub.

Fletcher H, Munson WA (1937) Relation between loudness and masking. The Journal of the Acoustical Society of America 9:1-10.

Fonollosa J, Gutierrez-Galvez A, Marco S (2012) Quality coding by neural populations in the early olfactory pathway: analysis using information theory and lessons for artificial olfactory systems. PLoS One 7:e37809.

Fourier JBJ (1822) Théorie analytique de la chaleur: Didot.

Gall MD, Lucas JR (2010) Sex differences in auditory filters of brown-headed cowbirds (Molothrus ater). J Comp Physiol A 196:559-567.

Galván VV, Chen J, Weinberger NM (2001) Long-term frequency tuning of local field potentials in the auditory cortex of the waking guinea pig. JARO-Journal of the Association for Research in Otolaryngology 2:199-215.

Gaucher Q, Edeline J-M, Gourévitch B (2012) How different are the local field potentials and spiking activities? Insights from multi-electrodes arrays. Journal of Physiology-Paris 106:93-103.

Geisler CD, Rhode WS, Kennedy DT (1974) Responses to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey. Journal of neurophysiology.

Gersuni G, Altman J, Maruseva A, Radionova E, Ratnikova G, Vartanian I (1971) Functional classification of neurons in the inferior colliculus of the cat according to their temporal characteristics. Sensory processes at the neuronal and behavioral level Academic Press, New York 157-179.

Gibson DJ, Young ED, Costalupes JA (1985) Similarity of dynamic range adjustment in auditory nerve and cochlear nuclei. Journal of neurophysiology 53:940-958.

Glasberg BR, Moore BC (1982) Auditory filter shapes in forward masking as a function of level. The Journal of the Acoustical Society of America 71:946.

Glasberg BR, Moore BC (1986) Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. J Acoust Soc Am 79:1020-1033.

Glasberg BR, Moore BC, Lutfi RA (1982) Off-frequency listening and masker uncertainty. The Journal of the Acoustical Society of America 72:273.

Glasberg BR, Moore BC, Nimmo-Smith I (1984a) Comparison of auditory filter shapes derived with three different maskers. The Journal of the Acoustical Society of America 75:536.

Glasberg BR, Moore BC, Patterson RD, Nimmo-Smith I (1984b) Dynamic range and asymmetry of the auditory filter. The Journal of the Acoustical Society of America 76:419.

Glasberg BR, Moore BC, Stone MA (1999) Modelling changes in frequency selectivity with level. Psychophysics, Physiology.

Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. Hearing research 47:103-138.

Glasberg BR, Moore BCJ (2000) Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. The Journal of the Acoustical Society of America 108:2318-2328.

Gleich O, Hamann I, Kittel MC, Klump GM, Strutz J (2006) A quantitative analysis of psychometric functions for different auditory tasks in gerbils. Hearing research 220:27-37.

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

Gorga MP, Stelmachowicz PG, Abbas PJ, Small Jr AM (1980) Some observations on simultaneous and nonsimultaneous masking. The Journal of the Acoustical Society of America 67:1821.

Gourevitch G (1965) Auditory masking in the rat. The Journal of the Acoustical Society of America 37:439.

Gourevitch G (1970) Detectability of tones in quiet and in noise by rats and monkeys. Animal psychophysics: The design and conduct of sensory experiments 67-97.

Green DM (1995) Maximum-likelihood procedures and the inattentive observer. The Journal of the Acoustical Society of America 97:3749.

Green DM, McKey MJ, Licklider J (1959) Detection of a pulsed sinusoid in noise as a function of frequency. The Journal of the Acoustical Society of America 31:1446.

Green DM, Swets JA (1966) Signal detection theory and psychophysics: Wiley New York.

Green S (1975) Auditory sensitivity and equal loudness in the squirrel monkey (Saimiri sciureus). Journal of the Experimental Analysis of Behavior 23:255.

Greenwood DD, Maruyama N (1965) Excitatory and inhibitory response areas of auditory neurons in the cochlear nucleus. Journal of neurophysiology 28:863-892.

Greenwood PE, Nikulin MS (1996) A guide to chi-squared testing: Wiley-Interscience.

Guinan Jr JJ (1996) Physiology of olivocochlear efferents. In: The cochlea, pp 435-502: Springer.

Guinan Jr JJ (2006) Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear and hearing 27:589-607.

Halpern DL, Dallos P (1986) Auditory filter shapes in the chinchilla. The Journal of the Acoustical Society of America 80:765.

Harris D, Dallos P, Kraus N (1976) Forward and simultaneous tonal suppression of single-fiber responses in the chinchilla auditory nerve. The Journal of the Acoustical Society of America 60:S81.

Harris DM, Dallos P (1979) Forward masking of auditory nerve fiber responses. Journal of Neurophysiology 42:1083-1107.

Harrison J (1992) Avoiding conflicts between the natural behavior of the animal and the demands of discrimination experiments. The Journal of the Acoustical Society of America 92:1331.

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

Heffner R, Heffner H, Masterton B (1971) Behavioral Measurements of Absolute and Frequency-Difference Thresholds in Guinea Pig. The Journal of the Acoustical Society of America 49:1888.

Heinz MG, Colburn HS, Carney LH (2002) Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates. The Journal of the Acoustical Society of America 111:996.

Helm J, Akgul G, Wollmuth LP (2013) Subgroups of parvalbumin-expressing interneurons in layers 2/3 of the visual cortex. J Neurophysiol 109:1600-1613.

Helmholtz Hv (1877) On the Sensations of Tone, trans. AJ Ellis New York: Dover.

Hen I, Sakov A, Kafkafi N, Golani I, Benjamini Y (2004) The dynamics of spatial behavior: how can robust smoothing techniques help? Journal of neuroscience methods 133:161-172.

Henry KS, Lucas JR (2010a) Auditory sensitivity and the frequency selectivity of auditory filters in the Carolina chickadee, Poecile carolinensis. Animal Behaviour 80:497-507.

Henry KS, Lucas JR (2010b) Habitat-related differences in the frequency selectivity of auditory filters in songbirds. Functional Ecology 24:614-624.

Hernandez O, Espinosa N, Perez-Gonzalez D, Malmierca MS (2005) The inferior colliculus of the rat: a quantitative analysis of monaural frequency response areas. Neuroscience 132:203-217.

Hicks ML, Bacon SP (1999) Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing. The Journal of the Acoustical Society of America 105:326.

Hine JE, Martin RL, Moore DR (1994) Free-field binaural unmasking in ferrets. Behavioral neuroscience 108:196.

Holdsworth J, Nimmo-Smith I, Patterson R, Rice P (1988) Implementing a gammatone filter bank. Annex C of the SVOS Final Report: Part A: The Auditory Filterbank.

Houtgast T (1973) Psychophysical experiments on tuning curves and two-tone inhibition. Acta Acustica united with Acustica 29:168-179.

Houtgast T (1974) Lateral suppression in hearing. Acad Pers BV, Amsterdam.

Houtgast T (1977) Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker. The Journal of the Acoustical Society of America 62:409.

Hudspeth AJ (2008) Making an Effort to Listen: Mechanical Amplification in the Ear. Neuron 59:530-545.

Huffman RF, Henson Jr O (1990) The descending auditory pathway and acousticomotor systems: connections with the inferior colliculus. Brain Research Reviews 15:295-323.

Iversen IH (1994) Reply to Harrison: The representative method. The Behavior Analyst 17:231.

Javel E (1996) Long-term adaptation in cat auditory-nerve fiber responses. The Journal of the Acoustical Society of America 99:1040.

Jesteadt W, Bacon SP, Lehman JR (1982) Forward masking as a function of frequency, masker level, and signal delay. The journal of the Acoustical Society of America 71:950.

Johnson-Davies D, Patterson RD (1979) Psychophysical tuning curves: Restricting the listening band to the signal region. The Journal of the Acoustical Society of America 65:765.

Johnson C, McManus M, Skaar D (1989) Masked tonal hearing thresholds in the beluga whale. The Journal of the Acoustical Society of America 85:2651.

Johnson CS (1971) Auditory masking of one pure tone by another in the bottlenosed porpoise. The Journal of the Acoustical Society of America 49:1317.

Johnson DH (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 68:1115-1122.

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proceedings of the National Academy of Sciences 108:17516-17520.

Kajikawa Y, Schroeder CE (2011) How local is the local field potential? Neuron 72:847-858.

Karikoski JO, Marttila TI, Jauhiainen T (1998) Behavioural observation audiometry in testing young hearing-impaired children. Scandinavian audiology 27:183-187.

Kate Jt, Bilsen F, Raatgever J, Buunen T (1974) Single unit responses in acoustic nuclei of cat to noise and its attenuated repetition. submitted to 8th ICA, London.

Katsuki Y, Sumi T, Uchiyama H, Watanabe T (1958) Electric responses of auditory neurons in cat to sound stimulation. J Neurophysiol 21:569-588.

Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M (2009) Local origin of field potentials in visual cortex. Neuron 61:35-41.

Kawase T, Delgutte B, Liberman MC (1993) Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. Journal of Neurophysiology 70:2533-2549.

Kelly JB, Glenn SL, Beaver CJ (1991) Sound frequency and binaural response properties of single neurons in rat inferior colliculus. Hearing research 56:273-280.

Kelly JB, Kavanagh GL, Dalton JCH (1986) Hearing in the ferret (Mustela putorius): Thresholds for pure tone detection. Hearing research 24:269-275.

Kemp D, Chum R (1980) Observations on the generator mechanism of stimulus frequency acoustic emissions—Two tone suppression. Psychophysical, Physiological and Behavioral Studies in Hearing 5:34-42.

Kiang NY-S (1965) Discharge Patterns of Single Fibers in the Cat's Auditory Nerve. DTIC Document.

Kidd G, Feth LL (1981) Patterns of residual masking. Hearing research 5:49-67.

Kidd Jr G, Feth L (1982) Effects of masker duration in pure-tone forward masking. The Journal of the Acoustical Society of America 72:1384.

King AJ (1998) Auditory system: A neural substrate for frequency selectivity? Current Biology 8:R25-R27.

Koay G, Kearns D, Heffner HE, Heffner RS (1998) Passive sound-localization ability of the big brown bat (Eptesicus fuscus). Hearing research 119:37-48.

Kuhn A, Saunders JC (1980) Psychophysical Tuning Curves in the Parakeet - a Comparison between Simultaneous and Forward Masking Procedures. Journal of the Acoustical Society of America 68:1892-1894.

Laming D, Laming J (1992) F. Hegelmaier: On memory for the length of a line. Psychological research 54:233-239.

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. Journal of neurophysiology 60:1799-1822.

Le Beau F, Rees A, Malmierca MS (1996) Contribution of GABA-and glycine-mediated inhibition to the monaural temporal response properties of neurons in the inferior colliculus. Journal of neurophysiology 75:902-919.

LeBeau FE, Malmierca MS, Rees A (2001) Iontophoresis in vivo demonstrates a key role for GABAA and glycinergic inhibition in shaping frequency response areas in the inferior colliculus of guinea pig. The Journal of neuroscience 21:7303-7312.

Leshowitz B, Lindstrom R (1977) Measurement of nonlinearities in listeners with sensorineural hearing loss. Psychophysics and physiology of hearing 283-292.

Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. The Journal of the Acoustical Society of America 63:442.

Liberman MC (1982) The cochlear frequency map for the cat: Labeling auditory-nerve fibers of known characteristic frequency. The Journal of the Acoustical Society of America 72:1441-1449.

Lilaonitkul W, Guinan JJ (2009) Reflex control of the human inner ear: a half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking. Journal of neurophysiology 101:1394-1406.

Lina IA, Lauer AM (2013) Rapid measurement of auditory filter shape in mice using the auditory brainstem response and notched noise. Hearing research.

Loeb GE, White MW, Merzenich MM (1983) Spatial cross-correlation. Biological cybernetics 47:149-163.

Lopez-Poveda EA, Eustaquio-Martin A (2013) On the Controversy About the Sharpness of Human Cochlear Tuning. JARO-Journal of the Association for Research in Otolaryngology 14:673-686.

Lu Y, Jen PH-S (2001) GABAergic and glycinergic neural inhibition in excitatory frequency tuning of bat inferior collicular neurons. Experimental brain research 141:331-339.

Lüscher E, Zwislocki J (1947) The decay of sensation and the remainder of adaptation after short pure-tone impulses on the ear. Acta Oto-Laryngologica 35:428-445.

Lutfi RA, Patterson RD (1984) On the growth of masking asymmetry with stimulus intensity. The Journal of the Acoustical Society of America 76:739.

Lyon RF, Katsiamis AG, Drakakis E (2010) History and future of auditory filter models. In: Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pp 3809-3812: IEEE.

Macmillan NA, Creelman CD (1991) Detection theory: Cambridge Univ. Press.

Macmillan NA, Creelman CD (2004) Detection theory: A user's guide: Psychology press.

Manley G, Gleich O, Leppelsack H-J, Oeckinghaus H (1985) Activity patterns of cochlear ganglion neurones in the starling. J Comp Physiol A 157:161-181.

Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. The Journal of neuroscience 29:13613-13620.

Margolis RH, Small AM (1975) The measurement of critical masking bands. Journal of Speech, Language and Hearing Research 18:571.

Mayer A (1894) Art. I.-Researches in acoustics. Am J Sci—3rd Series 47:1-28.

Mayer AM (1876) LXI. Researches in acoustics. Philosophical Magazine Series 5 2:500-507.

McGee T, Ryan A, Dallos P (1976) Psychophysical tuning curves of chinchillas. The Journal of the Acoustical Society of America 60:1146.

Meddis R, O'Mard LP, Lopez-Poveda EA (2001) A computational algorithm for computing nonlinear auditory frequency selectivity. The Journal of the Acoustical Society of America 109:2852.

Miller JD (1964) Auditory sensitivity of the chinchilla in quiet and in noise. The Journal of the Acoustical Society of America 36:2010.

Miller JD, Murray FS (1966) Guinea pig's immobility response to sound: Threshold and habituation. Journal of Comparative and Physiological Psychology 61:227.

Miller LM, Escabí MA, Read HL, Schreiner CE (2002) Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. Journal of neurophysiology 87:516-527.

Møller AR (1978) Frequency selectivity of the peripheral auditory analyzer studied using broad band noise. Acta Physiologica Scandinavica 104:24-32.

Moore BC (1978) Psychophysical tuning curves measured in simultaneous and forward masking. J Acoust Soc Am 63:524-532.

Moore BC (2003) An introduction to the psychology of hearing: Academic press San Diego.

Moore BC, Glasberg BR (1983a) Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise. The Journal of the Acoustical Society of America 73:1249.

Moore BC, Glasberg BR (1983b) Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. The Journal of the Acoustical Society of America 74:750.

Moore BC, Glasberg BR, Plack C, Biswas A (1988) The shape of the ear's temporal window. The Journal of the Acoustical Society of America 83:1102.

Moore BCJ (1987) Distribution of Auditory-Filter Bandwidths at 2-Khz in Young Normal Listeners. Journal of the Acoustical Society of America 81:1633-1635.

Moore BCJ (1995) Frequency analysis and masking. Hearing 161-205.

Moore BCJ, Glasberg BR (1981) Auditory filter shapes derived in simultaneous and forward masking. The Journal of the Acoustical Society of America 70:1003.

Moore BCJ, Glasberg BR, Baer T (1997) A model for the prediction of thresholds, loudness, and partial loudness. J Audio Eng Soc 45:224-240.

Moore BCJ, Glasberg BR, Roberts B (1984) Refining the measurement of psychophysical tuning curves. The Journal of the Acoustical Society of America 76:1057.

Moore BCJ, Poon PWF, Bacon SP, Glasberg BR (1987) The temporal course of masking and the auditory filter shape. The Journal of the Acoustical Society of America 81:1873.

Moshitch D, Las L, Ulanovsky N, Bar-Yosef O, Nelken I (2006) Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. Journal of neurophysiology 95:3756-3769.

Munson W, Gardner MB (1950) Loudness patterns—a new approach. The Journal of the Acoustical Society of America 22:177.

Murugasu E, Russell I (1995) Salicylate ototoxicity: the effects on basilar membrane displacement, cochlear microphonics, and neural responses in the basal turn of the guinea pig cochlea. Aud Neurosci 1:139-150.

Nelken I, Chechik G, Mrsic-Flogel TD, King AJ, Schnupp JW (2005) Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. Journal of computational neuroscience 19:199-221.

Nelson D, Swain A (1996) Temporal resolution within the upper accessory excitation of a masker. Acta Acustica united with Acustica 82:328-334.

Nelson DA (1991) High-level psychophysical tuning curves: Forward masking in normal-hearing and hearing-impaired listeners. Journal of Speech, Language and Hearing Research 34:1233.

Nelson PC, Smith ZM, Young ED (2009) Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. The Journal of neuroscience 29:2553-2562.

Nieder A, Klump GM (1999) Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. Hearing research 127:41-54.

Niemiec AJ, Yost WA, Shofner WP (1992) Behavioral measures of frequency selectivity in the chinchilla. The Journal of the Acoustical Society of America 92:2636.

Nienhuys TG, Clark GM (1979) Critical bands following the selective destruction of cochlear inner and outer hair cells. Acta oto-laryngologica 88:350-358.

Noreña A, Eggermont JJ (2002) Comparison between local field potentials and unit cluster activity in primary auditory cortex and anterior auditory field in the cat. Hearing research 166:202-213.

Noreña AJ, Gourévitch B, Pienkowski M, Shaw G, Eggermont JJ (2008) Increasing spectrotemporal sound density reveals an octave-based organization in cat primary auditory cortex. The Journal of neuroscience 28:8885-8896.

O'Loughlin B, Moore B (1981) Off frequency listening: Effects on psychoacoustical tuning curves obtained in simultaneous and forward masking. The Journal of the Acoustical Society of America 69:1119.

Ohlemiller K, Echteler S (1990) Functional correlates of characteristic frequency in single cochlear nerve fibers of the Mongolian gerbil. J Comp Physiol A 167:329-338.

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

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

Oxenham AJ (2001) Forward masking: Adaptation or integration? The Journal of the Acoustical Society of America 109:732.
Oxenham AJ (2012) Pitch perception. The Journal of Neuroscience 32:13335-13338.
Oxenham AJ, Moore BC (1994) Modeling the additivity of nonsimultaneous masking. Hearing research 80:105-118.
Oxenham AJ, Moore BC (1995) Additivity of masking in normally hearing and hearing-impaired subjects. The Journal of the Acoustical Society of America 98:1921.
Oxenham AJ, Moore BC (1997) Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing. Modeling sensorineural hearing loss 273-288.
Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. JARO-Journal of the Association for Research in Otolaryngology 4:541-554.
Palmer A, 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 AR, Jiang D, McAlpine D (2000) Neural responses in the inferior colliculus to binaural masking level differences created by inverting the noise in one ear. Journal of neurophysiology 84:844-852.
Palmer AR, Shackleton TM, Sumner CJ, Zobay O, Rees A (2013) Classification of frequency response areas in the IC reveals continua not discrete classes. The Journal of physiology.
Parker A, Hawken M (1985) Capabilities of monkey cortical cells in spatial-resolution tasks. JOSA A 2:1101-1114.
Patterson RD (1974) Auditory filter shape. The Journal of the Acoustical Society of America 55:802.
Patterson RD (1976) Auditory filter shapes derived with noise stimuli. The Journal of the Acoustical Society of America 59:640.
Patterson RD (1986) Auditory filters and excitation patterns as representations of frequency resolution. Frequency selectivity in hearing 123-177.
Patterson RD, Henning GB (1977) Stimulus variability and auditory filter shape. J Acoust Soc Am 62:649-664.
Patterson RD, Nimmo Smith I (1980) Off frequency Listening and auditory filter Asymmetry. The Journal of the Acoustical Society of America 67:229.
Patterson RD, Nimmo Smith I, Weber DL, 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 RD, Robinson K, Holdsworth J, McKeown D, Zhang C, Allerhand M (1992) Complex sounds and auditory images. Adv Biosci 83:429-446.
Penner M (1975) Persistence and integration: Two consequences of a sliding integrator. Perception & Psychophysics 18:114-120.
Phillips D, Cynader M (1985) Some neural mechanisms in the cat's auditory cortex underlying sensitivity to combined tone and wide-spectrum noise stimuli. Hearing research 18:87-102.
Phillips D, Irvine D (1981) Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. Journal of neurophysiology 45:48-58.
Phillips D, Semple M, Calford M, Kitzes L (1994) Level-dependent representation of stimulus frequency in cat primary auditory cortex. Experimental Brain Research 102:210-226.
Phillips DP (1985) Temporal response features of cat auditory cortex neurons contributing to sensitivity to tones delivered in the presence of continuous noise. Hear Res 19:253-268.

Phillips DP (1990) Neural representation of sound amplitude in the auditory cortex: effects of noise masking. Behavioural brain research 37:197-214.
Pick G, Evans E, Wilson J (1977) Frequency resolution in patients with hearing loss of cochlear origin. Psychophysics and physiology of hearing 273-281.
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 JO (1975) Normal critical bands in the cat. Acta Otolaryngol 80:245-254.
Plack CJ, Moore BC (1990) Temporal window shape as a function of frequency and level. The Journal of the Acoustical Society of America 87:2178.
Plack CJ, Oxenham AJ (1998) Basilar-membrane nonlinearity and the growth of forward masking. The Journal of the Acoustical Society of America 103:1598.
Plomp R (1964) Rate of decay of auditory sensation. The Journal of the Acoustical Society of America 36:277.
Plomp R, Steeneken H (1968) Interference between two simple tones. The Journal of the Acoustical Society of America 43:883.
Popov VV, Supin AY, Klishin VO (1997) Frequency tuning of the dolphin's hearing as revealed by auditory brain-stem response with notch-noise masking. The Journal of the Acoustical Society of America 102:3795.
Prijs VF (1989) Lower boundaries of two-tone suppression regions in the guinea pig. Hearing research 42:73-81.
Prosen CA, Petersen MR, Moody DB, Stebbins WC (1978) Auditory thresholds and kanamycin-induced hearing loss in the guinea pig assessed by a positive reinforcement procedure. The Journal of the Acoustical Society of America 63:559.
Qin MK, Oxenham AJ (2005) Effects of envelope-vocoder processing on F0 discrimination and concurrent-vowel identification. Ear and hearing 26:451-460.
Rabinowitz NC, Willmore BD, King AJ, Schnupp JW (2013) Constructing Noise-Invariant Representations of Sound in the Auditory Pathway. PLoS biology 11:e1001710.
Ramachandran R, Davis KA, May BJ (1999) Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. J Neurophysiol 82:152-163.
Ramachandran R, Davis KA, May BJ (2000) Rate representation of tones in noise in the inferior colliculus of decerebrate cats. Journal of the Association for Research in Otolaryngology 1:144-160.
Rapisarda C, Bacchelli B (1977) The brain of the guinea pig in stereotaxic coordinates. Arch Sci Biol (Bologna) 61:1-37.
Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. Journal of neurophysiology 83:2315-2331.
Rees A, Palmer AR (1988) Rate-intensity functions and their modification by broadband noise for neurons in the guinea pig inferior colliculus. The Journal of the Acoustical Society of America 83:1488.
Relkin EM, Pelli DG (1987) Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures. The Journal of the Acoustical Society of America 82:1679.
Relkin EM, Turner CW (1988) A reexamination of forward masking in the auditory nerve. The Journal of the Acoustical Society of America 84:584.
Rhode WS (1971) Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. The Journal of the Acoustical Society of America 49:1218.
Rhode WS (1973) An investigation of post-mortem cochlear mechanics using the Mössbauer effect. Basic mechanisms in hearing 49-63.

Rhode WS (1978) Some observations on cochlear mechanics. The Journal of the Acoustical Society of America 64:158.
Rhode WS, Robles L (1974) Evidence from Mössbauer experiments for nonlinear vibration in the cochlea. The Journal of the Acoustical Society of America 55:588.
Rife DD, Vanderkooy J (1989) Transfer-function measurement with maximum-length sequences. J Audio Eng Soc 37:419-444.
Robertson D, Cody A, Bredberg G, Johnstone B (1980) Response properties of spiral ganglion neurons in cochleas damaged by direct mechanical trauma. The Journal of the Acoustical Society of America 67:1295.
Rodenburg M, Verschuure J, Brocaar M (1974) Comparison of two masking methods. Acta Acustica united with Acustica 31:99-106.
Rose JE, Greenwood DD, Goldberg JM, Hind JE (1963) Some discharge characteristics of single neurons in the inferior colliculus of the cat. I. Tonotopical organization, relation of spike-counts to tone intensity, and firing patterns of single elements. Journal of neurophysiology 26:294-320.
Rosen S, Baker RJ (1994) Characterising auditory filter nonlinearity. Hear Res 73:231-243.
Rosen S, Baker RJ, Darling A (1998) Auditory filter nonlinearity at 2 kHz in normal hearing listeners. Journal of the Acoustical Society of America 103:2539-2550.
Rosen S, Baker RJ, Kramer S (1992) Characterizing Changes in Auditory Filter Bandwidth as a Function of Level. Adv Biosci 83:171-177.
Ruggero MA (1973) Response to noise of auditory nerve fibers in the squirrel monkey. Journal of neurophysiology 36:569-587.
Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar-membrane responses to tones at the base of the chinchilla cochlea. The Journal of the Acoustical Society of America 101:2151.
Ruggero MA, Robles L, Rich NC (1992) Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. Journal of neurophysiology 68:1087-1099.
Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proceedings of the National Academy of Sciences of the United States of America 102:18614-18619.
Russell I, Nilsen K (1997) The location of the cochlear amplifier: spatial representation of a single tone on the guinea pig basilar membrane. Proceedings of the National Academy of Sciences 94:2660-2664.
Sachs MB, Abbas PJ (1974) Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. The Journal of the Acoustical Society of America 56:1835.
Sachs MB, Kiang NY (1968) Two-tone inhibition in auditory-nerve fibers. J Acoust Soc Am 43:1120-1128.
Saint Marie RL, Shneiderman A, Stanforth DA (1997) Patterns of γ-aminobutyric acid and glycine immunoreactivities reflect structural and functional differences of the cat lateral lemniscal nuclei. Journal of Comparative Neurology 389:264-276.
Sally SL, Kelly JB (1988) Organization of auditory cortex in the albino rat: sound frequency. Journal of neurophysiology 59:1627-1638.
Salvi RJ, Ahroon WA, Perry JW, Gunnarson AD, Henderson D (1982) Comparison of psychophysical and evoked-potential tuning curves in the chinchilla*. American Journal of Otolaryngology 3:408-416.
Saunders J, Bock G, Fahrbach S (1978a) Frequency selectivity in the parakeet (Melopsittacus undulatus) studied with narrow-band noise masking. Sensory processes 2:80.
Saunders JC, Else PV, Bock GR (1978b) Frequency selectivity in the parakeet (Melopsittacus undulatus) studied with psychophysical tuning curves. Journal of Comparative and Physiological Psychology 92:406.

Saunders JC, Rintelmann WF, Bock GR (1979) Frequency selectivity in bird and man: A comparison among critical ratios, critical bands and psychophysical tuning curves. Hearing research 1:303-323.
Sayles M, Winter IM (2008) Reverberation challenges the temporal representation of the pitch of complex sounds. Neuron 58:789-801.
Sayles M, Winter IM (2010) Equivalent-rectangular bandwidth of single units in the anaesthetized guinea-pig ventral cochlear nucleus. Hear Res 262:26-33.
Schafer T, Gales R, Shewmaker C, Thompson P (1950) The frequency selectivity of the ear as determined by masking experiments. The Journal of the Acoustical Society of America 22:490.
Schouten JF, Ritsma R, Cardozo BL (2005) Pitch of the residue. The Journal of the Acoustical Society of America 34:1418-1424.
Schreiner CE (1995) Order and disorder in auditory cortical maps. Curr Opin Neurobiol 5:489-496.
Schreiner CE, Langner G (1997) Laminar fine structure of frequency organization in auditory midbrain. Nature 388:383-386.
Schreiner CE, Mendelson JR (1990) Functional topography of cat primary auditory cortex: distribution of integrated excitation. Journal of neurophysiology 64:1442-1459.
Schreiner CE, Winer JA (2005) The inferior colliculus: Springer.
Schroeder MR (2005) Period Histogram and Product Spectrum: New Methods for Fundamental-Frequency Measurement. The Journal of the Acoustical Society of America 43:829-834.
Seaton WH, Trahiotis C (1975) Comparison of critical ratios and critical bands in the monaural chinchilla. The Journal of the Acoustical Society of America 57:193.
Seki S, Eggermont JJ (2002) Changes in cat primary auditory cortex after minor-to-moderate pure-tone induced hearing loss. Hearing research 173:172-186.
Sellick P, Patuzzi R, Johnstone B (1982) Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. The Journal of the Acoustical Society of America 72:131.
Serafin JV, Moody DB, Stebbins WC (1982) Frequency selectivity of the monkey's auditory system: psychophysical tuning curves. J Acoust Soc Am 71:1513-1518.
Shackleton TM, Carlyon RP (1994) The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. The Journal of the Acoustical Society of America 95:3529-3540.
Shamma S, Klein D (2000) The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. The Journal of the Acoustical Society of America 107:2631-2644.
Shannon RV (1990) Forward masking in patients with cochlear implants. The Journal of the Acoustical Society of America 88:741.
Shen Y, Richards VM (2013) Bayesian adaptive estimation of the auditory filter. The Journal of the Acoustical Society of America 134:1134-1145.
Shera C, Guinan J, Oxenham A (2010a) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. JARO-Journal of the Association for Research in Otolaryngology 11:343-365.
Shera CA, Guinan JJ, Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proceedings of the National Academy of Sciences 99:3318.
Shera CA, Guinan Jr JJ, Oxenham AJ (2010b) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. Journal of the Association for Research in Otolaryngology 11:343-365.
Shower E, Biddulph R (1931) Differential pitch sensitivity of the ear. The Journal of the Acoustical Society of America 3:275-287.

Small Jr AM (1959) Pure-Tone Masking. The Journal of the Acoustical Society of America 31:1619.

Smith D, Moody DB, Stebbins WC, Norat M (1987) Effects of outer hair cell loss on the frequency selectivity of the patas monkey auditory system. Hearing research 29:125-138.

Smith RL (1977) Short-Term Adaptation in Single Auditory Nerve Fibers: Some Poststimu.

Smith RL (1979) Adaptation, saturation, and physiological masking in single auditory-nerve fibers. The Journal of the Acoustical Society of America 65:166.

Smith SW (2003) Digital signal processing: a practical guide for engineers and scientists: Access Online via Elsevier.

Sollini JA (2013) Behavioural and neural correlates of binaural hearing. vol. PhD UK: University of Nottingham.

Stüttgen MC, Schwarz C, Jäkel F (2011) Mapping spikes to sensations. Frontiers in neuroscience 5.

Suga N (1969) Classification of inferior collicular neurones of bats in terms of responses to pure tones, frequency-modulated sounds and noise bursts. The Journal of physiology 200:555-574.

Suga N (1977) Amplitude spectrum representation in the Doppler-shifted-CF processing area of the auditory cortex of the mustache bat. Science 196:64-67.

Sumner CJ, Palmer AR (2012) Auditory nerve fibre responses in the ferret. Eur J Neurosci 36:2428-2439.

Sutter ML (2000) Shapes and level tolerances of frequency tuning curves in primary auditory cortex: quantitative measures and population codes. Journal of neurophysiology 84:1012-1025.

Swets JA, Green DM, Tanner Jr WP (1962) On the width of critical bands. The Journal of the Acoustical Society of America 34:108.

Syka J, Popelář J, Kvašňák E, 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.

Taberner AM, Liberman MC (2005) Response properties of single auditory nerve fibers in the mouse. J Neurophysiol 93:557-569.

Tasaki I (1954) Nerve impulses in individual auditory nerve fibers of guinea pig. J Neurophysiol 17:97-122.

Terhardt E (1974) On the perception of periodic sound fluctuations (roughness). Acta Acustica united with Acustica 30:201-213.

Tolhurst DJ, Movshon JA, Dean A (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision research 23:775-785.

Tsuji J, Liberman M (1997) Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections. Journal of Comparative Neurology 381:188-202.

Turner CW, Relkin EM, Doucet J (1994) Psychophysical and physiological forward masking studies: Probe duration and rise-time effects. The Journal of the Acoustical Society of America 96:795.

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440-10453.

Vater M, Kössl M, Horn AK (1992) GAD-and GABA-immunoreactivity in the ascending auditory pathway of horseshoe and mustached bats. Journal of Comparative Neurology 325:183-206.

Viemeister NF (1988) Intensity coding and the dynamic range problem. Hearing research 34:267-274.

Vogten LL (1978) Low-level pure-tone masking: a comparison of "tuning curves" obtained with simultaneous and forward masking. J Acoust Soc Am 63:1520-1527.

Von Békésy G (1944) Über die mechanische Frequenzanalyse in der Schnecke verschiedener Tiere. Akust Zeits 9:3-11.
Von Békésy G, Wever EG (1960) Experiments in hearing: McGraw-Hill New York.
Walker KM, Schnupp JW, Hart-Schnupp SM, King AJ, Bizley JK (2009) Pitch discrimination by ferrets for simple and complex sounds. The Journal of the Acoustical Society of America 126:1321-1335.
Walker KM, McDermott JH, Kang J, King AJ (2014) Ferret pitch perception is dominated by temporal cues, unlike that of humans. ARO conference abstract
Wallace MN, Rutkowski RG, Palmer AR (2000) Identification and localisation of auditory areas in guinea pig cortex. Exp Brain Res 132:445-456.
Wang J, Salvi RJ, Powers N (1996) Plasticity of response properties of inferior colliculus neurons following acute cochlear damage. Journal of neurophysiology 75:171-183.
Watson CS (1963) Masking of tones by noise for the cat. The Journal of the Acoustical Society of America 35:167.
Weber DL (1977) Growth of masking and the auditory filter. The Journal of the Acoustical Society of America 62:424.
Weber DL, Patterson RD (1984) Sinusoidal and noise maskers in simultaneous and forward masking. J Acoust Soc Am 75:925-931.
Webster JC, Miller P, Thompson P, Davenport E (1952) The masking and pitch shifts of pure tones near abrupt changes in a thermal noise spectrum. Journal of the Acoustical Society of America.
Wegel RL, Lane C (1924) The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. Physical Review 23:266.
Wenstrup JJ, Ross LS, Pollak GD (1986) Binaural response organization within a frequency-band representation of the inferior colliculus: implications for sound localization. The Journal of neuroscience 6:962-973.
Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63:1293-1313.
Widin GP, Viemeister NF (1979) Intensive and temporal effects in pure-tone forward masking. The Journal of the Acoustical Society of America 66:388.
Wightman F, McGee T, Kramer M (1977) Factors influencing frequency selectivity in normal and hearing-impaired listeners. Psychophysics and physiology of hearing 295-306.
Willott JF (2001) Handbook of mouse auditory research: from behavior to molecular biology: CRC Press.
Willott JF, Kulig J, Satterfield T (1984) The acoustic startle response in DBA/2 and C57BL/6 mice: relationship to auditory neuronal response properties and hearing impairment. Hearing research 16:161-167.
Wilson J, Evans E (1971) Grating acuity of the ear: psychophysical and neurophysiological measures of frequency resolving power. In: Proc 7th Int Congr on Acoustics, vol. 3, pp 397-400.
Winer JA, Larue DT, Diehl JJ, Hefti BJ (1998) Auditory cortical projections to the cat inferior colliculus. Journal of Comparative Neurology 400:147-174.
Winslow RL, Sachs MB (1988) Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. Hearing research 35:165-189.
Yost WA, Shofner WP (2009) Critical bands and critical ratios in animal psychoacoustics: An example using chinchilla data. The Journal of the Acoustical Society of America 125:315.
Young ED, Barta PE (1986) Rate responses of auditory nerve fibers to tones in noise near masked threshold. The Journal of the Acoustical Society of America 79:426.

Young ED, Sachs MB (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. The Journal of the Acoustical Society of America 66:1381-1403.

Zwicker E (1954) Die Verdeckung von Schmalbandgeräuschen durch Sinustöne. Acustica 4:415-420.

Zwicker E (1974) On a psychoacoustical equivalent of tuning curves. In: Facts and models in hearing, pp 132-141: Springer.

Zwicker E (1984) Dependence of post-masking on masker duration and its relation to temporal effects in loudness. The Journal of the Acoustical Society of America 75:219.

Zwicker E, Flottorp G, Stevens SS (1957) Critical band width in loudness summation. The Journal of the Acoustical Society of America 29:548.

Zwislocki J, Pirodda E, Rubin H (1959) On some poststimulatory effects at the threshold of audibility. The Journal of the Acoustical Society of America 31:9.