INVESTIGATIONS OF THE EFFECTS OF SEQUENTIAL TONES ON THE RESPONSES OF NEURONS IN THE GUINEA PIG PRIMARY AUDITORY CORTEX

Chris D. Scholes, MNeuro

Thesis submitted to the University of Nottingham for

the degree of Doctor of Philosophy

March 2009

Abstract

The auditory system needs to be able to analyse complex acoustic waveforms. Many ecologically relevant sounds, for example speech and animal calls, vary over time. This thesis investigates how the auditory system processes sounds that occur sequentially. The focus is on how the responses of neurons in the primary auditory cortex 'adapt' when there are two or more tones.

When two sounds are presented in quick succession, the neural response to the second sound can decrease relative to when it is presented alone. Previous two-tone experiments have not determined whether the frequency tuning of cortical suppression was determined by the receptive field of the neuron or the exact relationship between the frequencies of the two tones. In the first experiment, it is shown that forward suppression does depend on the relationship between the two tones. This confirmed that cortical forward suppression is 'frequency specific' at the shortest possible timescale.

Sequences of interleaved tones with two different frequencies have been used to investigate the perceptual grouping of sequential sounds. A neural correlate of this auditory streaming has been demonstrated in awake monkeys, birds and bats. The second experiment investigates the responses of neurons in the primary auditory cortex of anaesthetised guinea pigs to alternating tone sequences. The responses are generally consistent with

awake recordings, though adaptation was more rapid and at fast rates, responses were often poorly synchronised to the tones.

In the third experiment, the way in which responses to tone sequences build up is investigated by varying the number of tones that are presented before a probe tone. The suppression that is observed is again strongest when the frequency of the two tones is similar. However, the frequencies to which a neuron preferentially responds remain irrespective of the frequency and number of preceding tones. This implies that through frequency specific adaptation neurons become more selective to their preferred stimuli in the presence of a preceding stimulus.

Declaration

This thesis is the candidate's own work, and has not, in the same or different form been submitted to this or any other University for a degree. The majority of neurophysiological experiments were performed by this candidate, though some were performed by Dr Christian J. Sumner or Dr Sylvie Baudoux. All data analysis was performed by this candidate.

Acknowledgements

Firstly, I would like to thank Chris Sumner for all the hours of supervision that he has bestowed upon me and for being a good mate to have around over the past few years. Thanks also to Mark Wallace and Trevor Shackleton for invaluable help with the technical side of things. Also, thanks to Secret Santa for the Juventus shirt, I doubt anyone will top that for less than a fiver. Cheers go to Alan Palmer for general advice on all things written, spoken and operated on, and for allowing me to run rings around him at football (insert winky smiley). Thanks to Jon Peirce for invaluable (though mostly forgotten) advice in the pub. Also, thanks to Deb Hall for good advice during the PhD and for being an understanding landlady.

Big up to the glue that keeps the IHR together. To the office staff who keep things ticking over and provide us with buns. To Kathie, Rachel and Zoe for keeping the guinea pigs happy. To the cleaners who are always there for a down-to-earth chat when you're the only person left in the building attempting to finish a thesis.

A last, but by no means least, shout out to all those people who have helped me to maintain a sunny outlook on life over the past few years. To distant friends Jazz and Calum, Mike, Dougie, Si, Sam and George. To Joe, Kerri, Nick, Jess, Damon, Eve, Jonathon and Darren for being in the Jono's at the same time as me. Finally, thanks to Angie who was always there for a ciggy and a chat.

Abbreviations

A1	=	Primary auditory cortex
AN	=	Auditory nerve
AVCN	=	Anteroventral cochlear nucleus
BF	=	Best frequency
BM	=	Basilar membrane
BSF	=	Best suppressed frequency
CF	=	Characteristic frequency
CFI	=	Conditioner following index
Ch	=	Chopper
CN	=	Cochlear nucleus
CNIC	=	Central nucleus of the inferior colliculus
CSF	=	Characteristic suppressed frequency
DCIC	=	Dorsal cortex of the inferior colliculus
DCN	=	Dorsal cochlear nucleus
EAM	=	External auditory meatus
EEG	=	Electroencephalography
ERP	=	Event related potential

FD	=	Frequency difference
FFT	=	Fast Fourier transform
*f*MRI	=	*functional* Magnetic resonance imaging
FR	=	Frequency response
FSR	=	Frequency specificity ratio
FTC	=	Frequency threshold curve
GABA	=	Gamma-aminobutyric acid
IC	=	Inferior colliculus
IHC	=	Inner hair cell
ISI	=	Inter-stimulus interval
LLR	=	Long latency response
LSO	=	Lateral superior olive
MEG	=	Magnetoencephalography
MGB	=	Medial geniculate body
MLR	=	Mid-latency response
MMN	=	Mismatch negativity
MSO	=	Medial superior olive
MTB	=	Medial nucleus of the trapezoid body

MU	=	Multi-unit
NESD	=	No-tone equivalent spike count distribution
OHC	=	Outer hair cell
On	=	Onset
PB	=	Pause build-up
PBC	=	Preferred binaural combinations
PCA	=	Principal components analysis
PFI	=	Probe following index
PL	=	Primary-like
PR	=	Presentation rate
PSTH	=	Peri-stimulus time histogram
PTC	=	Psychophysical tuning curve
PVCN	=	Posteroventral cochlear nucleus
RF	=	Receptive field
SFTC	=	Suppressed frequency threshold curve
SOA	=	Stimulus onset asynchrony
SOC	=	Superior olivary complex
SRF	=	Suppressed receptive field

SSA	=	Stimulus-specific adaptation
STG	=	Superior temporal gyrus
SU	=	Single unit
VCN	=	Ventral cochlear nucleus

Table of Contents

1. Introduction

This thesis aims to investigate how the auditory system processes sounds that occur sequentially over different time scales. Specifically, how the responses of neurons in the primary auditory cortex change when sounds are not presented alone but are presented in the context of other sounds.

The auditory system needs to be able to analyse complex acoustic waveforms that include elements from a variety of sources. Many ecologically relevant sounds, for example speech and animal vocal calls, vary in a specific manner over time. Furthermore, sounds rarely occur in isolation but arrive at the ears as a mixture and the auditory system seems well suited to the analysis of such sounds (Moore, 2003).

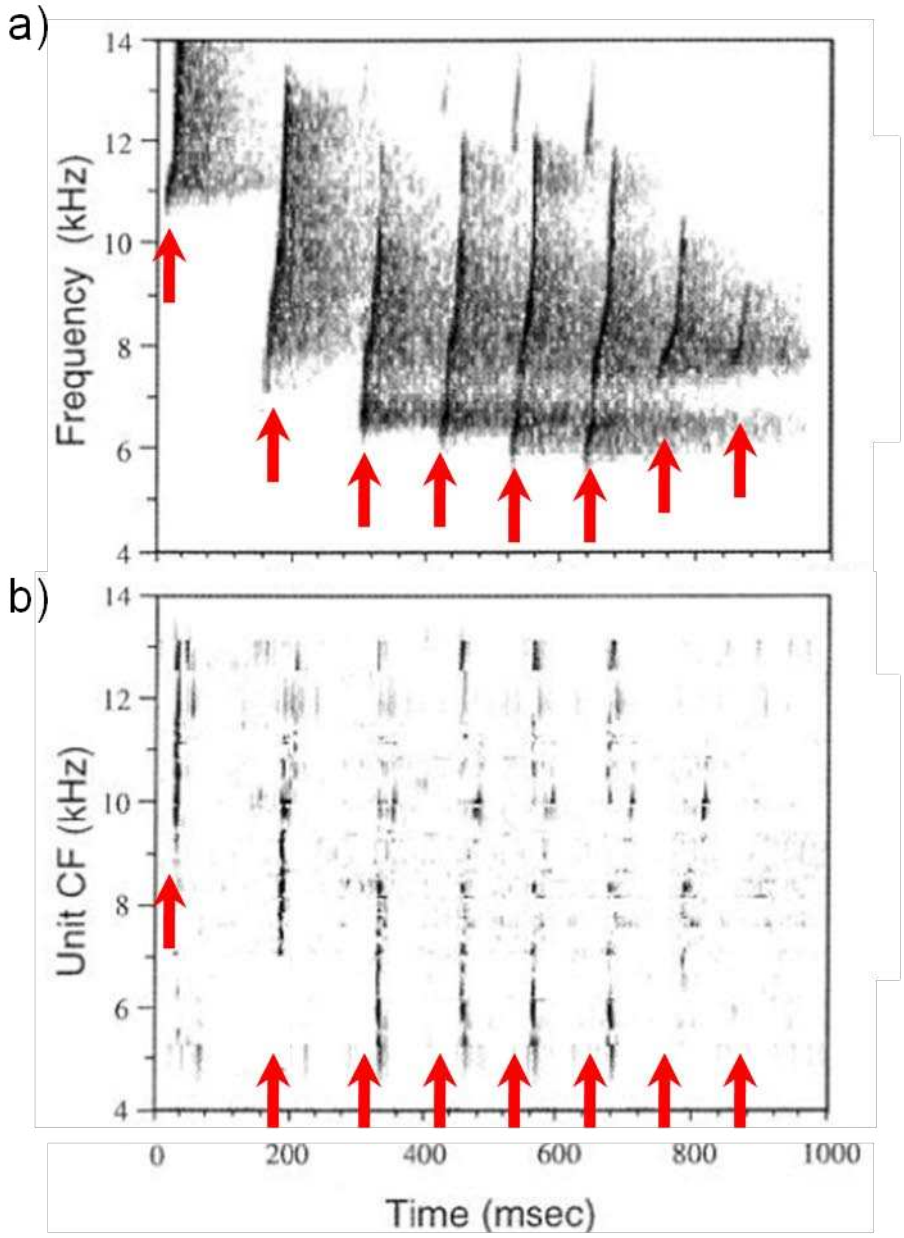


Figure 1-1. The neural representation of a twitter call in the marmoset auditory cortex. a) The spectrogram for a male marmoset twitter call (explanation in text). b) The neural population responses to the male marmoset's twitter call. Spikes as they occurred in time from individual cortical units are ordered according to their characteristic frequency (CF). The grey level of the plot is proportional to the number of spikes. The time scales for a) and b) have been aligned to allow direct comparison (adapted from Wang et al., 1995).

An example of how the frequency components of a sound change over time is displayed in figure 1-1a, which shows the spectrogram for the twitter

call of a male marmoset. A spectrogram shows the power of different frequency components of a sound as a function of time. The spectrogram for the twitter call is characterised by several phrases (the onset of each is highlighted by a red arrow) that consist of an upward sweep across frequency. In addition, there is a decrease in the frequency at which each sweep begins. The auditory system is suited to the analysis of such signals partly because it breaks down the input signal into different frequency components. If the twitter call is presented to a marmoset and responses of neurons in the auditory cortex (one of the regions of the brain that processes sound) are displayed then the neural representation is very similar to the spectrogram (figure 1-1b). Neurons in the auditory system each respond to a limited range of sound frequencies and there is usually one particular frequency to which they are most sensitive: the characteristic frequency (CF). In figure 1-1b, the auditory neurons have been lined up by their CF and the darkness of the grey lines denotes how strong the response was. The red arrows in figure 1-1b are aligned to those in figure 1-1a. After a small temporal gap the neural populations with CFs that are at the frequency of each sweep phrase respond vigorously to the twitter call. In addition, the neural responses follow the decrease in frequency with each phrase of the twitter call.

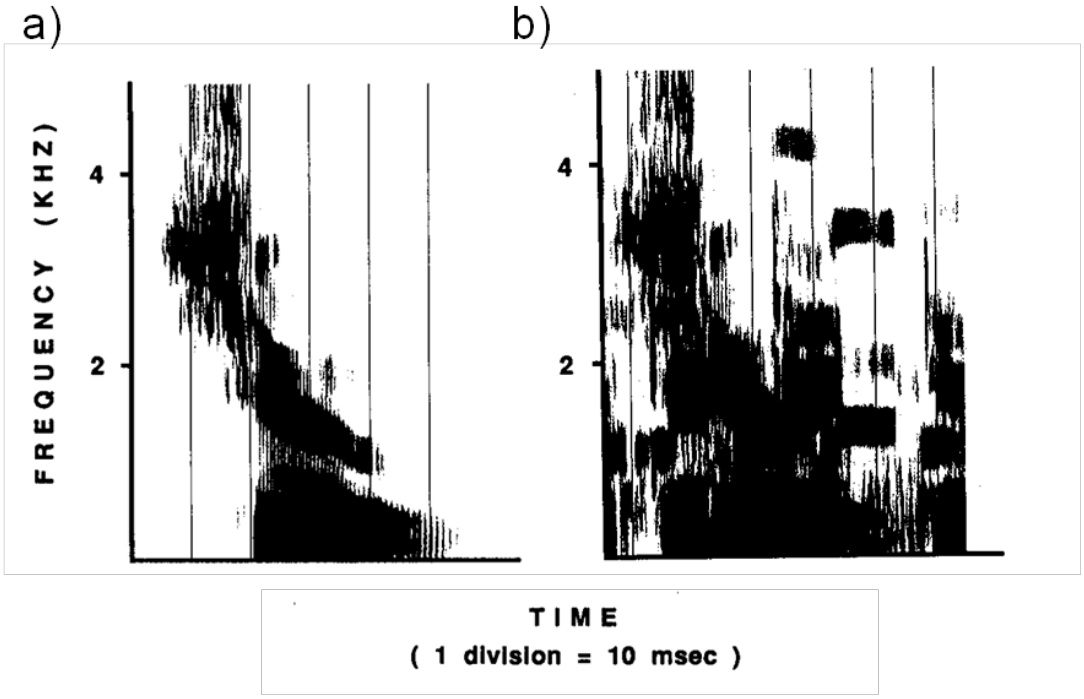


Figure 1-2. Spectrogram for the word 'shoe' spoke a) in isolation and b) in a mixture of sounds (reprinted from Bregman, 1990 fig. 1.3 and 1.4).

Human speech sounds display a similar complex relationship of frequency and time, as shown in the spectrogram representation of the word ‘shoe’ in figure 1-2a. The spectrogram initially shows frequency components at and above 3 kHz that correspond to the high frequency fricative ‘sh’ sound. Over time there is a descending sweep in the frequency to the lower frequency vowel sound. One can imagine that the responses of neurons in the human brain would show a similar pattern of activity to those in the marmoset brain. In trying to establish how the auditory system processes these sounds, one could imagine that a categorical system, or one which learns different categories of sound, could successfully partition sound input without the need for any other special mechanisms. A problem appears, however, if one considers that the frequency components for a given sound

rarely occur alone. The auditory system is instead faced with a complex input, such as that displayed in figure 1-2b, consisting of a mixture of sounds from different sources.

The method by which the auditory system deals with this task has been termed auditory scene analysis (Bregman, 1990). In essence, auditory scene analysis refers to the way in which the complex acoustic waveform is separated into different acoustic objects based on Gestalt principles such as proximity. For example, acoustic elements of similar frequencies might be grouped together because they are likely to come from the same sound source. In a similar way, acoustic elements that repeat over time and are of a similar frequency will be grouped together into 'auditory streams'. Characterising the formation of auditory streams is important in understanding how the auditory system processes time-varying signals, for example speech or species-specific vocalisations, in environments that are intrinsically noisy. Understanding speech in a noisy environment is, incidentally, a major problem faced by people with hearing problems (Plomp and Mimpen, 1979; Gatehouse and Noble, 2004).

A simple way of investigating how the auditory system groups sound inputs is to use pure tones, which are sinusoidal waveforms. While it is clear that tonal stimuli do not reflect natural acoustic stimuli very well, there are several advantages to using such simple stimuli. Firstly, any waveform can be uniquely described as a set of sinusoids, and the transformations of signals that result from the processing by a linear system can be completely

described as changes in the phase and amplitude of those sinusoids (Moore, 2003). The auditory system is often conceived as being made up of several stages with each stage considered as a device with an input and an output. If the device is linear then changing the magnitude of the input to the device by a factor of *k* should result in the output magnitude changing by a factor of *k*, but with the output otherwise remaining the same as the input. Additionally, the output of the device in the response to a number of simultaneously presented inputs should be the same as the sum of the outputs that would have been obtained if the inputs were presented separately. Indeed, the spectrogram representations in figure 1-1 and figure 1-2 are decompositions of those complex sounds to individual pure tones that change in amplitude over time. Although auditory neural responses are not linear, they sometimes behave as if the device driving them were linear (Moore, 2003). For this reason, the pure tone has been a popular tool for investigating the operations of the auditory system and so both the perception of and neural responses to tones have been extensively characterised. A second advantage is that tones can be characterised using a small number of parameters, for example frequency, level and duration. By varying one of these parameters investigators can be sure that any change in either perception or neural responses is due to the parameter change. With more complicated stimuli, for example speech or animal vocalisations (Wang et al., 1995), it is more difficult to infer the cause of any recorded changes because parametric control of the stimulus is complicated (Theunissen and Doupe, 1998). Thus, in this thesis

pure tones will be used to study how the auditory system processes sequential sounds.

To fully understand the way in which the auditory system performs the task of auditory scene analysis, it can be argued that it is necessary to understand how individual neurons process sound, and this approach is taken in this thesis. The nature of the neural processing of sound has been investigated at all levels of the auditory system. However this thesis focusses on the representation of sound at the level of the primary auditory cortex. The primary areas of the auditory cortex represent the last region of the brain that is clearly tonotopically organised (Read et al., 2002), that is the neurons are spatially arranged such that there is a regular gradient in the CF of neurons along one axis. This systematic arrangement of frequency reflects the manner in which sound is transduced at the first processing level of the auditory system (the cochlea) and this organization is maintained at all levels of the ascending auditory system up to the primary auditory cortex (Merzenich et al., 1977). The importance of the auditory cortex in the processing of temporal stimulus sequences has been demonstrated with behaviour-lesion studies in both cats and monkeys. Following ablation of parts of the temporal lobes (on which the auditory cortex is positioned) subjects were significantly impaired in discriminating changes in the temporal patterns of tone sequences (Diamond and Neff, 1957; Diamond et al., 1962; Kaas et al., 1967; Strominger et al., 1980), tone intensities (Oesterreich et al., 1971) and tone durations (Scharlock et al., 1965) .

Psychophysically, the processing of tones has been investigated extensively across many time scales (van Noorden, 1975; Jesteadt et al., 1982; Bregman, 1990). For example, interesting properties of perceptual organisation can be examined simply by linking individual tones together into sequences. A sequence of non-overlapping tones with two different frequencies (frequency 'A' and frequency 'B' in the sequence ABABABAB) has been used extensively to investigate how the auditory system forms separate streams. If the frequency difference between the 'A' and 'B' tone is small then the sequence is perceived as a single stream that moves up and down in frequency. However, if the frequency difference is increased then the 'A' and 'B' tones perceptually split to form two streams occurring at half the rate of the original sequence, only one of which can be attended to at any moment in time (Miller and Heise, 1950; van Noorden, 1975; Bregman, 1990). The ease with which the perception of this tone sequence can be changed by varying one parameter makes it a powerful tool for investigating auditory streaming. Another simple stimulus that has been used to investigate both the temporal and frequency resolution of the auditory system consists of two sequential tones. Psychophysically, previous tones tend to make any subsequent tones harder to hear in a process known as forward masking (Moore, 1978; Jesteadt et al., 1982). By varying the parameters of the two tones it has been possible to infer the way in which the peripheral auditory system breaks input sounds into different frequency components.

Physiological studies lag somewhat behind perceptual studies of tonal processing. The response to single tones has been extensively investigated (Kiang, 1965; Pfeiffer, 1966; Aitkin and Webster, 1972; Phillips and Orman, 1984; Micheyl et al., 2005; Moshitch et al., 2006) and is still the basis on which most basic characterisations of neurons are done (Blackburn and Sachs, 1989; Sutter, 2000). The response to two sequential tones has been investigated at all levels of the auditory system (Smith, 1977; Harris and Dallos, 1979; Boettcher et al., 1990; Calford and Semple, 1995; Shore, 1995; Brosch and Schreiner, 1997; Finlayson, 1999; Malone and Semple, 2001; Zhang et al., 2005). In contrast, the response to sequences of tones has only recently been studied (Fishman et al., 2001; Kanwal et al., 2003; Bee and Klump, 2004; Fishman et al., 2004; Bee and Klump, 2005; Micheyl et al., 2005; Micheyl et al., 2007). Thus there are many aspects of the processing of sequences of tones which remain poorly understood. Physiologically, tones tend to cause a decrease in the response to a subsequent tone (Harris and Dallos, 1979), a process which will be referred to as forward suppression within this thesis. Forward masking and forward suppression share some features, a fact which has led to forward suppression being posited as a neural correlate of forward masking. For example, both forward masking and forward suppression tend to be largest when the two tones have similar frequencies. In some ways this is at odds with the Gestalt idea that sound elements with similar properties should be grouped together. However, forward masking may play more of a role in processes such as echo suppression (Finlayson, 1999).

The usual physiological response to a repetitive stimulus is a gradual decrease or adaptation of the neural response to that stimulus. Physiological investigations using the simple stimuli discussed above have highlighted such adaptation as a possible neural correlate of streaming. This will be discussed in greater detail in the next chapter. Further, the adaptation can be specific to certain features of the stimulus (stimulus-specific adaptation) such that the response to novel stimulus features may be facilitated as a result of adaptation (Ulanovsky et al., 2003; Ulanovsky et al., 2004). This novelty detection mechanism can operate over long periods of time and can act in a general manner on the statistical probability that events will occur. The neural representation of tone sequences is likely governed by an interaction between forward suppression and adaptation so, in essence, there is a link between the neural mechanisms that are operating at all timescales.

This thesis aims to fill in some of the gaps in knowledge of how the auditory system responds to sequences of tones. Specifically, several aspects of cortical responses to tone sequences are investigated. The aim is to explore in what ways the neural mechanisms at different timescales interact and whether it is possible that single or multiple neural mechanisms can explain all of the different features at different timescales. To do this, extracellular recordings were made from anaesthetised guinea pigs to measure the neural responses to the various stimuli discussed. One might expect that phenomena that are governed by the same underlying mechanisms will exhibit similarities when stimuli are altered in certain ways while those that are not would

display differences. This will be the central hypothesis by which the processing of tones at different timescales will be assessed.

In the next chapter I review the body of literature relevant to the processing of tones in the auditory system and elaborate on the features highlighted in this introduction. There are three experimental chapters. In chapter 5, I compare the physiological responses to streaming stimuli in anaesthetised guinea pigs to those that have been reported in awake animals. The results are generally consistent with those observed in awake animals, but there are several differences which may stem from the use of anaesthetic. In chapter 4, I investigate how the characteristics of forward suppression change when the frequency of the second tone in a sequence of two tones is varied. In chapter 6, I go on to investigate how forward suppression builds up when the number of tones that precede another tone is increased. The results of chapters 4 and 6 point to a system whereby in the presence of prior stimulation cortical neurons increase the selectivity of their responses around the stimulus features to which they preferentially respond. So, the frequency to which a neuron preferentially responds is conserved even in the context of prior stimulation that causes forward suppression. In agreement with previous studies (Zhang et al., 2005; Nakamoto et al., 2006; Scholl et al., 2008), stimulus features which the neurons respond weakly to are suppressed by a larger range of stimuli and for longer than those that the neuron responds strongly to. Such a system may contribute to both the responses to streaming stimuli that I have demonstrated in the anaesthetised guinea pig

and the stimulus-specific adaptation of responses over long time scales (Condon and Weinberger, 1991; Ulanovsky et al., 2003; Ulanovsky et al., 2004).

2. Literature review

In this chapter I will review literature relevant to the thesis as well as introducing some important concepts involved with the processing of tones. I start with a brief introduction to sound and then introduce the various processing stations of the auditory system. I then go on to address the literature associated with the two main themes in the thesis. Firstly, I discuss the way in which the temporal resolution of the auditory system has been investigated using sequential two-tone stimuli both psychophysically and physiologically. I move on to discuss how the changes in the neural responses to sequential two-tone stimuli as the auditory system is ascended can also be observed as general changes in the responses to repetitive stimuli. In this section I also introduce the concept of adaptation and, in particular, stimulus-specific adaptation. Secondly, in section 2.3, I move on to consider how the auditory system processes sequences of tones and how such knowledge can be used to infer mechanisms for the creation of auditory streams. In section 2.4 I outline the possibility that the same mechanisms may underlie cortical suppression at both short and long time scales. Finally, section 2.5 is a summary that brings aspects of the literature review together and poses questions that will be addressed in the experimental chapters.

2.1. Introduction to the auditory system

2.1.1. The nature of sound

Sound is essentially a travelling pressure wave accompanied by a wave of displacement of the air molecules which causes them to vibrate around their mean positions. Sound waves can be characterised using their frequency and their amplitude. The frequency is the number of times per second that the waveform is repeated and is expressed in cycles per second or hertz (Hz). The frequency of the sound provides a subjective quality of pitch with higher frequencies having higher pitch. The amplitude is the magnitude of pressure variation about the mean and has a subjective correlate of loudness (Pickles, 1988).

Sound waves can also be characterised using their intensity, which is the sound energy transmitted per second (Moore, 2003). As the auditory system can process sounds over a large range of intensities, a log scale of the ratio between two sound intensities is used. So, it is important to not only know the intensity of a sound but also the reference intensity that the intensity was expressed relative to. The logarithmic unit with which the sound intensity ratio is expressed is the decibel and the function by which the sound level, in decibels, is calculated is:

$$\text{Number of decibels} = 10 \log_{10} (I_1 / I_0) \qquad \textbf{Equation 2-1}$$

where I_1 is the sound intensity and I_0 is the reference intensity. For sounds in air, a reference intensity of 10^{-12} W/m^2 (watts per square meter) is most

commonly used, which is equivalent to a pressure of 20 µPa (micropascal). If a sound level is expressed using this reference then it is referred to as a sound pressure level (SPL). In some cases the reference level is the threshold of an individual listener and in this case the sound level is referred to as sensation level or SL. As the intensity of a sound is proportional to the square of the pressure equation 2-1 can be rewritten as:

$$\text{Number of decibels} = 20 \log_{10} (P_1 / P_0) \qquad \textbf{Equation 2-2}$$

where P_1 is the pressure of the sound and P_0 is the pressure of the reference sound. So a 10-fold increase in pressure is equivalent to a 100-fold increase in intensity and is represented by +20 dB.

2.1.2. Overview of the auditory system

Figure 2-1 shows the ascending pathways of the auditory system. This thesis is mainly concerned with responses in primary auditory cortical neurons though these responses will be shaped by excitatory and inhibitory interactions within the various subcortical nuclei shown. There are also several descending auditory pathways which will undoubtedly shape the responses of neurons at each level of the auditory system. I will generally concentrate on ascending pathways, though where relevant reference will be made to these descending influences.

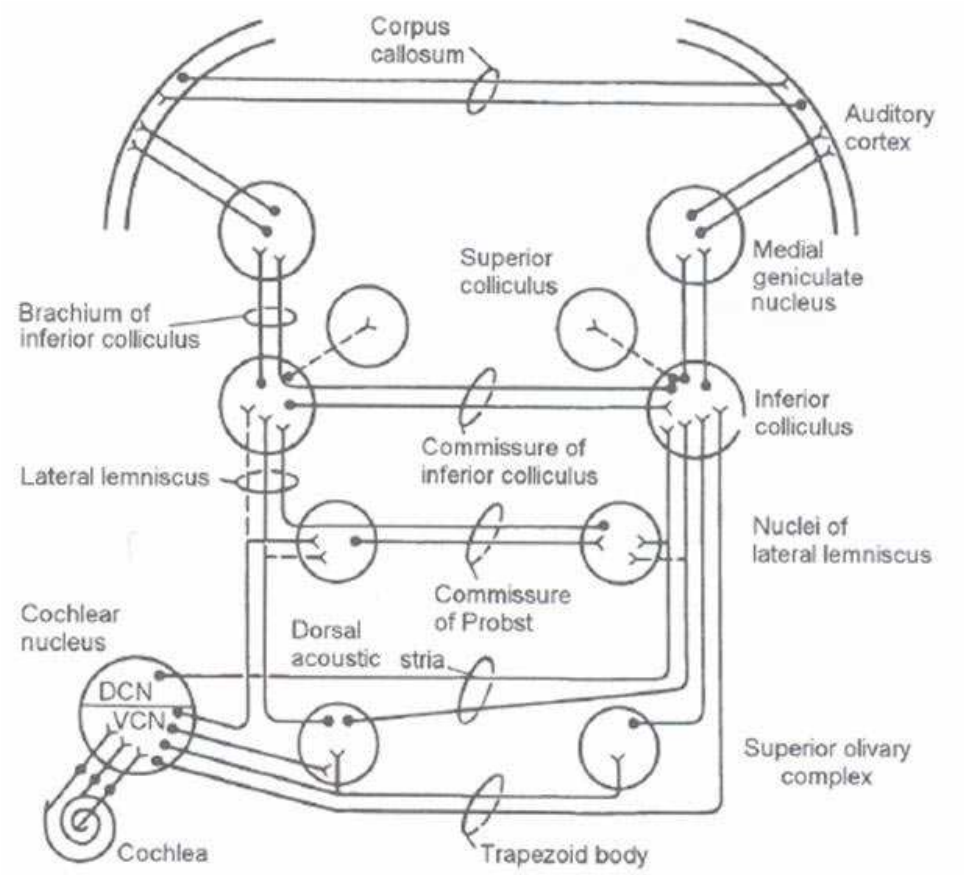


Figure 2-1. Diagram showing the ascending auditory pathways originating from the left cochlea. VCN: ventral cochlear nucleus. DCN: dorsal cochlear nucleus (adapted from Ehret and Romand, 1997).

2.1.2.1. The ear and the transduction of sound

The ear is the organ of hearing and balance in vertebrates. It consists of three main regions: the external, middle and inner ear. A schematic of the ear is displayed in figure 2-2. The following is a description of the functions of these regions with respect to the transduction of sound energy into nervous impulses.

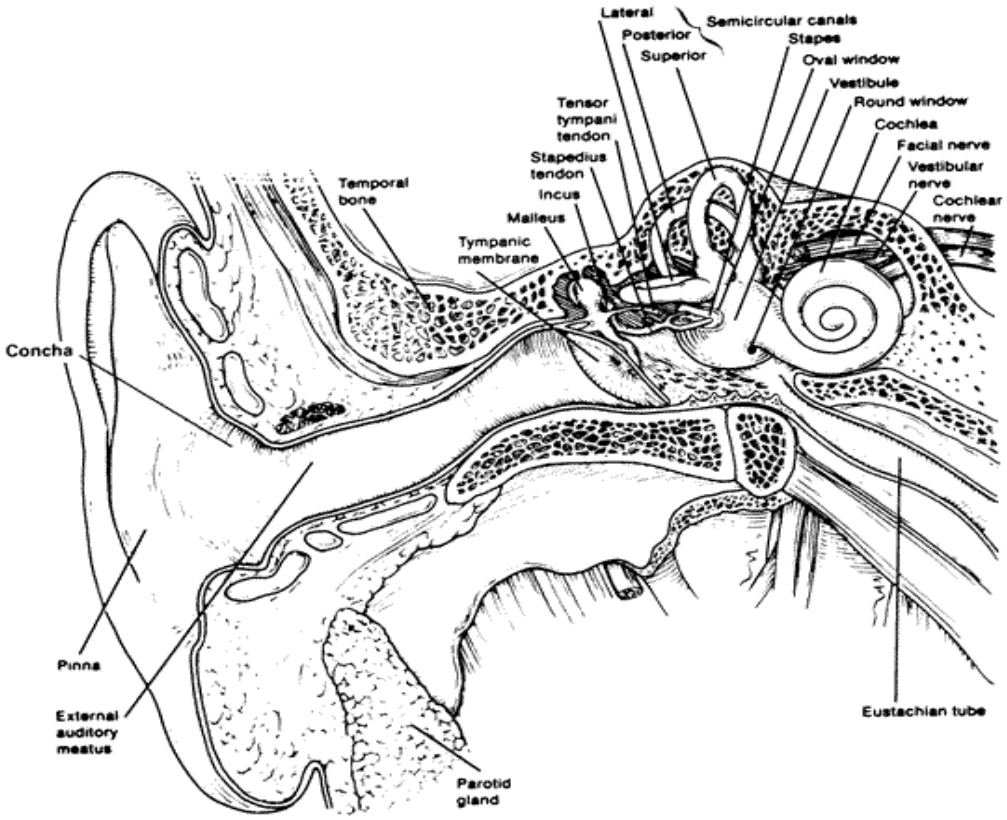


Figure 2-2. The external, middle and inner ears in man (reprinted from Pickles (1988), originally from Kessel and Kardon, 1979).

2.1.2.1.1. The external ear

The external ear consists of the pinna, the concha and the external auditory meatus (EAM). There are two functions of the external ear. Firstly, these three structures act to increase the sound pressure at the ear drum over the range of 2 to 7 kHz. Secondly, the modifications of sound waves in the external ear is direction-specific such that if confers directionality cues to the incoming sound (Phillips et al., 1982).

2.1.2.1.2. The middle ear

As little as 0.1 % of the energy is transmitted when sound travels from a low impedance medium such as air to a high impedance medium such as water. The impedance of the cochlear fluids is much higher than that of air which would result in a substantial loss of the sound energy if the energy carried in the sound waves were to be transformed directly into movements of the cochlear fluids. To counteract this energy loss the middle ear utilises three small bones (the malleus, incus and stapes), shown in figure 2-3, to act as an impedance transformer. The principle of this impedance transformer is that the area of the ear drum or tympanic membrane is much greater than that of the stapes footplate that is connected to the oval window. The pressure at the oval window is increased by the ratio of the two areas, leading to a matching of the impedances of air and the cochlear fluids. Two other factors, the buckling and the lever action of the ear bones also improve the impedance match but to a smaller degree (Pickles, 1988).

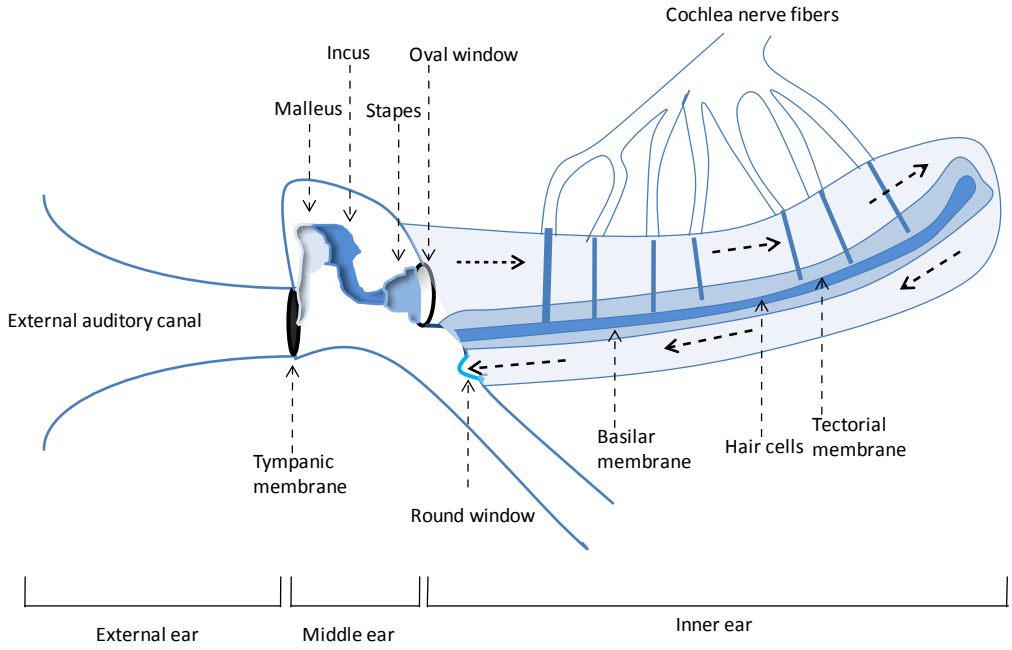


Figure 2-3. Schematic of the ear, with the cochlear spiral unravelled.

2.1.2.1.3. The inner ear

The cochlea is shaped like the spiral shell of a snail but in figure 2-3 it has been unravelled for clarity. The vibrations that are transmitted on to the oval window by the stapes cause a movement of the cochlear fluids, displacing fluid to the round window (see Figure 2-3). This, in turn, causes a wave of displacement of the basilar membrane (BM) which travels in an apical direction in the cochlea. One of the most important aspects of the manner in which the auditory system analyses sound is that the pattern of this displacement is different depending on the frequency of the sound wave. This is due in part to the mechanical properties of the BM; it is relatively narrow and stiff at the base and becomes gradually wider and less stiff towards its apex (Moore, 2003). For a high-frequency sound, the BM will show maximum displacement near the base while a low-frequency sound will elicit a pattern

of vibration along the length of the BM that reaches a maximum towards the apex. For a pure tone, with a single frequency, the vibration has a sharp peak which is confined to a narrow region of the basilar membrane. This frequency selectivity, caused by the mechanics of the cochlea, is the basis of the frequency selectivity that is observed at later stages of the auditory system.

Figure 2-4 shows how the velocity and estimated amplitude of the BM at the 18 kHz point vary with the frequency of stimulation. The points that are plotted on these curves show the level of tone that was necessary to cause an equal velocity or amplitude of BM vibration for each frequency. Figure 2-4 shows that at this point the BM is most sensitive at one frequency (the characteristic frequency or CF) because a lower level tone at 18 kHz elicits the same magnitude of response as higher level tones at other frequencies. The BM at this point is also sharply tuned to this frequency because the level of a tone away from 18 kHz has to be increased markedly for the same magnitude of response to be elicited. This manner of displaying data, the frequency-threshold curve (FTC), is useful because it allows comparisons with tuning from other stages of the auditory system, for example the auditory nerve (AN; shown as a dashed line in Figure 2-4).

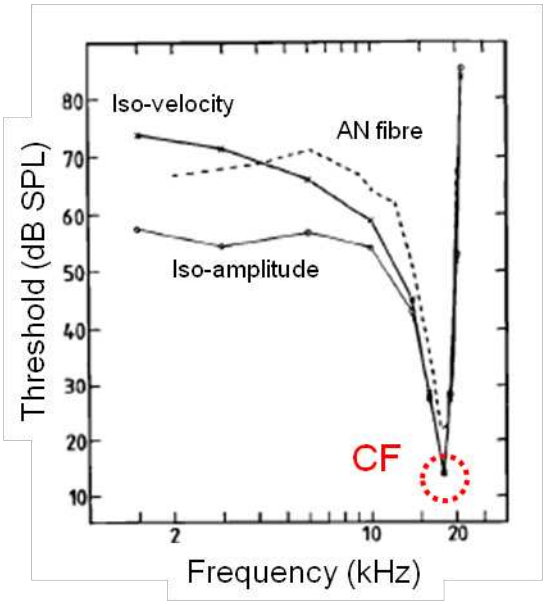


Figure 2-4. Tuning curves for the 18 kHz point on the guinea pig basilar membrane. The tuning curves were constructed by plotting the frequency points at which either the velocity or the estimated amplitude of the basilar membrane were equal (solid lines). The dashed line shows data from an auditory nerve fibre with a comparable characteristic frequency. The characteristic frequency (described in the text) is shown as a dashed red circle (adapted from Sellick et al., 1982).

Each point on the BM can be considered as a bandpass filter (one that attenuates frequencies that are lower and higher than a certain centre frequency). The bandwidth of the filter can be used to define how sharp the tuning is. It is common to use the 10 dB bandwidth which is the difference between the two frequencies at which the response has decreased by 10 dB from the minimum threshold. However, as the bandwidth shows an increase with the CF on the BM, it can be useful to define the relative bandwidth of the filter by dividing the bandwidth by the CF. The reciprocal of this ratio is the quality factor, which for a 10 dB bandwidth is denoted Q_{10}. BM Q_{10} values in guinea pig are usually between 0.2 and 6 (Evans et al., 1992). The notion that

the basilar membrane operates as a series of auditory filters is important for understanding the approaches that have been taken in investigating the auditory system and will be returned to later.

The motion of the BM is transformed into neural impulses by the inner hair cells located in the organ of Corti which is situated on the basilar membrane (see figure 2-5). The organ of Corti contains one row of inner hair cells (IHCs) and between 3 and 5 rows of outer hair cells (OHCs), depending on the species. The function of outer hair cells may be to actively influence the mechanics of the cochlea to produce high sensitivity and sharp tuning (Moore, 2003).

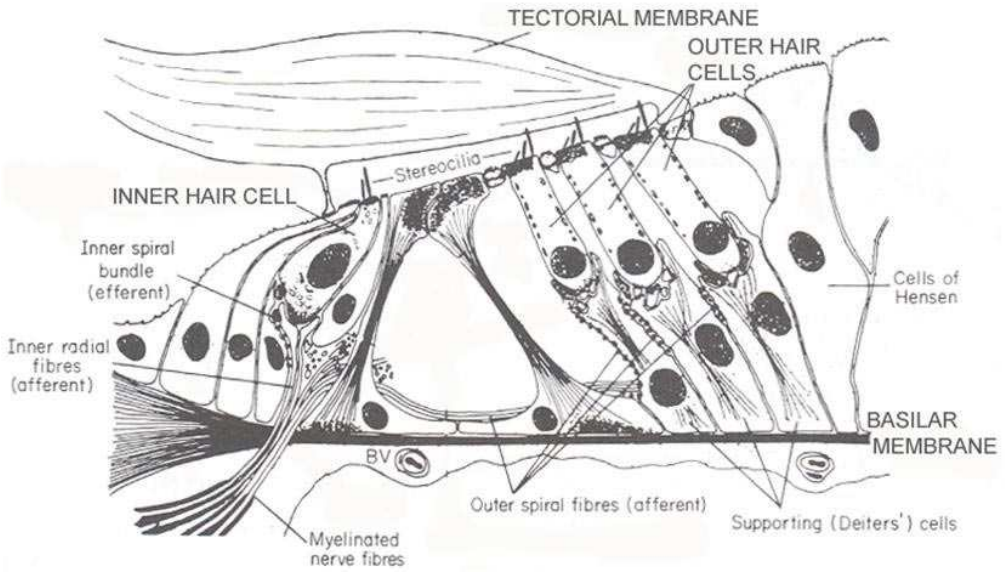


Figure 2-5. Cross-section of the cochlea, showing the basilar membrane, the tectorial membrane and the organ of Corti (adapted from Moore, 2003).

Figure 2-6 shows a schematic of an inner hair cell. The hairs or stereocilia are joined by tip links. Deflection of the stereocilia in the direction

indicated in figure 2-6 causes tension on the tip links leading to the opening of mechano-sensitive ion channels in the stereocilia. This leads to an influx of potassium ions (K^+) into the hair cell which causes a depolarisation. This in turn leads to the release of neurotransmitter by the hair cell which causes excitatory post-synaptic potentials in the afferent nerve fibres in figure 2-6 which may lead to action potential generation. In fact, the ion channels on the stereocilia are partly open at rest so a deflection of the stereocilia in the opposite direction causes them to close leading to a hyperpolarisation of the hair cell resting potential (Pickles, 1988). Generally, there is a close correspondence between the mechanical response of the basilar membrane and the inner hair cell responses so that, for example, tuning curves are similar (Cody and Russell, 1987).

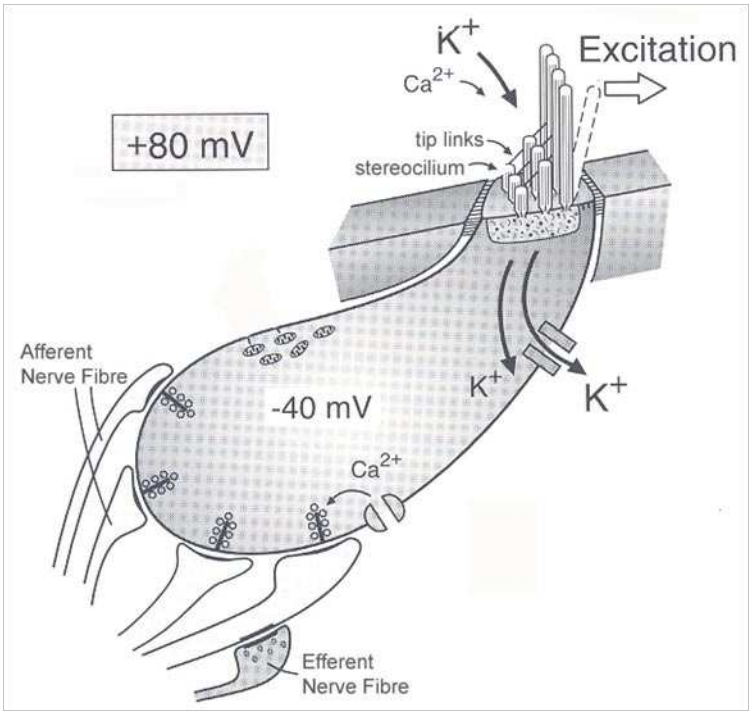


Figure 2-6. Physiology of an inner hair cell.

2.1.2.2. The auditory nerve

The afferent nerve fibres shown in figure 2-6 are also called auditory nerve (AN) fibres and lead directly from the hair cells to the cochlear nucleus. AN fibres can be classified depending on their firing rate at rest: the spontaneous rate (Liberman, 1978). Around 61% of fibres have high spontaneous rates (18 – 250 spikes per second), 23% have medium rates (0.5 – 18 spikes per second) and 16% have low spontaneous rates (less than 0.5 spikes per second).

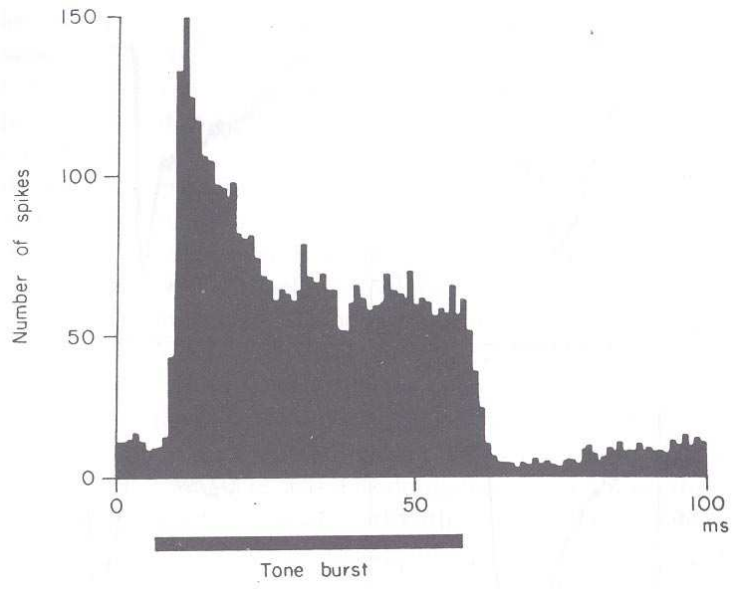


Figure 2-7. PSTH response to 50 ms tone presented at the CF (2 kHz) of an AN fibre (reprinted from Pickles (1988), originally from Kiang, 1965).

The response of an auditory nerve (AN) fibre to a tone is displayed as a peri-stimulus time histogram (PSTH) in figure 2-7. The PSTH was constructed by repeating the stimulus several times and then counting the number of action potential or spike responses contained within regularly-spaced time bins. The response consists of an initial sharp peak at the onset of the tone, followed by a gradual adaptation to a relatively stable rate of discharge. At the offset of the tone, there is a sharp decrease in activity that dips below and then recovers back up to the level of the spontaneous activity.

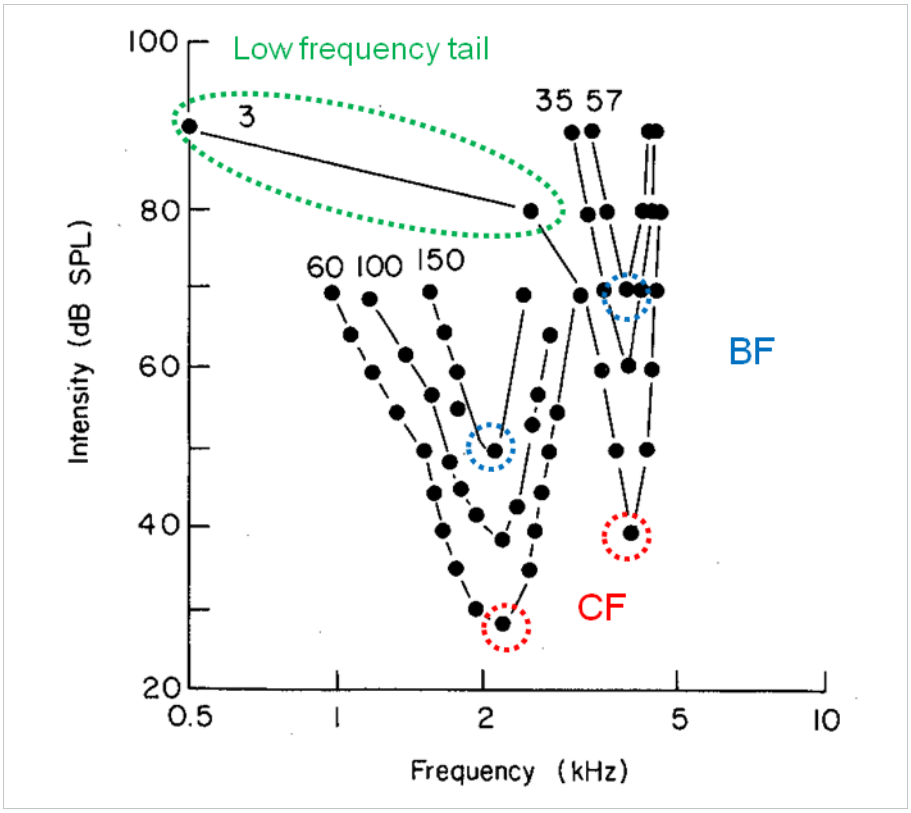


Figure 2-8. Tuning curves constructed at different firing rate criteria become sharper as the rate criterion is increased. Rate criterion is printed next to each tuning curve (adapted from Pickles (1988), originally from Evans, 1975).

Following from the BM and hair cell responses, AN fibres are frequency selective, that is they respond more to some frequencies than to others. Figure 2-8 shows examples of tuning curves constructed at different rate criteria for 2 AN fibres. The tuning curves are similar to those of the BM shown in figure 2-4, though figure 2-8 highlights further properties of the tuning. Firstly, AN fibres tends to fire more when the intensity of the tone is greater and when the frequency is nearer to the CF such that raising the rate criterion results in a tuning curve with a higher threshold and narrower tuning. A distinction can be made between the characteristic frequency and the best frequency (BF), as indicated by the dashed blue circles in figure 2-8,

though they tend to be similar. The BF is the frequency that elicits the maximal neural response. A further aspect of AN tuning curves, that was also apparent in the BM tuning curves of figure 2-4, is that they tend to be symmetrical when the CF is low whereas they become asymmetrical when the CF is higher, for example having a low frequency tail which is indicated by the dashed green oval in figure 2-8.

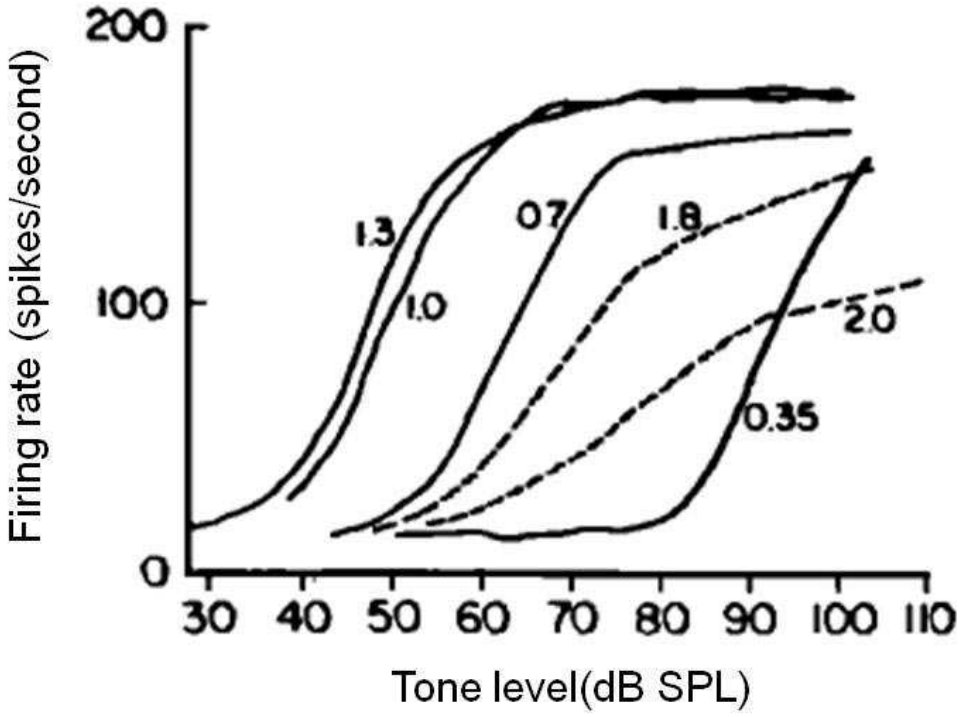


Figure 2-9. Rate-level functions for one auditory nerve fibre for different tone frequencies. The characteristic frequency of the fibre is 1.3 kHz and the tone frequency is indicated next to each function (adapted from Sachs and Abbas, 1974).

Figure 2-9 shows the relationship of firing rate to tone level (rate-intensity function) at different frequencies for an auditory nerve fibre. The firing rate increases as the tone level increases in a monotonic fashion but

tends to saturate at high sound levels. This pattern is observed at CF in all reported AN fibres (Kiang, 1965; Sachs and Abbas, 1974; Winter et al., 1990).

2.1.2.3. The cochlear nucleus

The first auditory processing station after the auditory nerve is the cochlear nucleus (CN). The CN is subdivided into the dorsal cochlear nucleus (DCN) and the ventral cochlear nucleus (VCN). The VCN can be further subdivided into anterior (AVCN) and posterior (PVCN) sections. Each of the subdivisions is tonotopically organised: they are arranged such that there is a regular gradient in the BF of neurons along one axis of the nucleus. Responses of neurons in the CN are more heterogeneous than those of AN fibres. Neurons are divided into four gross classifications with each group being subdivided depending on more specific cell response properties, for example PSTH classification and receptive fields. Primary-like (PL) cells have a similar response profile to those in the auditory nerve and are found throughout the VCN. The responses are particularly similar to AN responses in the AVCN where cells exhibit, for example, similar tuning curve shapes and monotonic rate-level functions. Chopper (Ch) units are found throughout the CN and produce a rapid, consistent response at the beginning of a stimulus followed by a train of spikes with specific inter-spike intervals for the duration of the stimulus. Onset (On) units are found throughout the CN and respond only to the onset of a stimulus at specific frequencies (or frequency ranges) and levels. Finally, pauser and build-up (PB) units are found predominantly in the DCN. Pauser units demonstrate an initial onset response, a silent period and

then a gradual build-up of activity while build-up units demonstrate a gradual build-up without the onset response (Pickles, 1988).

2.1.2.4. The Superior Olivary Complex

The superior olivary complex (SOC) are some of the first nuclei of the auditory brainstem to receive input from both ears and so their main function is to play a role in auditory localisation (Pickles, 1988). The SOC is further divided into the lateral and medial nuclei of the superior olive (LSO and MSO) and the medial nucleus of the trapezoid body (MTB). The LSO receives direct input from PL units in the ipsilateral AVCN (Joris et al., 2004) and input from PL units in the contralateral AVCN that has been relayed through the ipsilateral MTB. Most of the binaural LSO neurons are excited by ipsilateral stimuli and inhibited by contralateral ones (EI cells). The MSO receives direct input from the AVCN on both sides of the brain. In contrast to the LSO, the majority of neurons are excited by inputs from both ears (EE cells) and the minority are EI cells (Pickles, 1988).

2.1.2.5. The Inferior Colliculus

The inferior colliculus (IC) is an obligatory processing centre for most information that ascends the auditory system through the medial geniculate body to the auditory cortex (Joris et al., 2004). It has two tonotopically-organised structures concerned with auditory relay: the central nucleus (CNIC) and the dorsal cortex (DCIC). I will concentrate on the neural responses

in these regions though there are also the paracentral nuclei which are primarily somasthetic and auditory integrative areas (Pickles, 1988).

The CNIC has a pronounced laminar structure (Malmierca et al., 1995) that confers upon the nucleus a highly tonotopic frequency organization (Schreiner and Langner, 1997). The laminae within the IC are approximately isofrequency and consist of two cell types: disc-shaped and stellate (Malmierca et al., 1993). The CNIC is the location of a convergence of temporal, spectral and spatial information that has been extracted in parallel earlier in the auditory pathway (Joris et al., 2004) and receives ascending inputs from both the SOC and CN. Neurons in the CNIC exhibit a diversity of tuning and temporal responses that probably reflect complex excitatory and inhibitory interactions. For example, half of the neurons exhibit non-monotonic rate-level functions (Ryan and Miller, 1978). In addition, response types can vary from onset and pauser types in anaesthetised preparations (Rose et al., 1963; Ryan and Miller, 1978) to more sustained responses in awake animals (Bock et al., 1972). As in the SOC, many neurons display binaural response interactions (Rose et al., 1963; Yin et al., 1986; Caird and Klinke, 1987). Tuning in CNIC neurons is varied, though some neurons exhibit extremely narrow tuning with Q_{10} values between 25 and 40 (Aitkin et al., 1975). Many neurons in the DCIC receive somatosensory along with auditory input, as well as extensive descending projections from the auditory cortex (Saldana et al., 1996) and the cells are tonotopically organised, having broad or complex tuning curves (Aitkin et al., 1975; Aitkin et al., 1981).

2.1.2.6. The Medial Geniculate Body

The medial geniculate body (MGB) is an obligatory relay of information from the midbrain to the auditory cortex (Joris et al., 2004). It consists of 3 main regions: the dorsal, medial and ventral MGB. Only the ventral MGB is a specific auditory relay; its inputs are mainly from the ipsilateral central nucleus of the IC (Calford and Aitkin, 1983) and it projects principally to the primary auditory cortex. The dorsal and medial divisions of the MGB represent more diffuse or nonspecific auditory areas and so will not be considered further here. All of the features highlighted above in the central nucleus of the IC are apparent in the ventral MGB, for example sustained responses in awake preparations (Aitkin and Prain, 1974; Calford, 1983), non-monotonic rate-level functions (Aitkin and Prain, 1974) and binaural sensitivity (Aitkin and Webster, 1972). However, tuning as narrow as that observed in the central nucleus of the IC has not been reported (Aitkin and Webster, 1972).

2.1.2.7. The auditory cortex

Like lower nuclei, the auditory cortex can be separated into several regions or cortical fields depending on the underlying connectivity and the physiological responses (Wallace et al., 2000; Read et al., 2002). An important feature that is used to define primary cortical fields is a tonotopic organisation which reflects the fact that a majority of their input comes from the ventral MGB (Read et al., 2002). This thesis is concerned with the

response properties of the primary auditory cortex (A1) and so I will not elaborate on the other cortical areas further.

Neurons in A1 display several functional differences to those in the AN and reflect the output of the multiple layers of processing which I have briefly described. A key feature of A1 neurons is the diversity of response characteristics formed due to convergence and integration of excitatory and inhibitory inputs that presumably reflect functional specialisations; in fact it has been shown that neurons in the auditory cortex of awake animals respond most strongly to a specific preferred stimulus (Wang et al., 2005). In this section I discuss the properties of A1 neurons and how these properties may reflect functional specialisations as a precursor to describing the responses to two-tone stimuli.

2.1.2.7.1. Frequency response areas

The receptive fields (RFs) of A1 neurons are not necessarily V-shaped as they are in the AN, reflecting excitatory and inhibitory modulations of the AN output in the subcortical nuclei. Instead, RF shapes seem to form a continuous distribution that differs along the level and frequency dimensions. Further, there seems to be little correlation between RF parameters, for example CF, bandwidth and threshold, suggesting that essentially any combination of shape parameters could be achieved (Moshitch et al., 2006). An additional property of the auditory cortex is that neurons with different response types may form separate subpopulations in different parts of A1.

Along the level dimension, cortical neurons can be broadly classified as monotonic or non-monotonic depending on the rate-intensity function, and the frequency of these two response types seems to be fairly equal, at least in cat (Phillips et al., 1985; Sutter and Schreiner, 1995; Moshitch et al., 2006). Also, non-monotonic responses seem to be segregated in central A1 (Heil et al., 1994; Phillips et al., 1994). Monotonic neurons tend to display a V-shaped tuning curve (see example in figure 2-10a), like those in the AN (Phillips et al., 1985), though neurons with other RF shapes display a range of rate-level functions, from strongly non-monotonic to monotonic. Two types of RF shape that have not been observed in the AN are circumscribed (see figure 2-10c) and multi-peaked (see figure 2-10b). Circumscribed or closed RFs are characterised by non-monotonic rate-level functions (Sutter, 2000) and in some cases highly specific frequency tuning (Phillips et al., 1994; Sadagopan and Wang, 2008). Multi-peaked RFs are characterised by two or more distinct excitatory frequency ranges and are mainly found in dorsal A1 (Sutter and Schreiner, 1991; Sutter, 2000; Moshitch et al., 2006).

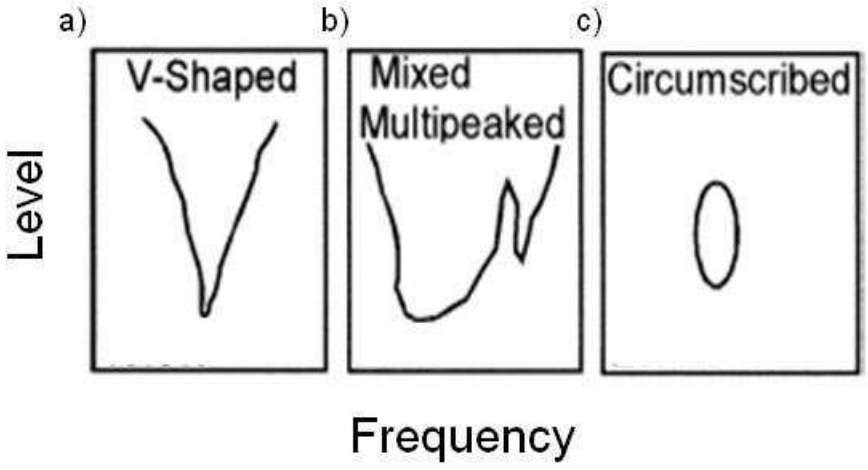


Figure 2-10. Selected exemplars of RF shape in the primary auditory cortex (adapted from Sutter, 2000).

The exact proportions of various neural RF types in the auditory cortex seem to differ between studies, presumably due to sampling, anaesthetic and possibly species differences. For example, it is well established that anaesthesia strongly affects the responses of neurons at several levels of the auditory system including the auditory cortex (Sally and Kelly, 1988; Sutter and Schreiner, 1991; Schreiner et al., 1992; Sutter and Schreiner, 1995; Gaese and Ostwald, 2001). Studies conducted in awake preparations (Qin et al., 2003) or using halothane, which seemed to generally maintain the properties of awake responses (Moshitch et al., 2006), indicated that bandwidths of auditory cortical neurons could be up to three or four times wider than when barbiturate or ketamine anaesthesia were used (Sutter and Schreiner, 1991; Schreiner and Sutter, 1992). A recent study has also called into question the proportion of circumscribed RFs in auditory cortex (Sadagopan and Wang, 2008). Specifically, 64% of neurons in A1 of awake marmoset monkeys were classified as circumscribed, being narrowly tuned to both frequency and level.

Such highly specific neurons could have been missed in previous studies due to sampling that was too coarse to excite these neurons, and more likely to excite neurons with V-shaped tuning. From a functional perspective, these neurons could be involved in the level-invariant coding of sound frequency, where variations in frequency could be coded by firing rate without the confounding influence of level. In contrast, neurons with V-shaped RFs could act as detectors of sounds over a wide range of frequency and level (Sadagopan and Wang, 2008).

2.1.2.7.2. Temporal responses

Another feature of auditory cortical responses that demonstrates differences due to anaesthesia is the temporal response to sound (Moshitch et al., 2006). As mentioned, the response to tones in the AN is always like that displayed in figure 2-7, a large increase in spike count at the onset of a tone and then a sustained response while the tone is still present. In contrast, neurons in the auditory cortex can display a variety of temporal response patterns. In many anaesthetised preparations responses are more phasic, consisting of a short burst of action potentials shortly after tone onset, and possibly after tone offset, but with no sustained activity (Eggermont, 1991; Phillips and Sark, 1991). In awake animals (Evans and Whitfield, 1964; Abeles and Goldstein, 1972; Pfingst and O'Connor, 1981; Shamma and Symmes, 1985; Qin et al., 2003), or those under halothane anaesthesia (Moshitch et al., 2006), temporal responses are more rich, ranging from phasic responses, as in barbiturate-anaesthetised animals, to sustained responses, like those in the

auditory nerve, to through responses, where the response continues even after the tone has finished. When combined with the possibility that the RFs of cortical neurons could assume nearly any shape (Moshitch et al., 2006), such a rich set of temporal response patterns conveys a large degree of flexibility to cortical responses that is not possible at the level of the AN. Indeed, it is likely that neurons or populations of neurons in the cortex are specialised to perform specific auditory processing tasks.

2.2. Temporal resolution in the auditory system

In this section I introduce one method that has been used psychophysically to estimate the temporal and frequency resolution of the auditory system: forward masking. I then describe studies that have investigated a neural correlate of forward masking in several nuclei of the auditory system. I finish this section by discussing other stimuli that have been used to investigate the temporal resolution physiologically at several levels of the auditory system and by introducing the concept of adaptation.

2.2.1. Psychophysical masking

Psychophysical masking refers to an increase in the detection threshold of one sound (probe) when it is simultaneous with, preceded by (forward masking) or succeeded by (backward masking) another sound (masker). Masking studies have been used to estimate both the tuning properties and the temporal resolution of the auditory system (Moore, 1993; Viemeister and Plack, 1993).

To estimate the tuning properties, a probe tone was chosen at a certain frequency with a low level, for example 10 dB SL. As the probe tone was at a low level, it was assumed that it would primarily excite a single auditory filter. The masker was either a noise or a tone and, for each of several masker centre frequencies, the level of the masker was adjusted such that the probe was only just masked. The function of the masker levels against the masker centre frequency is known as a psychophysical tuning curve (PTC), and it gives an idea of the tuning resolution of the channel at that probe tone frequency.

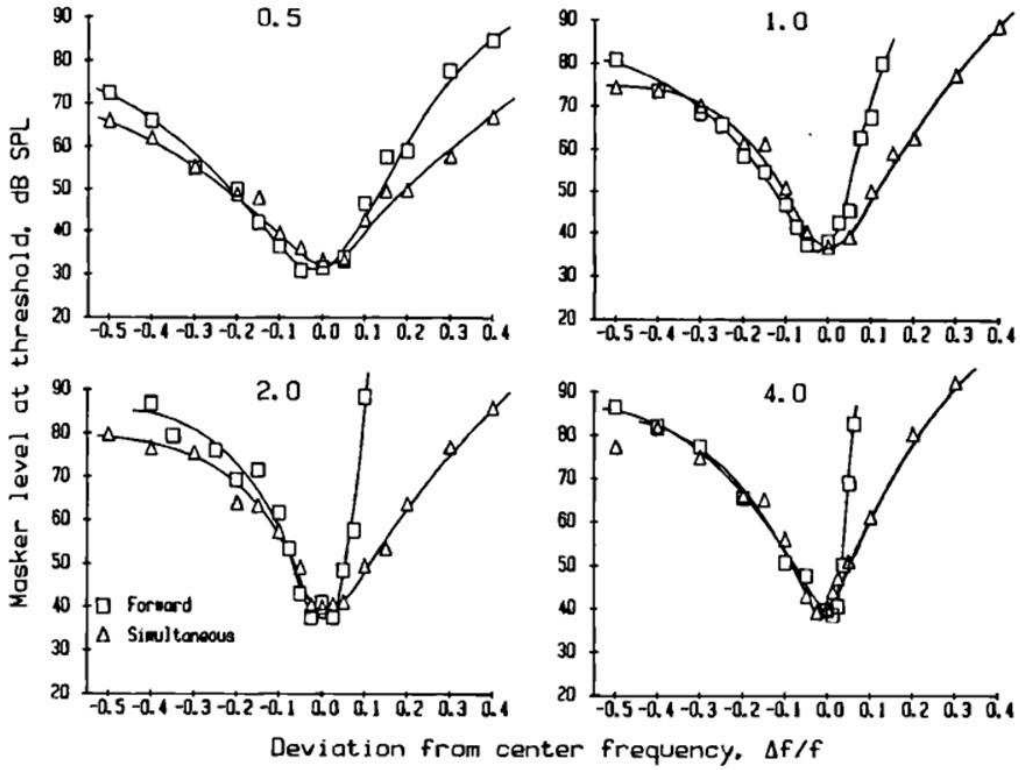


Figure 2-11. Comparison of psychophysical tuning curves measured using simultaneous (triangles) and forward (squares) masking (reprinted from Moore et al., 1984).

Examples of PTCs are displayed in figure 2-11 for different probe centre frequencies when both simultaneous and forward masking are used to

determine the PTC. The PTCs determined using forward masking are narrower and more asymmetrical than those determined using simultaneous masking. In brief, this may be because the masker and probe responses interact in a nonlinear fashion on the basilar membrane when the tones are presented at the same time (see Moore, 1993 for a more in-depth discussion). The amount of interaction between the masker and probe sounds on the basilar membrane in forward masking is negligible except for the shortest masker-probe delays (Duifhuis, 1973).

As this thesis is concerned with how the auditory system processes non-simultaneous sequential sounds, I will henceforth focus on the characteristics of forward and backward masking. The characteristics of forward masking have been investigated by varying the parameters, for example duration and level, of both the masker and probe sounds. Manipulation of these parameters can lead to changes both in the magnitude of forward masking, manifested as changes in the detection threshold, and in the recovery from forward masking, which is characterized by increasing the silent gap between the masker offset and probe onset (the inter-stimulus interval or ISI).

2.2.1.1. The effect of varying masker and probe characteristics in forward masking studies

Figure 2-12 shows the effect of varying the masker level and duration on the recovery from forward masking. As displayed in figure 2-12a, the effect of increasing the masker level is to increase the initial magnitude of and the

amount of time to recovery from forward masking. Extrapolation of the curves to the right gives a maximal duration of forward masking between 200 and 300 ms. A further point from figure 2-12a is that the function relating masker level and the amount of masking is compressive. Although the masker levels cover a range of 60 dB, the thresholds cover a range of less than 40 dB. A similar relationship is observed when the masker duration is increased, as shown in figure 2-12b. The effect of increasing masker duration can be observed up to durations of 300-500 ms (Weber and Green, 1978; Kidd and Feth, 1982; Carlyon, 1988).

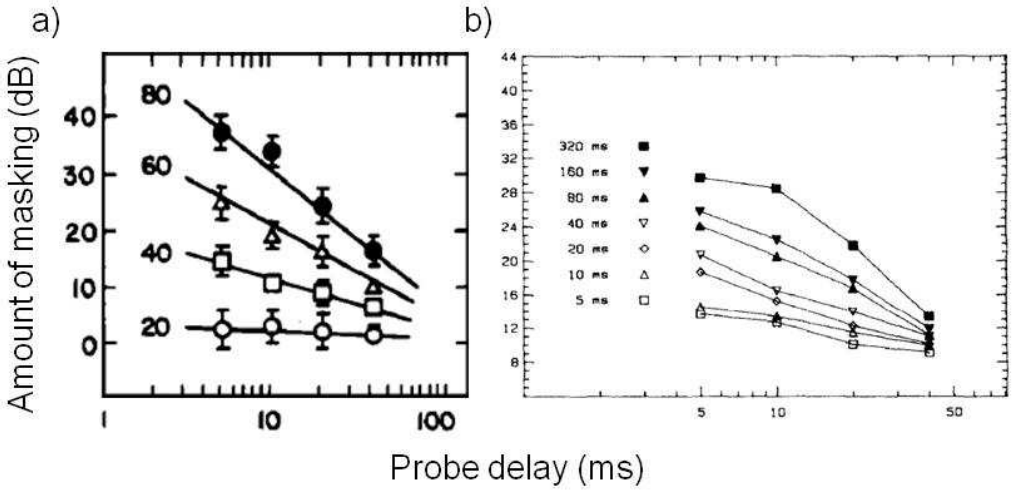


Figure 2-12. The effect of varying a) the level and b) the duration of the masker on the amount of forward masking (Data from Jesteadt et al., 1982 (a) and Kidd and Feth, 1982 (b)).

As noted, the probe tones used in psychophysical forward masking studies generally have a short duration and a low level, so that they mainly excite a single frequency channel (Moore, 1993). The characteristics of the probe tone have not been investigated systematically, though it has been

reported that increasing the duration of the probe tone can decrease the amount of forward masking that occurs (Turner et al., 1994). Furthermore, more forward masking has been reported for low frequency than for high frequency probe tones in the range of 125 to 4000 Hz (Jesteadt et al., 1982).

2.2.1.2. Backward masking

Although backward masking has been shown to affect suprathreshold tasks such as lateralization and frequency discrimination (Massaro et al., 1976), it is not clear whether backward masking has any discernible effect when two sounds are presented in isolation (Viemeister and Plack, 1993). However, the amount of masking when both backward and forward masking are combined is greater than the sum of the masking from identical forward and backward maskers separately (Bilger, 1959).

2.2.2. Physiological sequential two-tone interactions

The following section addresses how the neural response to the second tone in a sequential two-tone paradigm is influenced by the first tone. A summary of the development of two-tone interactions as the auditory system is ascended is given in table 2-1.

A note on nomenclature: The term ‘forward masking’ has been used both to refer to the increase in psychophysical detection threshold and to the decrease in physiological response associated with a prior sound. To avoid any confusion with psychophysical effects, the decrease in physiological response considered in this thesis will be referred to as suppression and the

increase in psychophysical threshold will be referred to as masking. Also, for physiological responses, the first tone in a two-tone paradigm will be referred to as the conditioner tone (rather than the masker) and the second as the probe tone.

2.2.2.1. Sequential two-tone interactions in the AN

2.2.2.1.1. Forward suppression in the AN

In the AN there is a simple relationship between both the magnitude and time course of forward suppression and the response to the conditioner tone. Forward suppression can be considered as a process of recovery from the adaptation that occurs during the response to the conditioner tone. An increase in the adaptation to the conditioner tone leads to an increase in the magnitude of suppression and the time constant of recovery from suppression. Essentially, the suppression that is observed varies depending on the adaptation to the conditioner tone so that conditioner tone characteristics, for example duration, intensity and spectral content only influence the suppression indirectly (Smith, 1977; Harris and Dallos, 1979). No relationships have been found between the probe characteristics, for example level (Smith, 1977) and frequency (Abbas, 1979), and the amount or duration of suppression. Also, there was no significant difference between the recovery time constants for fibres with different CFs (Harris and Dallos, 1979), though differences have been reported between AN fibres with different spontaneous rates (Relkin and Doucet, 1991).

Time course of forward suppression

The recovery from suppression can be estimated with a single exponential function. A single logarithmic function fits data from 10 to 100 ms, with time on a logarithmic scale, as shown in figure 2-13. From this figure, it can be seen that forward suppression lasted around 100 – 200 ms in chinchilla AN fibres (Harris and Dallos, 1979). In the Mongolian gerbil, recovery times ranged from 40 to 310 ms, with a mean of 115 ms (Smith, 1977).

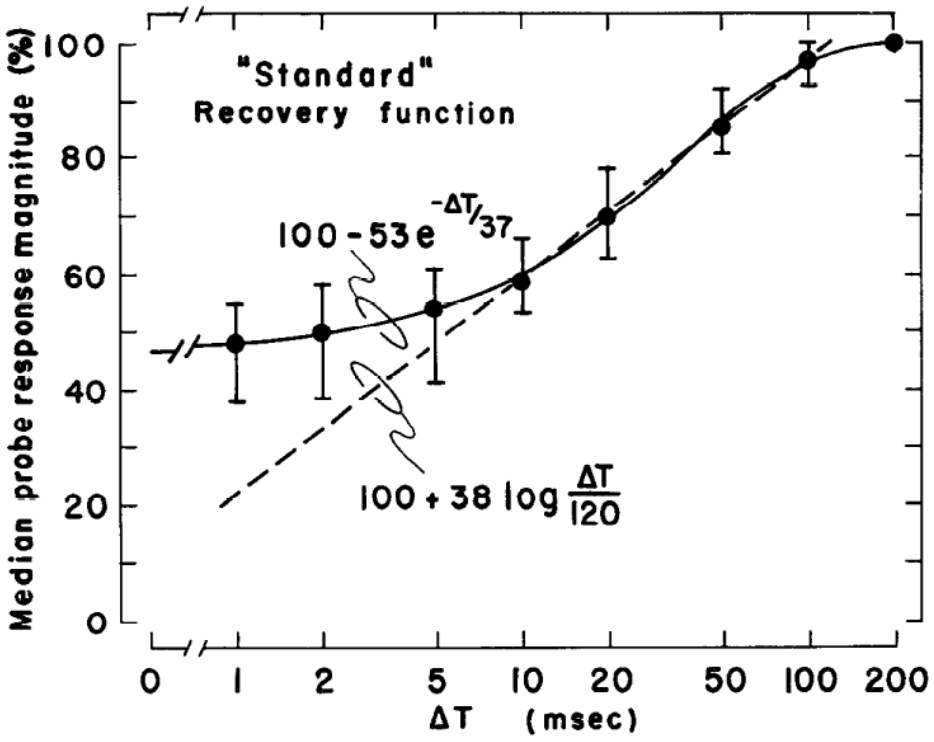


Figure 2-13. The recovery from suppression showing the medians (filled circles) and interquartile ranges (vertical bars) calculated from 24 chinchilla AN fibres (reprinted from Harris and Dallos, 1979)

Conditioner tone duration

If the duration of the conditioner tone was increased, more adaptation was observed and the response decrement to the probe tone increased (Smith, 1977). This relationship held until the adaptation reached the steady-state phase, after which increases in conditioner tone duration did not result in further suppression. This saturation occurred at conditioner tone durations of around 100 – 200 ms (Smith, 1977; Harris and Dallos, 1979).

Conditioner tone level

In a similar way, increases in conditioner tone level led to increases in the magnitude of suppression and the time constant of recovery from suppression. It was demonstrated in the chinchilla that the suppression/level function was in general agreement with the rate/level function for the median data from 12 neurons, with both responses saturating at between 40 and 50 dB above the neuron's threshold (Harris and Dallos, 1979). In the Mongolian gerbil, this saturation seems to occur at around the same conditioner tone intensity range, though only two example units were displayed (Smith, 1977).

Conditioner tone frequency

The effect of the frequency of the conditioner tone has also been investigated. The suppressed threshold at a given frequency could be defined by starting with a conditioner tone with a low intensity and then incrementing the intensity of the conditioner tone until a probe tone decrement of 20% was observed (Harris and Dallos, 1979). Joining the points at different

frequencies gave the suppressed threshold curve. With a 20% suppression criterion, these curves were found to be similar to frequency threshold curves constructed using a criterion of a 5% increase in rate above the spontaneous activity. So, suppression was only observed when the neuron responded to the conditioner tone. To further investigate whether suppression was independent of the frequency of the conditioner tone, conditioner tones with different frequencies that elicited the same magnitude of suppression were investigated, along with noise. The recovery functions were similar indicating that the spectral content of the conditioner tone had no direct effect; rather suppression was dependent on the adaptation to the conditioner tone alone (Harris and Dallos, 1979).

Relation to spontaneous rate

For AN neurons with a high spontaneous rate the time course of recovery from suppression and the time course of recovery from adaptation, as indexed from the time course of recovery of spontaneous activity, were similar. For neurons with lower spontaneous rates the recovery from suppression was also similar to the recovery from adaptation, but the initial part of the curve was truncated because the spontaneous activity had been reduced to zero (Harris and Dallos, 1979).

Mechanisms of forward suppression in the AN

The forward suppression that is observed in the AN is most likely due to a mechanism operating at the synapse between the hair cell and the AN and/or axonal transmission characteristics of AN fibres (Chimento and

Schreiner, 1991). This is because adaptation is not observed in the hair cell (Russell and Sellick, 1978).

Comparison to psychophysical masking

As described previously, forward masking is manifested psychophysically as an increase in the detection threshold of a signal which is preceded by a masker. The detectability of a signal depends not only on the response but also on the variance of the response. Physiological studies of forward suppression in the AN demonstrated a decrease in response to the probe tone but ignored the variance of the response and thus the detectability of the probe tone. This was addressed in AN fibres in the chinchilla using signal detection theory (Relkin and Pelli, 1987; Relkin and Turner, 1988). They used a two-interval forced-choice paradigm; one interval contained a conditioner tone alone while the other interval contained both a conditioner tone and a probe tone. The response to the probe tone was then compared to an equivalent time window in the interval where no probe tone was presented. So, in essence, the probe tone response was compared to the activity in an equivalent time period after an identical conditioner tone. By making the comparison on a trial-by-trial basis the variability of the response was taken account of. The probe tone was defined as detected if the number of spikes in the interval containing the tone exceeded the number of spikes in the equivalent interval with no probe tone. The detectability of the probe tone was thus dependent on the ratio of the mean response to the variance of the response. A neurometric function was then created by plotting the

percent correct identifications against probe tone intensity. A threshold was then set at 66% of correct responses.

It was found that the threshold shifts for AN fibres were less than those observed psychophysically, so AN fibres could detect the presence of a probe tone under stimulus conditions for which the behavioral observer could not. This was especially true for high spontaneous rate neurons, for example in one neuron an 80 dB SPL conditioner tone elicited a threshold shift of 15 dB. This can be compared to a threshold shift of around 30-50 dB in humans (Widin and Viemeister, 1979; Jesteadt et al., 1982). It was also found that the slopes of growth of masking functions (amount of masking as a function of masker level) for AN fibres were less (0.1 dB/dB for high SR to 0.5 dB/dB for low and medium SR) than those observed psychophysically (0.6 dB/dB from Jesteadt et al., 1982). Differences in the stimuli used in physiological and psychophysical studies limit the comparisons between the two sets of experimental data. A further study aimed to compare psychophysical and physiological threshold shifts using the same stimuli (Turner et al., 1994). It was found that changes in the rise/fall time and probe duration did not affect thresholds in the physiological preparation, but markedly affected psychophysical results. Additionally, a psychophysical study demonstrated that the time course of recovery from forward suppression was similar for patients with cochlear implants and normal-hearing listeners (Shannon, 1990). From this, the author infers that the time course of forward suppression must not be dependent on any of the missing details of cochlear

function, and thus that the mechanism of psychophysical forward masking is retrocochlear.

2.2.2.2. Sequential two-tone interactions in the CN

2.2.2.2.1. Forward suppression in the CN

In some neurons of the CN, forward suppression is similar to that observed in the auditory nerve. However, other CN neurons display marked differences in their responses to sequential two-tone stimuli that are likely due to synaptic interactions and inhibitory modulations of AN output. The following is a brief summary of the differences that have been reported.

Sequential two-tone interactions in the CN are correlated with unit type, such that primary-like and some chopper units displayed similar features to AN fibres, for example, similar time courses of recovery (Boettcher et al., 1990; Bleeck et al., 2006). However, onset units displayed complete suppression at short delays (Shore, 1995; Bleeck et al., 2006) and sustained chopper units displayed little or no suppression (Bleeck et al., 2006). Some units also displayed delayed suppression, a feature not observed in AN responses, where the suppression did not become evident until some time after the conditioner tone offset (Boettcher et al., 1990; Kaltenbach et al., 1993).

The simple relationship between the adaptation to the conditioner and the amount of forward suppression that was observed in the AN is not observed in some units of the CN. For example, in the DCN, suppression has

been reported for conditioner tones that were outside of the RF and would thus elicit no response from the neuron (Kaltenbach et al., 1993). Also, some units, for example pause build-up (Boettcher et al., 1990) and onset units (Shore, 1995), exhibited non-monotonic conditioner-level functions, such that increases in conditioner level did not necessarily result in a greater magnitude of suppression.

2.2.2.2.2. Forward facilitation in the DCN

A further change from two-tone interactions in the AN is the occurrence of forward facilitation, where the response to a tone *increases* when it is preceded by a conditioner tone. Facilitation of responses occurred mainly in pause build-up units in the DCN (Palombi et al., 1994) and lasted for around 250 ms (Bleeck et al., 2006).

2.2.2.3. Sequential two-tone interactions in the SOC

Forward suppression in the SOC (the majority of cells studied using sequential two-tone stimuli were from the LSO) was similar to that observed in the CN, except that the initial response decrements were larger and suppression generally lasted for longer (Finlayson and Adam, 1997). Primary-like neurons in the LSO exhibited the lowest magnitude of suppression implying that they received input from the VCN (Finlayson and Adam, 1997). As in the CN, delayed suppression was observed in a small number of neurons; however no facilitation was reported in the LSO (Finlayson and Adam, 1997). This could be due to sampling or to the fact that the LSO

receives the majority of its input from the AVCN, where facilitation is rare (Palombi et al., 1994). No studies have varied the conditioner tone frequency so it is unknown whether suppression would be observed when the conditioner tone was outside of the RF. Again, because the LSO receives input from the AVCN, it is unclear whether this aspect of suppression would be observed, given that it has only been reported in the DCN (Kaltenbach et al., 1993).

A novel feature of forward suppression in the LSO is the interaction of inputs from the two ears. This thesis concerns only diotic stimuli (the same stimulus at both ears) so the specific details of the interaction between inputs from the two ears will not be reviewed. However, the principal conclusion from the LSO study was that binaural stimulation only resulted in changes to the respective monaural suppression patterns when the suppression in one monaural input was greater (Finlayson and Adam, 1997).

2.2.2.4. Sequential two-tone interactions in the IC

The characteristics of forward suppression reported in the CN and LSO are present in the IC, in addition to a greater initial magnitude of forward suppression and forward suppression that lasted longer (Finlayson, 1999). Further, in contrast to the LSO, facilitation was observed in a subset of units, mainly with pauser/build-up response patterns. Both delayed suppression and delayed facilitation were also observed in two subsets of units. Binaural interactions were present in the IC however the interactive effects were weaker than in the LSO. The magnitude of suppression increased with

conditioner tone duration up to durations of around 100-200 ms. Suppression and facilitation were also observed when the conditioner tone was outside of the RF (Malone and Semple, 2001).

2.2.2.5. Sequential two-tone interactions in the MGB

Only one study of two-tone interactions in the MGB exists (Schreiner, 1981), in the awake guinea pig. Although the study did not systematically analyse sequential two-tone interactions in the way that has been done in other nuclei, the time constant of suppression was around 390 ms and the initial suppression was around 45%.

2.2.2.6. Sequential two-tone interactions in the cortex

2.2.2.6.1. Forward suppression

Forward suppression in the primary auditory cortex resembles forward masking to a greater extent than in the AN and is thus thought to underlie the perceptual phenomenon (Brosch and Schreiner, 1997). It displays the additional features of forward suppression that I have discussed in the subcortical nuclei. An important further aspect of cortical forward suppression is that it lasts for longer in general than that observed in the lower auditory nuclei. The following is a summary of effects in cortex that have already been noted in lower auditory nuclei.

Firstly, responses to the probe occurred independently of responses to the preceding tone such that suppression could occur when the conditioner tone was outside of the RF of a neuron and thus elicited no response (Calford

and Semple, 1995; Brosch and Schreiner, 1997). As discussed above, cortical RFs are more complex than those displayed by AN fibres, reflecting convergence and integration of excitatory and inhibitory inputs at several levels of the auditory system (Sutter, 2000). It seems that forward suppression in cortical neurons is determined by similar parameters as the RF tuning. This is implied by the fact that all neurons with non-monotonic rate-level functions displayed non-monotonic suppression-level functions, where the strongest suppression was not seen with the highest level conditioners. Similarly, the majority of neurons that displayed monotonic rate-level functions displayed monotonic suppression-level functions (Calford and Semple, 1995). In addition, it was found that varying the probe frequency in 7 neurons that displayed broad or multi-peaked tuning led to suppression of a restricted region of the excitatory RF (Calford and Semple, 1995). The different types of suppression observed could come from two distinct subpopulations of neurons: those that display non-monotonic rate-level functions and those with broad or multi-peaked tuning. As noted, neurons with broad and multi-peaked tuning are mostly found in dorsal A1 (Sutter and Schreiner, 1991) while neurons with non-monotonic rate-level functions are mostly found in central A1 (Heil et al., 1994; Phillips et al., 1994).

Suppression lasted for longer in cortical neurons than in subcortical neurons though there were differences between cat and macaque. In cat, the maximal duration of suppression varied between 40 and 430 ms with a mean of 143 ms (Brosch and Schreiner, 1997), while the median duration of

suppression was 177 ms in anaesthetized (Brosch et al., 1999) and 189 ms in awake macaque (Brosch and Scheich, 2008). While species differences in suppression duration may be involved, it is also possible that shorter intervals between each tone pair in the cat study led to forward suppression being measured in a habituated state (Brosch and Schreiner, 1997; Brosch and Scheich, 2008). Previous studies have shown suppression effects at interpair intervals of over 1 second (Hocherman and Gilat, 1981; Werner-Reiss et al., 2006). The duration of suppression in awake and anaesthetized macaque was similar even though the conditioner tone durations were different, 100 and 30 ms respectively. This led Brosch and Scheich (2008) to postulate that the important factor determining cortical forward suppression was the stimulus onset asynchrony (SOA: the time between tone onsets) rather than the ISI. In support of this idea is the fact that the median suppression duration was less in anaesthetized (77 ms) than in awake (159 ms) macaque if the ISI was used as the measure of interest.

Another difference between cortical forward suppression and that in the AN is that suppression lasted longer in neurons with low CFs than those with high CFs (Brosch and Schreiner, 1997). This parallels psychophysical findings described in section 2.2.1.1 that the duration of forward masking decreased with the frequency of the stimulus (Jesteadt et al., 1982). At subcortical levels, the duration of forward suppression is independent of CF, indicating a cortical basis for this aspect of forward suppression (Brosch and Schreiner, 1997).

In addition to effects on the spike rate response to the probe, suppression can be manifested as an increase in the latency to the first spike as the intensity of the conditioner tone is increased (Brosch and Schreiner, 1997). This latency effect was mainly observed when there was a decrease in the magnitude of the probe response and the SOA was 30 ms though latency changes could occur in the absence of a change in response strength to the probe.

It has been established that cortical forward suppression depends on both the level and frequency of the conditioner tone (Calford and Semple, 1995; Brosch and Schreiner, 1997), but the parameters of the probe tone have been investigated in relatively few studies. It has been found that varying the level of the probe tone can have an effect on forward suppression (Scholl et al., 2008). Suppression of low-level probe tones was stronger than that of high-level probe tones in both the spiking response and in sub-threshold responses measured intracellularly. Facilitation of responses was rarely observed and did not depend on the level or frequency of the conditioner.

In cortex forward suppression is not limited to the frequency receptive field of neurons. Suppression of binaural level-response areas, constructed by systematically varying the level of a CF tone at the two ears, has also been observed (Zhang et al., 2005). The preferred binaural combinations (PBCs) were the binaural stimulus combinations that evoked the top 20% of responses. Suppression of a PBC tone was greatest when the conditioner tone

was in the PBC. Varying the binaural probe tone characteristics showed that tones within the PBC were more robust to suppression; in fact suppression acted to increase the selectivity of a neuron for its preferred binaural tone characteristics. There was a significant correlation between the response to the conditioner tone and the amount of suppression when the total numbers of spikes was averaged across the 30 repetitions of each trial. But, on a trial-by-trial basis there was no significant correlation between the response to the conditioner and the response to the probe, indicating that suppression was not dependent on the response to the conditioner. Reasons for this will be discussed in the summary section below. Non-monotonic suppression patterns were observed for non-monotonic neurons. When the ISI was increased, recovery of responses occurred earliest for tones within the PBC and later for tones outside of the PBC (Nakamoto et al., 2006).

2.2.2.6.2. Backward Suppression

In some cortical neurons responses to a tone can occur after the tone has ended (e.g. Moshitch et al., 2006). In these neurons it is possible that backward suppression occurs, in which the second tone in a two-tone paradigm can suppress these late responses (Brosch et al., 1998). The response to the second tone can overlap with the late response to the first tone and so a method of linear superimposition was used to find statistically significant suppression of the late response. Although no population statistics were shown, backward suppression could be observed up to an ISI of around 77 ms in one example. Slightly longer backward suppression effects have

been described in an EEG study, where a late potential related to the first tone in a two-tone paradigm could be reset if the second tone was presented within 100 ms (Winkler and Naatanen, 1992).

2.2.2.6.3. Forward facilitation

As in the subcortical nuclei, the response to the second of two tones can be facilitated as well as suppressed. Forward facilitation has been observed in anaesthetized cat (Brosch and Schreiner, 2000) and in anaesthetised (Brosch et al., 1999) and awake (Brosch and Scheich, 2008) macaque.

The spectral and temporal characteristics of forward facilitation differed from those of forward suppression. Firstly, facilitation was rarely observed when the spectral characteristics of the two tones were similar. In cat, the greatest amount of facilitation was observed when there was around a 1 octave frequency difference between the first and second tones (Brosch and Schreiner, 2000). In macaque, there seems to be more of a spread of frequency differences that elicited forward facilitation, but these were rarely the same frequencies that elicited suppression; in awake macaque the median difference between tones that evoked suppression and enhancement was 0.92 octaves (Brosch and Scheich, 2008). In cat, the greatest amount of facilitation was observed when the probe tone was 14 dB lower than the first tone (Brosch and Schreiner, 2000). The effect of varying the probe tone frequency and intensity was also investigated in the cat. Generally, facilitation was only observed with probe tones which were in the RF of the neuron and

thus evoked a reliable response. These probe tones were seemingly facilitated by a similar range of first tones, though with probe tones that evoked suboptimal firing, the measure of facilitation was less reliable. The authors also concluded that unlike suppression, facilitation has no backward component (Brosch and Schreiner, 2000).

There are two main differences between suppression and facilitation effects in the temporal domain. Firstly, the dependence of facilitation on the temporal separation of the tones was nonmonotonic with the greatest amount of facilitation occurring at an ISI of around 20 ms in macaque (Brosch et al., 1999; Brosch and Scheich, 2008) and 70 ms in cat (Brosch and Schreiner, 2000). As stated, forward suppression is strongest at the shortest ISIs and decreases monotonically with time (Brosch and Schreiner, 1997; Brosch and Scheich, 2008). A second temporal difference is that facilitation showed a greater dependence on ISI than on SOA, opposite to the trend previously described for suppression in section 2.2.2.6.1. This is suggested by the fact that neurons show greatest facilitation at an ISI of around 20ms in both awake (Brosch and Scheich, 2008) and anaesthetized macaque (Brosch et al., 1999).

Table 2-1. The characteristics of sequential two-tone interactions at various levels of the auditory system.

Structure	Initial suppression for equal level tones	Time course of suppression	Facilitation present?	Effects directly correlated with response history?	Binaural interactions?
AN		Mean 115 ms (Range 40 – 310 ms)	No.	Yes.	No.
AVCN	21%		No.	Yes.	No.
DCN		Mean ~50 ms	Yes.	No. Suppression from outside the RF and delayed suppression.	No.
SOC	50%	Mean 105 ± 20 ms	Not reported.	Not reported.	Yes.
IC	77± 16 %	Mean 271 ms (Range 7 – 1370 ms; Median 73 ms)	Yes.	No. Delayed suppression and facilitation from outside the RF.	Yes.
A1	Around 100% for most neurons	Mean 143 ms (Range 40 – 430 ms)	Yes.	No correlation between conditioner response and suppression.	Not reported but likely.

2.2.2.7. Summary

In this section I have shown how sequential two-tone interactions develop as the auditory system is ascended, from a simple relationship based on the adaptation to the conditioner tone in the AN to a more complex relationship shaped by excitatory and inhibitory modulations in A1. There was also an increase in the time that neurons took to recover from suppression that has lead to the cortex being posited as the neural basis of psychophysical forward masking. However, if one refers to table 2-1, which displays various features of suppression at different levels of the auditory system, the increase in the time course of suppression is far from systematic. One question, then, concerns exactly how different temporal processing is in cortical responses as compared to subcortical. A clearer story comes from a study that used clicks to characterise suppression at various stages of the auditory system (Fitzpatrick et al., 1999). It was found that the recovery to 50% of the control response occurred later (20 ms) in the cortex than in the IC (7 ms) or the AN, AVCN or SOC (2 ms). In the next section I expand on the responses to sequential two-tone stimuli by discussing the responses to repetitive stimuli.

2.2.3. Temporal coding in the auditory system and adaptation

In the previous section I introduced the concept of forward suppression where the response to a tone could be decreased by presenting another tone before it. The forward suppression observed with two tones has important implications for the ability of the auditory system to process sounds that repeat regularly over time (periodic sounds). Sounds that repeat

at a slow rate (less than around 10 per second) are perceived as distinct events (Miller and Taylor, 1948; Symmes et al., 1955). As the rate of presentation is increased to around 10-12 Hz the sounds begin to blur into a single rough percept and then at even faster rates of around 30-50 Hz a tonal percept develops with a pitch corresponding to the repetition rate. The following section begins by considering the means by which neurons in the auditory system process periodic sounds. I then discuss an important feature of the response to repetitive stimuli, adaptation, and how the auditory system may utilise this functionally.

2.2.3.1. Temporal coding in animals

AN fibres respond to trains of successive sounds by synchronising responses to each sound. AN fibres can synchronise responses to click stimuli at rates of up to 1000 Hz (Kiang, 1965) and can synchronise responses to amplitude modulated tones to a similar degree (Joris and Yin, 1992; Langner, 1992). The maximum rate at which neurons can synchronise responses generally decreases as the auditory system is ascended, at least above the level of the CN (Langner, 1992). CN neurons generally display responses to amplitude modulations that are similar to or enhanced relative to those of AN fibres (Langner, 1992). Further, at the level of the IC there is a change in the type of neural code that is used to characterise periodic stimuli that occur at high rates (Schreiner and Langner, 1988; Rees and Palmer, 1989; Langner, 1992). Instead of synchronising to the events within a periodic stimulus, neurons respond to increasing rates by increasing their overall spike rate, so

that the mean firing rate is tuned to the modulation rate. This transformation is initiated in the IC but becomes more developed at the level of the MGB and above.

While the transformation from a synchronisation code to a rate code at higher repetition rates in the MGB and cortex is an interesting feature of periodicity coding, it is somewhat outside the focus of this thesis. The main feature of periodicity coding that impacts on the studies presented within this thesis is that the ability of auditory neurons to synchronise to the stimulus deteriorates as the auditory system is ascended. In a study that directly compared thalamic and cortical responses to AM tones, it was found that cortical neurons did not phase-lock to stimuli with modulation frequencies above 20 Hz (Creutzfeldt et al., 1980). The maximal rate at which cortical neurons can phase-lock differs depending on several factors, for example the species and the state of the animal (anaesthetised or awake) that is used in a study (Oshurkova et al., 2008). It has been shown, for example, that under ketamine anaesthesia no multi-unit clusters responded to click rates greater than 33 Hz, compared to around 80% of multi-unit clusters in the same animal when it was awake (Rennaker et al., 2007). While the main point here is that A1 neurons in general cannot synchronise to high rates, such effects of anaesthesia must be considered when analysing data concerning periodic synchronisation in the auditory cortex.

2.2.3.2. Temporal coding and adaptation in humans: EEG, MEG and *f*MRI studies

The ability of neurons to phase-lock to periodic stimuli is reflected somewhat in the population signals that are accessed using techniques such as EEG and MEG. In general terms, synchronisation ability is indexed by the amount of overlap of the ERP responses and the extent to which components of the ERP are reduced or adapt as the presentation rate is increased. In agreement with the electrophysiological results discussed above, components of the EEG and MEG signals that reflect later stages of auditory processing are more likely to adapt when the stimulus repetition rate is increased. For example, for both MEG and EEG, the amplitude of components of the long-latency response (LLR: >50 ms after stimulus onset (Picton et al., 1974; Carver et al., 2002), which likely originate at the level of the auditory cortex (Naatanen and Winkler, 1999; Carver et al., 2002), decreases as the presentation rate increases (Hari et al., 1982; Lu et al., 1992; Carver et al., 2002; Snyder and Large, 2004). In contrast, there is no significant adaptation of the middle-latency response (MLR: 10-50 ms (Picton et al., 1974; Carver et al., 2002)), sources of which are thought to be subcortical and early cortical (Naatanen and Winkler, 1999), at the rates tested (Snyder and Large, 2004) implying that populations of neurons in subcortical structures can process periodic stimuli at faster rates than can cortical populations of neurons.

While EEG and MEG offer high temporal resolution, the spatial resolution is such that it is difficult to ascribe the observed effects to a precise

location in the brain. In contrast, *f*MRI has a spatial resolution that allows more precise locations to be ascribed. In Heschl's gyrus, the location of the primary auditory cortex, both the percent signal change and the extent of activation of the haemodynamic response exhibit an *increase* as the stimulus presentation rate is increased up to around 5-10 Hz (Binder et al., 1994; Tanaka et al., 2000; Harms and Melcher, 2002). This is likely due to an increase in the integrated neuronal response as the number of stimuli in a given time is increased (Binder et al., 1994). As the rate is increased further, however, both measures begin to *decrease* (Tanaka et al., 2000; Harms and Melcher, 2002). This pattern is observed both in the MGB and the superior temporal gyrus (STG: the location of secondary auditory cortical areas), though the presentation rate at which a decrease begins is higher (20 Hz) in the MGB and lower (2 Hz) in the STG (Harms and Melcher, 2002). In the IC, the signal percent change increases as the rate increases up to the maximal rate tested, which was 35 Hz (Harms and Melcher, 2002). So, *f*MRI provides further evidence for a deterioration in the ability of neurons and neuronal populations to phase-lock to periodic stimuli as the auditory system is ascended.

A further change in the cortical *f*MRI signals that accompanied an increase in rate was observed in the shape of the waveform. At lower rates the wave shape was sustained, however as the rate was increased the wave began to have a more phasic shape, with prominent peaks after the stimulus train onset and offset. In the IC only sustained wave shapes were observed so

there was a systematic shift in the form of the response-rate dependencies from one of amplitude to one of wave shape. Due to the qualitative correlation between the wave shape and the perceptual attributes of the stimuli, it was suggested that the wave shape may be a neural correlate of the perception. Given the idea that population neural activity in the auditory cortex codes the onset and offset of auditory events, the transformation of a sustained to a phasic wave shape may coincide with the point at which individual auditory events fuse to create a single auditory percept (Harms and Melcher, 2002). The notion that changes in response measures can be correlated with changes in perception is returned to in several instances in section 2.3**.**

2.2.3.3. Mismatch Negativity

Another component of the EEG signal that is affected by repetitive stimulation is the mismatch negativity (MMN). The MMN is a component of the auditory event-related potential (ERP), occurring between 100 and 200 ms after stimulus onset, which is elicited by any discriminable change in some repetitive aspect of auditory stimulation (Naatanen et al., 2001). The MMN is generally measured in an oddball paradigm, where the probability of a 'deviant' stimulus being presented is lower than the probability of a 'standard' stimulus being presented. Major sources of the MMN are the primary and secondary auditory cortices, though the exact location may change depending on which sound feature changes (Javitt et al., 1994; Alho, 1995). In addition, MMN is pre-attentive, being observed both during REM

sleep (Atienza and Cantero, 2001) and in comatose patients (Fischer et al., 2000). The MMN has also been shown to exist under ketamine anaesthesia in guinea pigs (Kraus et al., 1994; King et al., 1995). Until recently, it was proposed that a MMN is not elicited for an infrequent sound alone; the infrequent sound must occur within the context of a frequently occurring sound (Naatanen and Winkler, 1999). However, it has been reported that a MMN can be elicited for a 'variant' tone presented after only one 'standard' tone (Jaaskelainen et al., 2004). It seems likely that the MMN is an N1 response (a negative component of the LLR that occurs near to 100 ms after stimulus onset (Carver et al., 2002)) that has been suppressed and delayed by adaptation that is specific to the standard stimulus (Jaaskelainen et al., 2004).

2.2.3.4. *f*MRI-adaptation and models of adaptation

The adaptation of the haemodynamic signal that I described above has been utilised to characterise the nature of neural representations (Grill-Spector et al., 2006). The procedure involves presenting a stimulus repeatedly and measuring the decrease in the *f*MRI signal. Then, a novel stimulus that differs along one dimension is presented. If the underlying neural representation is insensitive to the feature that is changed then the *f*MRI signal will be reduced similar to the reduction produced by the repeated stimulus. Conversely, if the neural representation is sensitive to the transformation, the *f*MRI signal will be similar to the non-adapted level. Thus, inferences can be made about the representation of the two stimuli.

*f*MRI-adaptation has mainly been used in the visual domain, though there is at least one example of its use in audition in the investigation of repetition priming (Bergerbest et al., 2004). Additionally, the inference of neuronal selectivity using this technique is complicated by evidence that, at least in the inferior temporal cortex, adaptation at the single-cell level demonstrated a greater degree of selectivity than the responses (Sawamura et al., 2006). Due to these factors, I will not discuss *f*MRI-adaptation in any more depth, except to describe several models of adaptation that have been posited because of it.

Based on *f*MRI-adaptation, three models of repetition-related reductions in neural activity have been suggested (Grill-Spector et al., 2006). The first is a fatigue model that suggests that repetition leads to an overall reduction in response. In *f*MRI this would relate to a reduction in population activity but with no changes in the pattern of responses across neurons or in the temporal window in which neurons are responding. Secondly, a sharpening model predicts a narrowing of the tuning curve around the preferred stimulus. From a global perspective this would involve some, but not all, of the neurons that were initially active exhibiting a reduction in response with repeated stimulation. Importantly, those neurons which responded optimally to the repeated stimulus would not display adaptation; adaptation would mainly occur in those neurons which coded irrelevant features of the stimulus, leading to a sparser representation and thus a lower *f*MRI signal. Finally, a facilitation model predicts that repetition causes faster

processing of stimuli. So, although the neurons that responded initially will continue to respond, the *f*MRI signal will decrease because the amount of activity integrated over time will decrease. These three models were suggested given results from *f*MRI studies, but they may be applicable to the responses of single neurons. This will be addressed in chapter 6.

2.2.3.5. Adaptation in animal models

Adaptation, or a decrease in the response to repetitive stimuli, was briefly mentioned in the previous section with reference to EEG, MEG and *f*MRI signals. Adaptation of the response to single tones has been observed over a variety of time courses in the AN (Smith and Zwislocki, 1975; Westerman and Smith, 1984; Chimento and Schreiner, 1991; Javel, 1996). Of increasing interest recently, is the way in which neural adaptation at higher levels contributes to various functions of the auditory system. For example, in the IC the coding of sound level has been shown to adapt depending on stimulus statistics (Dean et al., 2005). In a paradigm in which the probability of one sound level being presented was much higher than other sound levels, both single neurons and the population displayed shifts in the rate-level functions such that coding was improved at the most probable sound level. Further, neurons were also shown to adapt to second order stimulus statistics such as variance, in agreement with a previous report (Kvale and Schreiner, 2004). The adaptation to stimulus statistics had several time constants, one of which was fast (160 ms), a finding that correlates well to other forms of adaptation which will be discussed below (Ulanovsky et al., 2004; Dean et al.,

2008). Using a similar paradigm, it has been shown that non-monotonic cortical neurons adapt their rate-level functions in a similar manner, except at high intensities where the rate-level function is relaxed such that the maximal coding occurs at lower levels (Watkins and Barbour, 2008). This suggests that the two neural populations – monotonic and non-monotonic, may serve different functions in their adaptive properties as well as their static properties. Non-monotonic neurons encoded lower sound levels more effectively than monotonic neurons. In contrast, monotonic neurons adapted to higher sound levels while non-monotonic neurons did not. Thus, between the two populations a wide range of sound levels could be encoded both statically and in dynamic environments.

2.2.3.6. Stimulus Specific Adaptation

One type of adaptation that has been linked to sensory memory and the MMN is stimulus-specific adaptation (SSA). SSA has been demonstrated in single units in the cat primary auditory cortex (Ulanovsky et al., 2003). The responses to two tones with frequencies equidistant from the BF of the cell (see figure 2-14a) were measured as the probability of each tone occurring was varied: the probability of occurrence of the standard/deviant tones was either 90/10%, 70/30% or, as a control condition, 50/50%. It was found that the response to the standard tone was decreased compared to the control condition and that the response to the deviant tone was increased relative to the control condition (see figure 2-14b). Further, the magnitude of this trend was greater when the probability of occurrence of the tones was 90/10% than

when it was 70/30%. These effects were also observed when the standard and deviant tones differed by amplitude rather than frequency. The change-specific neural responses that accompany this SSA have led to it being posited as a neural correlate of the MMN (Ulanovsky et al., 2003). The SSA decreased with increasing ISI and was not observed at ISIs above 1 second for single units and 2 seconds for multi-units. It was also not observed at the level of the thalamus, in contradiction to previous reports that SSA existed in the IC (see below for a discussion: Sanes et al., 1998; Malone and Semple, 2001). However, different stimulation paradigms were used in the IC studies and the SSA that was observed in the cortex differed from that reported in the IC in mostly affecting sustained rather than onset responses, existing for longer (present for ISIs >1700 ms) and being highly stimulus specific. This high stimulus specificity was investigated by varying the frequency difference between the standard and deviant tones at values of 0.37, 0.10 and 0.04 (0.53, 0.14 and 0.06 octaves). Effects of stimulus probability could be observed even at the smallest frequency difference which is an order of magnitude smaller than both cortical and AN receptive fields (Ulanovsky et al., 2003). Such hyperacuity over long timescales had been previously demonstrated in multi-unit clusters, where the bandwidth of SSA observed was 0.125 octaves, the lowest frequency difference tested (Condon and Weinberger, 1991). Here it was observed in single units. A further study, using a similar oddball paradigm, demonstrated that SSA operated on several time scales, from hundreds of milliseconds to tens of seconds (Ulanovsky et al., 2004).

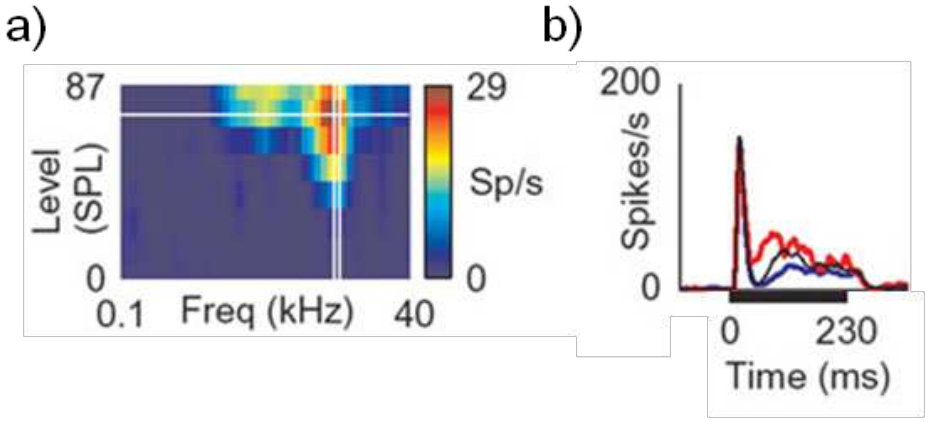


Figure 2-14. Stimulus specific adaptation in A1. a) The receptive field of an A1 neuron showing the level (horizontal white line) and frequencies (vertical white lines) of two tones equidistant from the CF that were used in the oddball stimuli. b) The response to the deviant tone which was presented with a probability of 10% (red line), the standard tone, presented with a 90% probability (blue line) and the control (50 % probability). The responses have been averaged over the conditions where each tone constituted the standard and the deviant. The tones had a frequency difference of 0.14 octaves. Black bar under the ordinate indicates when the tone was presented (adapted from Ulanovsky et al., 2003).

The presence of some type of SSA in the IC is further supported by a study that reported a subpopulation of neurons in IC that responded only when some characteristic of the stimulus changed (Perez-Gonzalez et al., 2005). These neurons represented only 6% (25) of the total recorded population (409) and were found only in lateral, dorsomedial and rostral portions of the IC, areas which receive extensive cortical projections (Herbert et al., 1991; Saldana et al., 1996). SSA within these neurons was rapid, sometimes occurring after the first presentation of a stimulus and the neuron only responded again when a different stimulus was presented. Parameters

which caused SSA included frequency, intensity, and the modulation rate of sinusoidally amplitude and frequency modulated stimuli.

2.3. The processing of tone sequences

In the previous section, I considered the way in which the auditory system responded to repetitive stimuli and how adaptation to repeated stimuli may be an important aspect of change detection. In this section I consider how the auditory system responds to a specific sequence of tones that has been used to investigate how auditory streams are formed. Again, the main impetus for investigating this neurophysiologically comes from effects that have been described psychophysically. In particular, a great deal of psychophysical work on tone sequences has been aimed at investigating how the auditory system organizes the incoming acoustic signal into auditory objects, in this case auditory streams. I will begin by introducing the concept of auditory streaming from a psychophysical perspective. I will then discuss electrophysiological and fMRI investigations of streaming in humans. Finally, I will discuss neurophysiological investigations of streaming in animals.

2.3.1. Introduction to auditory streaming

Auditory streaming refers to the separation of sound components into perceptually distinct auditory objects or streams, each of which can be attended to independently amidst other sound sources. When sequentially presented sound elements are grouped into separate streams this is referred to as stream segregation. Conversely, when sound elements are grouped

together into a single stream the process is referred to as stream integration (Bregman, 1990).

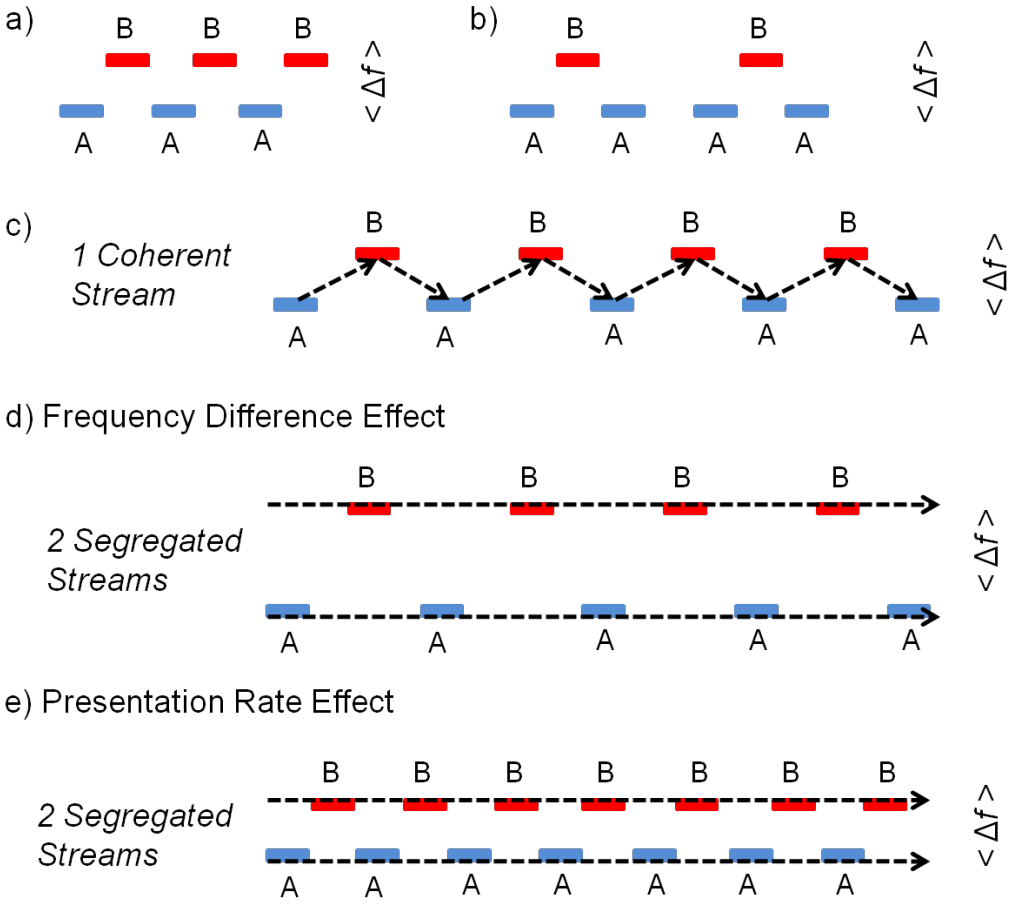


Figure 2-15. a) and b) The stimuli used to investigate streaming consist of two interleaved, non-overlapping sequences of tones at different frequencies. c) With a small frequency difference a single stream is likely to be perceived. d) With a larger frequency difference it is more likely that two streams are perceived. e) At a faster presentation rate it is more likely that two streams will be perceived. Further explanation in text.

Auditory streaming has been demonstrated in the laboratory using stimuli such as those shown in figure 2-15a and b. The stimuli consist of interleaved non-overlapping sequences of tones of different frequencies. When the difference in frequency between the tones is small, the sequence is

perceived as a single coherent stream with a continuous up-and-down movement in pitch, as shown in figure 2-15c. When the frequency difference (FD) is larger, as shown in figure 2-15d, then the sequence is perceived as two segregated streams of unrelated, interrupted tones of different frequencies. The transition point between these percepts is known as the trill threshold (Miller and Heise, 1950). The ABA_ABA_ stimulus of figure 2-15b adds a rhythm cue to the percept so that a 'galloping' rhythm is perceived with one stream that is lost when two streams are perceived. Using this stimulus it was demonstrated that the streaming percept depends not only on the frequency difference but also on the rate at which the tones are presented (van Noorden, 1975). When the presentation rate (PR) is faster, subjects tend to report a segregated percept, as shown in figure 2-15e.

2.3.1.1. Perceptual boundaries

In essence, then, the perception of the streaming stimulus as one or two auditory streams can be changed by simply altering two parameters – the FD and the PR. Systematically varying the two parameters has highlighted three perceptual regions, as shown in figure 2-16. As mentioned, when the frequency difference is small the sequence is always perceived as a single coherent stream, as indicated by the light blue region in figure 2-16. As the frequency difference is increased, it becomes possible to selectively attend to two distinct streams, labelled as the fission boundary in figure 2-16. The pink region of figure 2-16 represents the region of perceptual space where it is only possible to segregate the sequence into two separate streams. The

temporal coherence boundary delineates this region from the white 'ambiguous' region, where it is possible to hear either one or two streams.

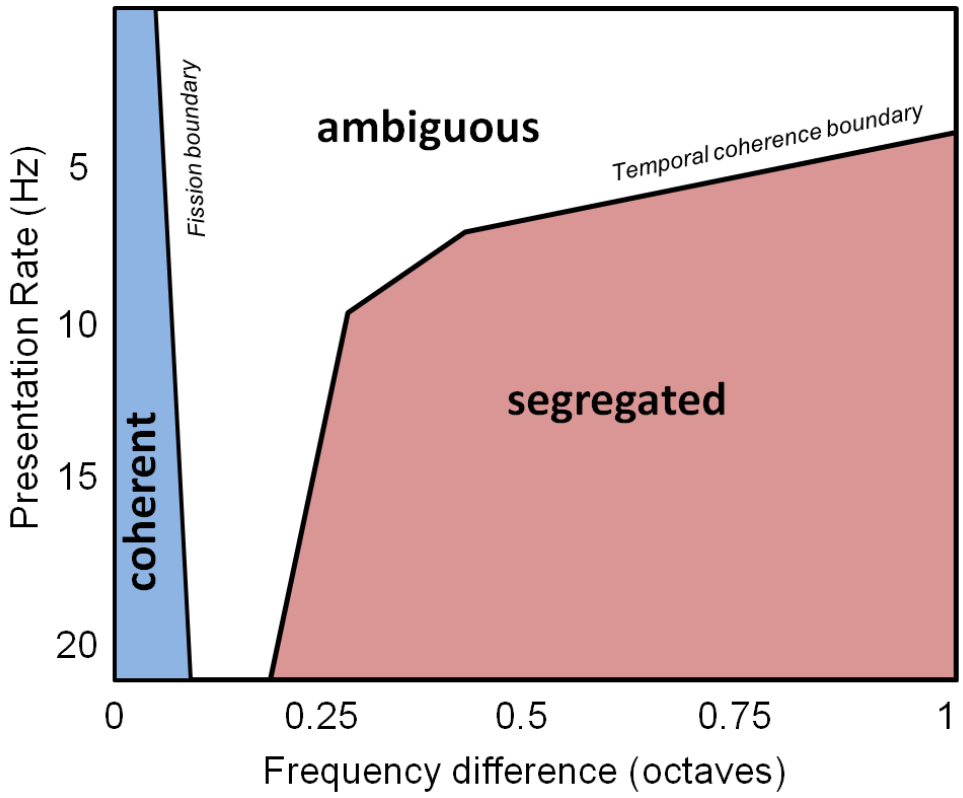


Figure 2-16. The perceptual boundaries of streaming (adapted from McAdams and Bregman, 1979).

2.3.1.2. Build-up of streaming and bistability

Another aspect of streaming is that, initially, a single coherent stream is always perceived: the percept of two segregated streams takes time to build up (Bregman and Pinker, 1978; Anstis and Saida, 1985; Carlyon et al., 2001). Further, within the ambiguous region the percept evoked by a sequence presented over the course of several seconds or minutes can spontaneously switch from coherent to segregated or vice versa (Pressnitzer and Hupe, 2006). This auditory bistability to tone sequences shares several

features with a visual bistability to plaids, indicating that visual and auditory perceptual organization could develop from functionally similar neural mechanisms (Pressnitzer and Hupe, 2006). An important caveat suggested by Pressnitzer and Hupe (2006) is that the psychophysically observed build-up of streaming may be an artefact of averaging across many subjects.

2.3.1.3. The role of attention

Streaming reflects a number of different processes, some of which have been characterized as automatic, stimulus-driven and not requiring attention and some of which are schema-based, requiring attention and prior knowledge (Bregman, 1990). The precise role of higher order processes such as attention has been difficult to characterize. It has been known from early on in streaming investigations that attention can affect streaming judgments (van Noorden, 1975). Subsequent investigations into the role of attention have given seemingly contrasting results.

Evidence for the necessity of attention in streaming comes from a study where it was found that the build up of streaming in normal subjects was greatly reduced or absent when they attended to a competing task in the contralateral ear (Carlyon et al., 2001). This was corroborated by a further study which demonstrated that more streaming occurred when subjects attended to the auditory stimulus than when they attended to a visual stimulus (Carlyon et al., 2003). However, a subsequent study challenged the notion that attention was necessary for streaming to occur by showing that a sequence that would elicit a two stream percept disrupted performance less

than one which would elicit a single stream percept in a visual memory task (Macken et al. 2003). In this study the subjects were asked to ignore the sound stimulus and due to the difficulty of the visual task it is unlikely that they could attend to the sound stimulus for long enough for build-up to take place (Snyder and Alain, 2007). These results could be reconciled if one considers that two distinct processes of streaming are being indexed: the build-up of streaming (Carlyon et al., 2001) and the segregation of streams (Macken et al., 2003). Alternatively, it is possible that switching attention in the dichotic listening task could reset the build-up of streaming (Carlyon, 2004; Cusack et al., 2004), consistent with studies that have shown resetting due to abrupt changes of the stimulus pattern (Anstis and Saida, 1985; Rogers and Bregman, 1993; Cusack et al., 2004) .

2.3.1.4. Theories of streaming

A popular theory of streaming that has led to several computational models (e.g. Beauvois and Meddis, 1996; McCabe and Denham, 1997) posits that segregation of streams occurs when the A and B tones excite different peripheral tonotopic channels (van Noorden, 1975). Evidence for this theory, further to that already discussed, comes from the finding that the percept can be biased towards segregation by presenting an induction sequence of tones at either the A or B frequency before the AB sequence is presented (Rogers and Bregman, 1993; Beauvois and Meddis, 1996). Several psychophysical studies have challenged the peripheral channel theory. Firstly, sounds that excite the same peripheral filter can elicit a percept of segregated streams

(Vliegen et al., 1999; Vliegen and Oxenham, 1999; Grimault et al., 2000). Secondly, the percept of segregated streams can be elicited by sequences on the basis of temporal cues alone, even though the sequences have identical long-term power spectra. Examples of such stimuli are amplitude-modulated bursts of white noise (Grimault et al., 2002) and harmonic complexes differing only by the phase relationship of their harmonics (Roberts et al., 2002). Thirdly, a percept of integration can be elicited with sounds presented diotically (same stimulus at both ears) and thus exciting distinct peripheral channels (Deutsch, 1974; Rogers and Bregman, 1993).

The challenging evidence has led to a call for a re-evaluation of the peripheral channelling theory in terms of a functional separation of neural populations, an important example of which is tonotopy. The population theory of streaming posits that stream segregation occurs whenever sounds excite distinct neural populations (Micheyl et al., 2007). These populations could be located peripherally and be sensitive to differences in frequency, or they could be more centrally based and sensitive to either pitch or temporal envelope. For example, neurons with non-monotonic rate-level functions that are essentially tuned to intensity and are found in the CN and above could mediate streaming based on level differences (van Noorden, 1975).

In the following sections I discuss physiological investigations of streaming that provide evidence for the population theory. Firstly, I will briefly describe how streaming was shown to exist behaviourally in animals and then the electrophysiological investigations that followed. I finish the

streaming section by considering physiological studies of EEG, MEG and fMRI in humans.

2.3.2. Investigations of streaming in animals

Using indirect behavioural measures, it has been shown that auditory streaming is experienced by non-human species as diverse as European starlings (Hulse et al., 1997; MacDougall-Shackleton et al., 1998), goldfish (Fay, 1998) and macaque (Izumi, 2002). This suggests that streaming is a process that is utilized by a wide range of species in various environments and opens up the possibility that the neural mechanisms that underlie streaming in humans may be similar to those that underlie streaming in other species.

The first demonstration of streaming in a non-human species was in the European starling (MacDougall-Shackleton et al., 1998). In this study, starlings were trained to recognize the temporal pattern of a stimulus and then to make a different response depending on the pattern. Specifically, the training stimulus could either be the repeating sequence AAA_, A_A_ or _A__, where A was a tone at a single frequency. The birds were trained to respond differently to the first pattern (peck left key) than to the latter two patterns (peck right key). The test stimulus was a repeating ABA_ pattern as used in streaming studies. Varying the frequency difference caused the starlings to respond differently. When the frequency difference was low and the sequence would be perceived as coherent, the response was more likely to be a left key peck. In contrast, when the frequency difference was increased the birds were more likely to respond with a right key peck, indicating that they

perceived either the segregated A_A_ or _B__ sequences. This demonstration of streaming in a non-human species highlighted the possibility of using electrophysiology in animal models to investigate the neural response to streaming stimuli. I address this approach in the following section.

2.3.2.1. Physiological investigations of streaming

2.3.2.1.1. The parameter space

The first studies to investigate the neural mechanisms of stream segregation did so by recording ensemble neural activity (multi-unit recordings and current source density) from A1 neurons in awake macaques (Fishman et al., 2001; Fishman et al., 2004). The frequency of the 'A' tone was set to be at the best frequency of the recording site, as defined using the frequency response curve (displayed as an inset in figure 2-17). The frequency of the 'B' tone was positioned away from the BF, differing from the 'A' tone by an amount which ranged from 10 to 50% (1.65 – 7.00 semitones) of the 'A' tone frequency. The presentation rate of the tones was varied between 5 and 40Hz. In the first study (Fishman et al., 2001), the tone sequences were presented in 490ms epochs with 992ms onset-to-onset intervals. This approach had the problem that streaming would not have built up in this short time and was remedied in Fishman et al. (2004), where the streaming stimulus was presented for 5 seconds before data collection began.

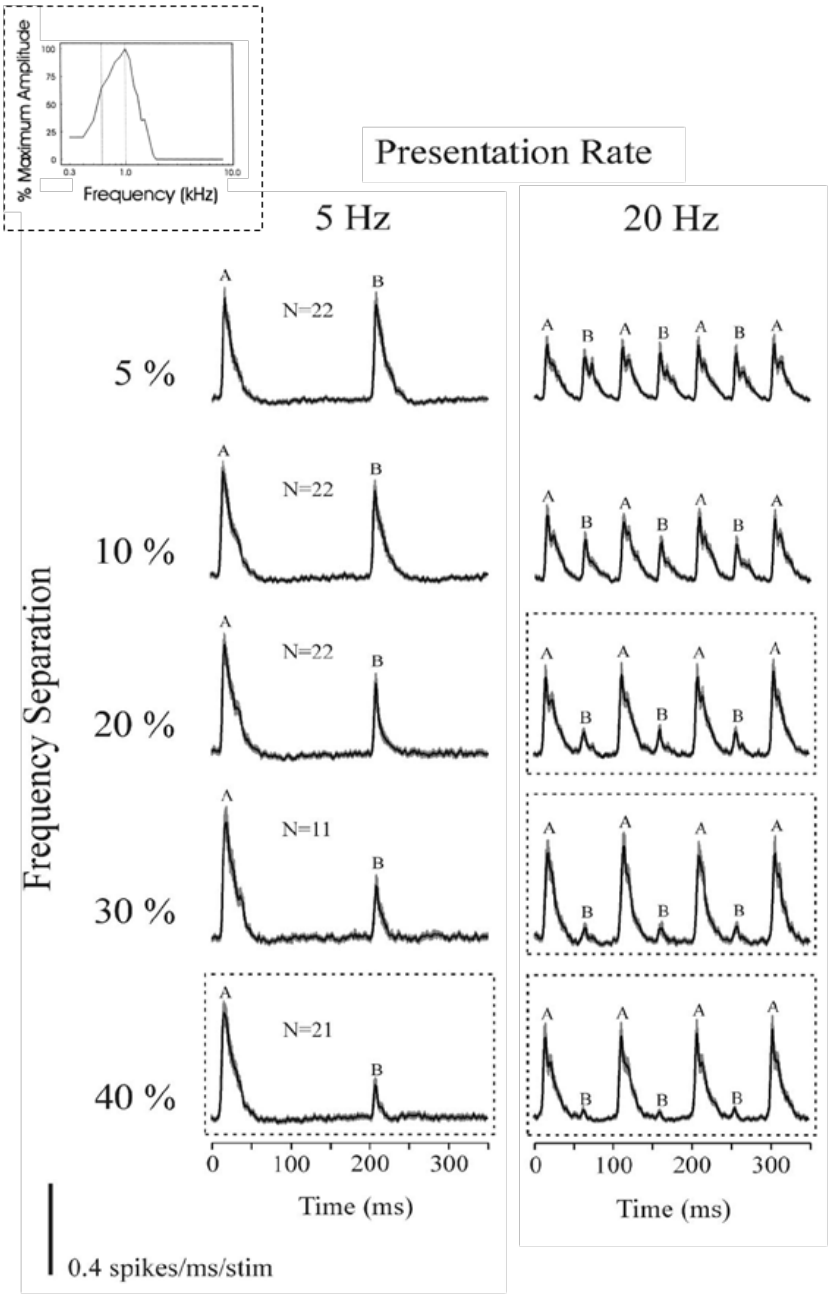


Figure 2-17. PSTHs of multi-unit spike activity evoked by alternating tone sequences averaged across A1 electrode penetrations with a 25 ms tone duration. PSTHs under the different FD and PR conditions are arranged in columns and rows, as indicated. Responses to 'A' and 'B' tones are labelled. Dashed boxes enclose PSTHs in which the 'B' tone response amplitudes are less than -9 dB relative to the amplitudes of 'A' tone responses. The frequency response area of a unit is displayed as an inset, with the frequency of the 'A' tone and the 40% FD 'B' tone indicated by vertical dashed lines (adapted from Fishman et al., 2001; Fishman et al., 2004)

Figure 2-17 shows the effect of increasing the frequency difference on the response to the 'A' and 'B' tones at two presentation rates. As the frequency difference was increased, the response to the 'B' tone decreased.

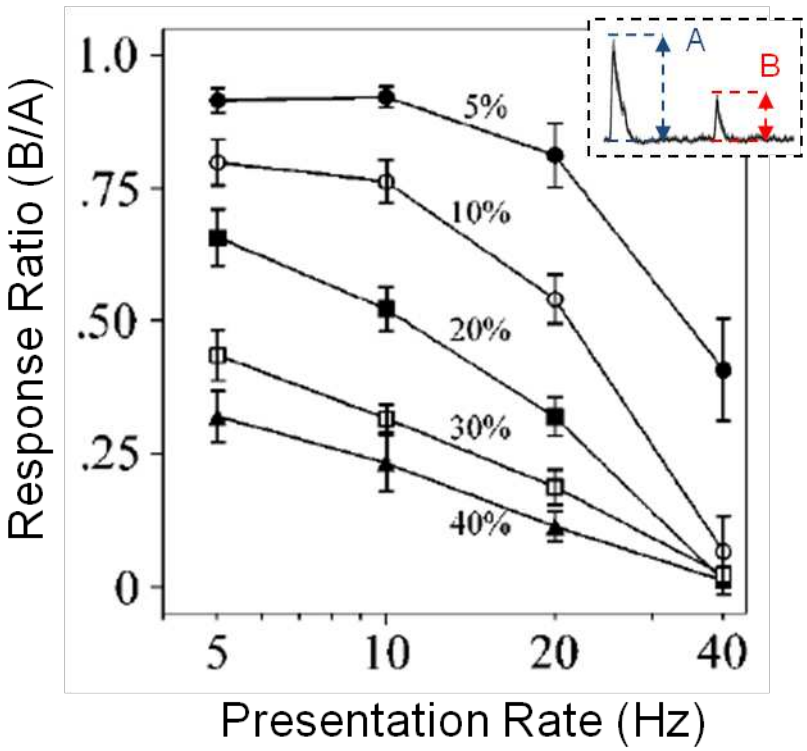


Figure 2-18. The response ratio as a function of presentation rate with the frequency difference as the parameter. The inset shows a schematic of the way in which the response ratio was calculated (adapted from Fishman et al., 2004)

This is explained by the 'B' tone dropping out of the frequency response area of the recording site. More difficult to explain is the greater decrease in response to the 'B' tone than the 'A' tone as the presentation rate is increased (compare the relative responses at the two PRs shown in figure 2-17). This was characterised by measuring the amplitude of the average response to both tones for each condition and then computing a ratio of the 'B' response to the 'A' response. This response ratio is displayed as a function

of presentation rate, with frequency difference as the parameter, in figure 2-18. It is obvious that as the PR is increased at all frequency differences, the response ratio decreases, indicating a larger decrease in 'B' response than 'A' response. The reason for this greater suppression of the 'B' tone is likely due to forward suppression (Fishman et al., 2001; Bee and Klump, 2004; Fishman et al., 2004). For example, it has been shown that the response to the 'B' tone decreases to a greater extent when the FD is smaller, similar to cortical forward suppression which is greatest when the conditioner and probe tones have a similar frequency (Bee and Klump, 2004).

Taking the measure of interest as the response ratio, the dependence of neurophysiological responses on PR and FD parallel the dependence of streaming on these parameters, as measured in psychophysical investigations (Fishman et al., 2001). In conditions where a coherent stream would be perceived psychophysically, for example when the FD between the tones is small, the response to both tones is similar indicating that a single neural population is responding to both tones. In conditions where stream segregation would be observed psychophysically, for example when the frequency difference is large, the neural population responds only to the 'A' tone which is set at its BF. Due to the tonotopic organisation of A1, another spatially distinct neural population, with a BF closer to the 'B' tone frequency would respond preferentially to the 'B' tone, with little or no response to the 'A' tone. This led to a theory of streaming which posits that stream segregation occurs when distinct neural populations respond to each stream

separately, with a single stream being perceived when overlapping neural populations respond to both ‘A’ and ‘B’ tones. This theory is similar to the peripheral channelling theory as it depends on tonotopic separation within the auditory system, the difference being that it applies to the central auditory system.

Other studies have replicated these findings, and thus provided support for the neural population theory of streaming, in the mustached bat (Kanwal et al., 2003) and the European starling (Bee and Klump, 2004, 2005). Studies in which the temporal parameters of the stimuli, for example tone duration and PR, have been independently varied have shown that the inter-stimulus interval (the time between a tone offset and a successive tone onset) is responsible for the observed effects rather the PR (Fishman et al., 2004; Bee and Klump, 2005). This is consistent with psychophysical results, though there is some debate as to whether the ISI within or between streams is responsible (Bregman et al., 2000).

2.3.2.1.2. The build-up of streaming

Although these studies provide evidence for the neural population theory, it is difficult to relate neurophysiological results from animals directly to the psychophysical results from humans. A study in which a nearly identical paradigm was used in both a human psychophysical and a macaque neurophysiological experiment demonstrated that the build-up of streaming could be accounted for physiologically (Micheyl et al., 2005). The only difference between the two paradigms was that for the physiological

experiment the frequency of the ‘A’ tone in an ABA_ sequence was set to the BF of the unit that was being measured, and the ‘B’ tones varied by some frequency difference from the ‘A’ tone. Sequences were presented for 10 seconds in both cases. The human subjects indicated whether they perceived one or two streams while the macaques performed an auditory discrimination task to ensure that they attended to the stimuli, while the responses from single units in A1 were recorded.

Figure 2-19 shows the response of cortical neurons to the ABA_ABA_ stimulus at various frequency differences. The response to both the ‘A’ tone and the ‘B’ tone in the triplet adapts as the sequence elapses. Though the response to the ‘B’ tone adapts differentially as the frequency difference increases, the difference between the tone responses (as used by Bee and Klump, 2004, 2005) is insufficient to account for the build-up of stream segregation (Micheyl et al., 2007). However, by taking account of the variability of the response, using a method based on signal detection theory, it is possible to fit the data from the psychophysical study on the build-up of stream segregation (Micheyl et al., 2005).

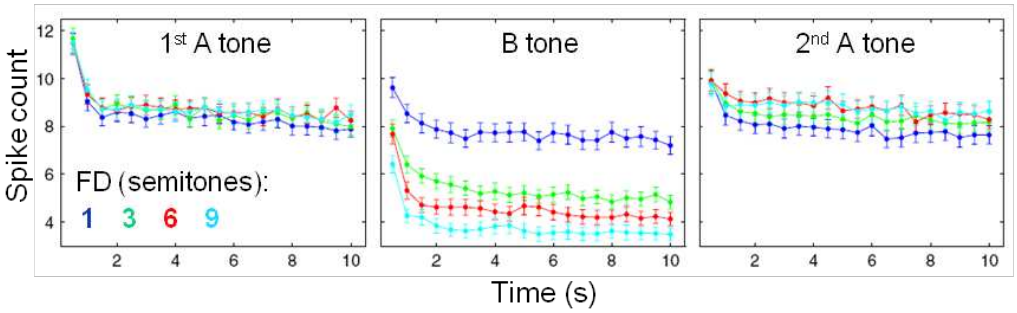


Figure 2-19. The response over time to each tone in the ABA_ triplet (adapted from Micheyl et al., 2007)

In essence, the signal detection model considers the spike count distributions in response to a 'B' tone at a given time in a certain condition. A threshold can be set, above which the 'B' tone is considered to be present in the response and below which the 'B' tone is considered to be absent. By using the spike count population distributions a probability of the 'B' tone being represented can be calculated. Both tones are represented if the 'B' tone is represented and thus the percept would be expected to be coherent. The reciprocal of this probability can be thought of as the probability that two streams are represented and is plotted against the psychophysical data in figure 2-20. As can be seen, the important aspects of build-up are accounted for, namely that segregation is more likely to be reported as time elapses and as the frequency difference increases.

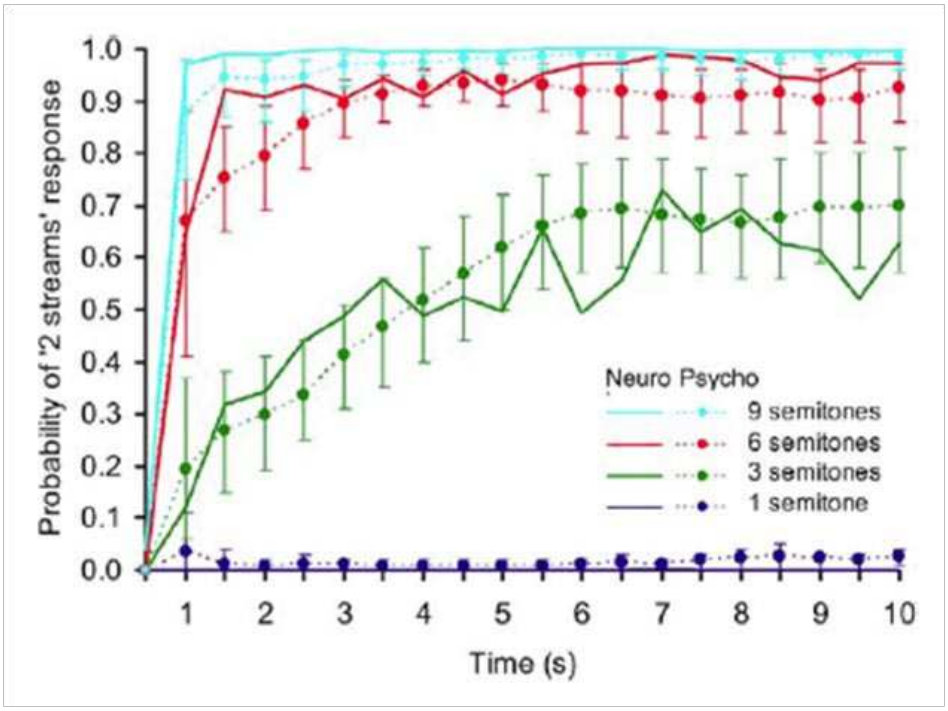


Figure 2-20. Comparison of psychometric and neurometric curves (reprinted from Micheyl et al., 2005)

Using a similar model based on signal detection theory, the build-up of stream segregation has also been accounted for in the cortex of the starling (Micheyl et al., 2008) and in the ventral cochlear nucleus of the anaesthetized guinea pig (Pressnitzer et al., 2008). In the VCN it was shown that chopper cell responses could to some extent predict the psychophysically observed build-up of streaming. However, using this analysis technique, primary-like units did not exhibit build-up. This study highlighted both the possibility that neural circuits responsible for streaming could be found at early auditory processing stations and that responses to streaming stimuli could be observed in anaesthetized preparations.

2.3.3. Investigations of streaming in humans

Both MEG (Gutschalk et al., 2005) and EEG (Snyder et al 2006) recordings have demonstrated an increase in the magnitude of the time-locked response to the 'B' tone as the frequency difference increases. Both MEG and EEG signals probably reflect activity from multiple non-primary areas of the auditory cortex and so differ in location from the animal physiological data (Micheyl et al., 2007). A further difference between these studies is that single BF sites were investigated physiologically while EEG and MEG data reflect more global activity from a variety of BF sites. One problem with these studies is that the effects may reflect the change in stimulus rather than a change in perception *per se*, even though the changes were statistically correlated with a psychophysical measure of streaming in each study (Micheyl et al., 2007). To take account of this, in a further MEG study a single stimulus from within the ambiguous region was presented and the subjects were asked to report whether they perceived one or two streams. It was found that the time-locked response to the B tone increased when two streams were reported. Although the effect was smaller than in the previous study it was demonstrated that this effect was not simply due to stimulus manipulations and so must have been influenced by perceived changes (Gutschalk et al., 2005).

Studies of streaming using *f*MRI have given contradictory results. One study found no changes in the haemodynamic signal from the auditory cortex when the streaming perception changed (Cusack, 2005). Instead, increased

activity was found for a two-stream percept in the intraparietal sulcus, an extra-auditory cortical region that has been implicated in feature binding in the visual domain and cross-modal integration. Conversely, another study did find changes in the *f*MRI activation of auditory cortex that correlated with the streaming perception (Wilson et al., 2007). With small FDs the sequence was perceived as a single stream occurring at the rate of the sequence and the *f*MRI activation in the auditory cortex was phasic. At higher FDs the sequence was perceived as two streams occurring at half the rate of the sequence, the *f*MRI activation became more sustained and the extent and magnitude of the activation increased. Interestingly, the shift from a phasic to a sustained response is comparable to the effect of decreasing the rate of a stimulus on the *f*MRI signal (see section 2.2.3.2). The reason for the discrepancy in the results of Cusack (2005) and Wilson et al. (2007) is unclear but may be due to the use of different streaming stimuli: ABA_ABA_ in the former and ABAB in the latter (Micheyl et al., 2007).

A recent study found that both MEG and *f*MRI signals correlated with the streaming perception even with stimuli where spectral cues were absent (Gutschalk et al., 2007), such as those mentioned in section 2.3.1.4. An increase in the difference between the fundamental frequencies of sequences of harmonic complex tones with identical spectral envelopes led to an increase in stream segregation. This correlated with an increase in both the MEG time-locked response to the 'B' tone and an increase in the cortical *f*MRI signal, which also became more sustained.

2.4. Mechanisms of cortical suppression

I have described the features of the cortical processing of tones over several timescales. In this section I focus on the mechanisms that may be responsible for both cortical forward suppression and stimulus-specific adaptation. These mechanisms are may also contribute to the physiological responses to streaming stimuli that I have described, as forward suppression is undoubtedly involved in the responses to tone sequences.

Several possible contributors to cortical suppression have been suggested and may be applicable to both short-term forward suppression and the long-term stimulus-specific adaptation that has been demonstrated (Ulanovsky et al. 2003; Condon and Wienberger 1991). Suppression mechanisms can be divided into two classes (Gollisch and Herz, 2004): 1) those mechanisms that operate at the output of a neuron, for example the activation of voltage-dependent conductances (Sanchez-Vives et al., 2000a, b); 2) those mechanisms that operate at the input to the neuron, for example synaptic depression (Abbott et al., 1997; Tsodyks and Markram, 1997) or inhibition (Zhang et al., 2003). While forward suppression could be inherited from the thalamus or earlier auditory processing stations it is likely that SSA originates from thalamocortical or corticocortical processes as it is not observed in the thalamus (Ulanovsky et al., 2003; Ulanovsky et al., 2004). Further, the long time course of SSA probably implicates corticocortical rather than thalamocortical generation. A thalamocortical or corticocortical generation of cortical forward suppression is supported by the fact that

thalamic neurons can follow click trains at several tens of hertz or higher, while cortical neurons can only follow click trains up to around 15-20 Hz (Creutzfeldt et al., 1980). Also, although some thalamic neurons demonstrated forward suppression with a time course similar to that observed in the cortex, the population of thalamic units recovered much more quickly than cortical cells (Wehr and Zador, 2005). This indicates that some temporal information is lost or recoded as the signal is passed from the thalamus to the cortex.

Intrinsic intracellular mechanisms such as post-discharge adaptation are unlikely mechanisms for either cortical forward suppression or SSA. The stimulus-specific nature of SSA precludes those mechanisms that operate at the output of the neuron and points strongly at mechanisms that operate at the input to the neuron (Ulanovsky et al., 2004). In cortical forward suppression discharges to the probe occurred independently of discharges to the conditioner tone, indicating that even when the conditioner tone was in the neuron's receptive field, forward suppression was not a post-discharge adaptation phenomenon.

A third possibility is that cortical forward suppression and SSA arise at the circuit level, from the action of either pre- or postsynaptic inputs driving a neuron's response. Historically, the most popular mechanism to explain the extended time course of cortical forward suppression was postsynaptic GABAergic ($GABA_A$) inhibition (Calford and Semple, 1995; Brosch and Schreiner, 1997). In addition, stimulus-specific inhibitory mechanisms have

been suggested to account for the SSA observed in A1 (Ulanovsky et al., 2004). An analogue of SSA has been demonstrated in a network of cortical neurons that developed *ex vivo* (Eytan et al., 2003). By stimulating two different sites, one at a fast rate (standard) and one at a slow rate (deviant), it was shown that responses to the standard decreased and those to the deviant increased relative to the initial response at each site. The response to the deviant only increased when accompanied by the frequent stimulation of the standard and was reduced markedly with the addition of bicuculline, a $GABA_A$ antagonist. This led the authors to postulate that facilitation of the response at the deviant site could be explained by depression of both the excitatory and inhibitory synapses stemming from the frequently stimulated site (Eytan et al., 2003). The fact that the excitatory pathway to the rarely stimulated site would be relatively unaffected, coupled with a general depression of inhibitory synapses in the network, could lead to an increased response to the deviant. This idea is supported by the effect of the blockade of GABAergic inhibitory transmission to reduce the response to the deviant. So, the facilitated response to the deviant occurs due to an imbalance of excitation and inhibition, consistent with intracellular studies in A1 that have shown a balanced inhibition and excitation with the inhibitory conductance occurring a short time after the excitatory one (Wehr and Zador, 2003; Zhang et al., 2003).

Recent evidence has demonstrated that previous reports of long-lasting inhibitory conductances (Volkov and Galazyuk, 1992; Tan et al., 2004)

which could account for the long time course of cortical forward suppression were probably due to the use of barbiturate anaesthesia (Wehr and Zador, 2005). Direct measurement of synaptic conductances to click stimuli, under ketamine anaesthesia, have shown that the inhibitory conductance only lasted for 50-100 ms, while the suppression of spike responses and synaptic input lasted for several hundred milliseconds (Wehr and Zador, 2005). This suggests that postsynaptic GABAergic inhibition only contributes to forward suppression up to around 100 ms and probably has little effect in the generation of SSA, which only occurs after hundreds of milliseconds or longer (Ulanovsky et al., 2004).

Given the evidence, the most likely candidate for both SSA and cortical forward suppression after around 100 ms is synaptic depression. Synaptic depression of thalamocortical synapses has been shown to be involved in suppression in somatosensory and olfactory cortices (Chung et al., 2002; Best and Wilson, 2004). Also, evidence from the somatosensory and visual cortices shows that recovery from depression at corticocortical synapses is best described by several time constants between hundreds of milliseconds and tens of seconds (Tsodyks and Markram, 1997; Varela et al., 1997; Markram et al., 1998), in accordance with those noted in A1 (Ulanovsky et al., 2004). In addition, in the visual cortex, suppression seems to be more consistent with thalamocortical synaptic depression than with inhibition (Carandini et al., 2002; Freeman et al., 2002). Computational models of the auditory cortex have shown that synaptic depression can account for forward suppression

and the low-pass characteristics of cortical responses to repetitive stimuli (Eggermont, 1999; Denham, 2001). Synaptic depression could act as a kind of cortical gain control that operates by balancing contributions from slowly and rapidly firing afferents (Abbott et al., 1997). This would reduce the chances of slowly firing afferents being masked by random fluctuations of high-firing afferents thus increasing the sensitivity of neurons to small rate changes.

2.5. Summary

This summary is organised into a series of sections each of which consider a particular question that I will attempt to address in this thesis.

2.5.1. Does cortical forward suppression contribute to SSA?

In section 2.4 I outlined the possibility that one of the mechanisms that shaped both cortical forward suppression and SSA could be synaptic depression. If the two are indeed connected by a similar mechanism, one question would concern whether they were also similar in other respects. One way of examining whether forward suppression and SSA are linked is to investigate whether they display the same features. One of the defining features of SSA is that the decrease in response to the standard stimulus is highly specific to the features of that stimulus (Condon and Weinberger, 1991; Ulanovsky et al., 2003). Cortical forward suppression has been shown to be frequency-specific in that the magnitude and duration of suppression that is observed differs depending on the characteristics of the conditioner tones (Calford and Semple, 1995; Brosch and Schreiner, 1997). However, forward suppression has generally only been investigated with the probe tone at the

CF of a neuron. Studies in IC (Malone and Semple, 2001) and cortex (Calford and Semple, 1995) have shown effects of varying the probe tone frequency, but have not investigated the effects systematically. Previous studies, on binaural levels (Zhang et al., 2005; Nakamoto et al., 2006) and intensity (Scholl et al., 2008) for example, have shown that varying the characteristics of the probe tone can have marked effects on the suppression pattern that is observed. So, a pertinent question is whether the suppression pattern will change when the frequency of the probe tone is varied. This will help to disambiguate CF-specific suppression from the frequency-specific adaptation that has been observed over long time scales. This question is addressed in chapter 4.

2.5.2. Are the responses of neurons in the anaesthetised guinea pig consistent with those in awake preparations?

When I began this PhD, there was no published literature in which the streaming stimulus had been presented to an anaesthetised animal and the responses measured. In order to investigate the possible mechanisms by which the neural responses to streaming stimuli were created, it was first necessary to confirm that those responses existed in the anaesthetised preparation. In addition, there were also several gaps in the knowledge about the physiological responses to streaming stimuli. Although the effect of varying the FD and PR parameters had been extensively investigated (Fishman et al., 2001; Bee and Klump, 2004; Fishman et al., 2004; Bee and Klump, 2005), the responses were from multi-units or neural ensembles so the

question of whether the same responses would be observed at the level of the single unit remained. Also, although a model based on signal detection theory had been used to account for the build-up of streaming (Micheyl et al., 2007; Micheyl et al., 2008; Pressnitzer et al., 2008), it had not been used to account for the changes in perception as the FD and PR parameters were varied. I address these issues in chapter 5.

2.5.3. How does forward suppression develop into the neural responses observed using streaming stimuli?

It is generally accepted that forward suppression plays some role in the neural responses to the tone sequences that constitute the streaming stimulus (Fishman et al., 2001; Bee and Klump, 2004; Fishman et al., 2004; Bee and Klump, 2005; Micheyl et al., 2007). For example, it has been shown that the average response to the 'B' tone decreased as the FD decreased, consistent with a greater magnitude of forward suppression of tones with similar frequencies (Bee and Klump, 2004). However, the way in which forward suppression builds up over a sequence of tones on the time scale of the streaming stimulus has not been investigated. There are many cortical studies aimed at characterising the interactions between two sequential tones both over short timescales (100s of ms - Calford and Semple, 1995; Wehr and Zador, 2005; Zhang et al., 2005; Brosch and Scheich, 2008; Scholl et al., 2008) and long time scales (seconds to minutes - Condon and Weinberger, 1991; Ulanovsky et al., 2003; Ulanovsky et al., 2004). There is little knowledge of the effects of suppression and adaptation in between

these two time scales. Investigations of suppression over the course of several tones could provide hints as to how both the responses to streaming stimuli and SSA develop. This is investigated in chapter 6.

3. Materials and Methods

3.1. Subjects

Neurophysiological recordings were performed in tricolour pigmented guinea pigs (*Cavia porcellus*).

3.2. Surgical procedure

Neurophysiological recordings were performed in a non-recovery experiment and were preceded by a preparatory surgical procedure. This consisted of anaesthetising the animal, placing it into a stereotaxic head frame and exposing the primary auditory cortex. Anaesthesia was initially induced by an intra-peritoneal injection of urethane (0.9 – 1.3 $g.kg^{-1}$ in a 20% solution, Sigma). A pre-medication of 0.2ml Atropine Sulphate (600 $\mu g.ml^{-1}$ Phoenix Pharma) was administered subcutaneously to suppress bronchial secretions. Supplementary analgesia was maintained using intra-muscular injections of 0.2ml Hypnorm (Fentanyl citrate 0.315 $mg.ml^{-1}$, fluanisone 10 $mg.ml^{-1}$, Janssen), whenever a forepaw withdrawal reflex could be elicited. The withdrawal reflex was checked every 30 minutes in order to assess the depth of anaesthesia. Once surgical anaesthesia had been established the animal was moved to a sound attenuated chamber (Whittingham Acoustics Ltd.) and its throat, ears and head were shaved using animal clippers (Sterling 2 plus, Wahl). The trachea was then cannulated using a piece of polythene tubing with an external diameter of 2.4 mm and an internal diameter of 2 mm (Portex). The tragi were excised bilaterally to allow access to the external

auditory meatus of each ear. Any blockages of the ear canal, for example wax, were then removed using forceps.

The animal was placed into a stereotaxic head frame and hollow Perspex specula were inserted into the ear canals. The animal's head was manoeuvred until the tympanic membrane for each of the ears could be visualised with a microscope directed down the end of each speculum. The animal was artificially respired by connecting the tracheal cannula to a respiratory pump supplying 100% oxygen (BOC) mixed with air. Both the respiratory rate and the end tidal CO_2 were monitored at the cannula and maintained in the normal range of 28 – 38 mm Hg. The core temperature of the animal was monitored with a rectal thermometer inserted into the rectum using petroleum jelly (Vaseline) as a lubricant. The thermometer connected to an external temperature control unit which operated using feedback: the amount of current to a thermal blanket (Harvard Apparatus) was increased if the core body temperature dropped below 38°C and *vice versa*. The cardiac function of the animal was monitored using an electrocardiogram recorded via electrodes inserted into the skin on either side of the thorax.

To expose the skull, a midline sagittal incision was made along the scalp, the skin was retracted and the periostium and temporalis muscles were removed. To equalise the middle ear pressure with the surrounding air a length of 0.5 mm diameter polythene tubing was inserted into a hole made in the auditory bulla on each side and sealed with petroleum jelly (Vaseline). In

order to increase the stability of the recording, a small hole was made in the connective tissue above the foramen magnum, thus releasing the pressure on the cerebro-spinal fluid.

To expose the primary auditory cortex, a craniotomy was performed on the right side using a dental drill (Minitech Dentimex). In a small number of experiments (less than 10), a craniotomy was instead performed on the left side. Due to the small number of such experiments a comparison of the effects of lateralisation was not performed. The dura mater was removed and the exposed cortex was covered with agar (1.5% agar in 0.9% saline). The position of the craniotomy was assessed using the coordinates displayed in figure 3-1 which are based on previous studies of guinea pig auditory cortex (for example Wallace et al. 2000). Further information that helped to identify the primary auditory cortex was provided once recording had begun. For example, the direction of tonotopy, the preference for tones over noise and the response latency of the neurons (typically 10 ms) could be used to disambiguate A1 from the other cortical areas (Wallace 2000).

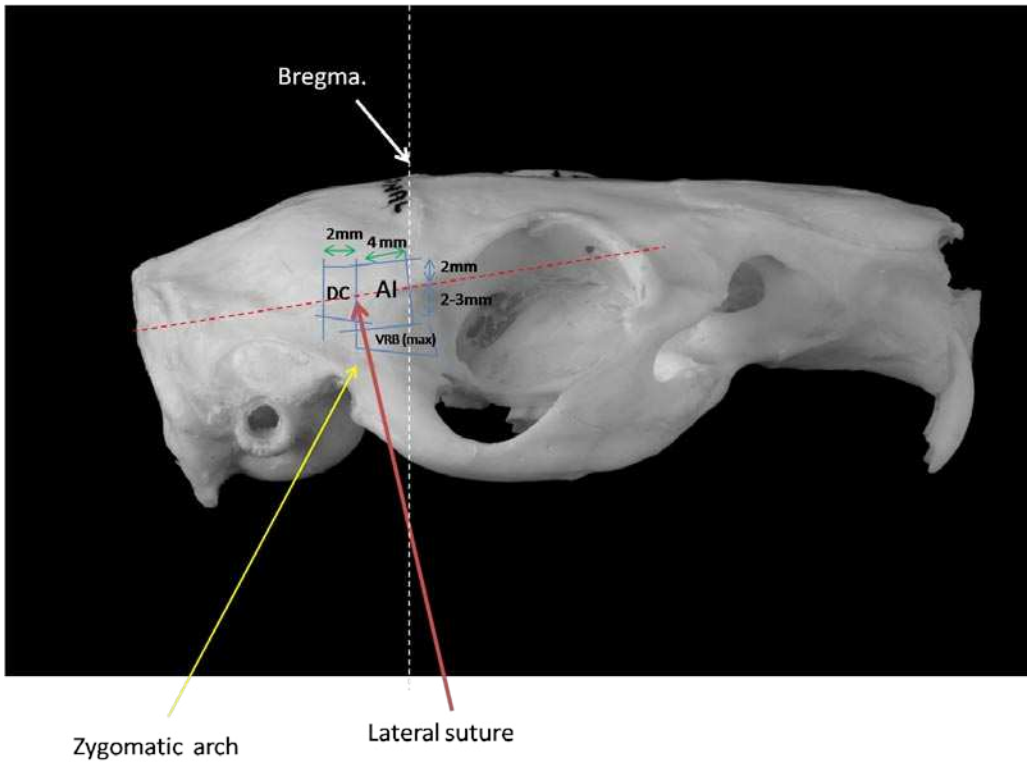


Figure 3-1. A photograph of the guinea pig skull showing the approximate position of the primary auditory cortex (A1) with respect to the dorsal cortex (DC) and the ventro-rostral belt (VRB).

3.3. Stimulus generation

Figure 3-2 shows a schematic of the hardware that was used to generate stimuli. The RX6 digital signal processor (Tucker-Davis Technologies) was used to generate auditory stimuli (noise and tones) under the control of the BrainWare software program, run from a Viglen PC (Pentium III, 600 MHz processor with 256 MB RAM). The output then passed through a 24 bit sigma-delta digital-to-analog converter before being attenuated with the programmable attenuator version 5 (PA5 in figure 3-2). The attenuated signals were then fed through SM3 mixers and further attenuated by 20 dB by a power attenuator, to limit the maximum output to 100 dB SPL. Stimuli were

then presented binaurally to the animal through custom-modified Radioshack 40-1377 tweeters (M. Ravicz, Eaton Peabody Laboratory, Boston, MA) which fitted into the hollow speculae via a damped tube (diameter 2.5 mm, length 34 mm).

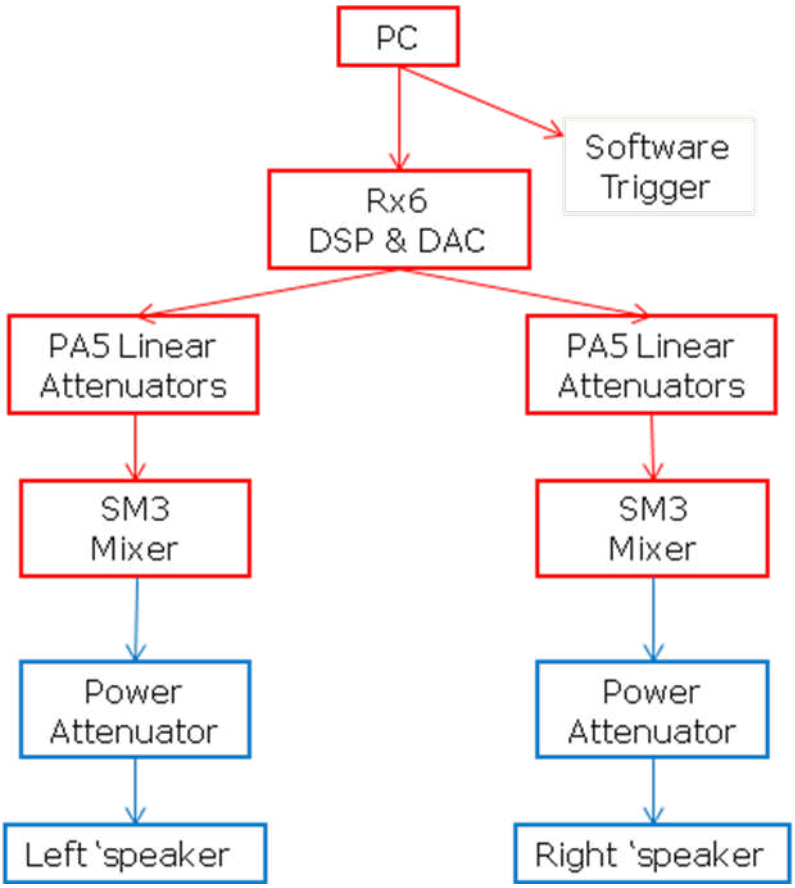


Figure 3-2. Schematic of the equipment that was used to generate the stimuli. Equipment in red boxes was positioned outside of the attenuated chamber while equipment in the blue boxes was situated inside.

3.4. Sound system calibration

The level of the stimuli was controlled by varying the amount of attenuation that was applied by the PA5 attenuators. To calibrate the sound system, white noise bursts were presented to each ear separately at full

system output (120 dB SPL). The bursts were recorded by a condenser microphone (Brüel and Kjaer 4134) positioned around 3-4 mm from the tympanic membrane. The sound waveform at the ear drum was averaged over 20 repetitions, Fourier transformed and corrected for microphone sensitivity and probe tube characteristics. The sound level at the ear drum is given by subtracting the attenuation from 100 dB SPL (maximum output of 120 dB SPL minus 20 dB from the power attenuator). However, where it has been necessary to convert attenuation to level for a given frequency (for example see chapter 4), an average calibration, based on the calibrations from 20 experiments, was used.

3.5. Neurophysiological recording

Neurophysiological recordings were performed with a multi-electrode array (MEA), to maximise the number of units that could be recorded at one time. The array consisted of between 4 and 8 glass-insulated tungsten electrodes, manufactured in-house (Merrill and Ainsworth, 1972; Ainsworth et al., 1977; Bullock et al., 1988) and attached to a printed circuit board using epoxy resin (as shown in figure 3-3). The electrodes were individually connected to different channels of the circuit board using electro-conductive paint.

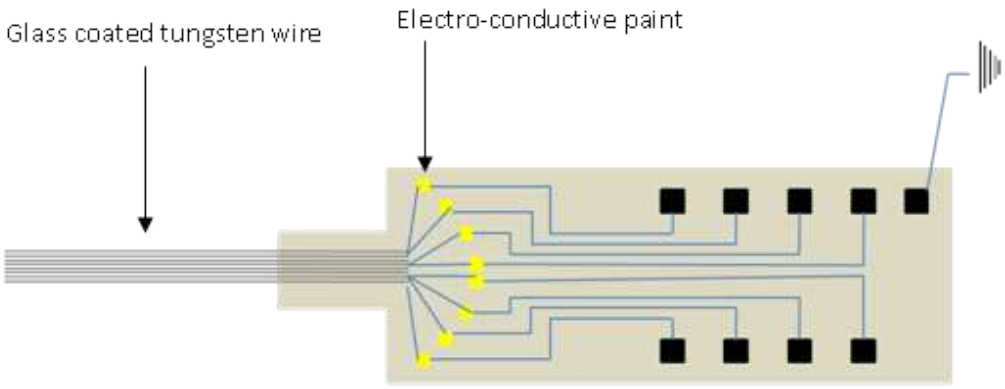


Figure 3-3. Illustration of the multi-electrode array as built on the printed circuit board.

Figure 3-4 shows the neurophysiological recording setup. The signals from the MEA were initially band-pass filtered (0.16 – 6000 Hz) using a high impedance head-stage (TDT RA16AC). The output from the head-stage was then digitised using a pre-amplifier (TDT RA16PA) and band-pass filtered again (300 – 3000 Hz) and amplified using a digital signal processor (RX7). The RX7 was controlled by BrainWare software as was the online control of the recordings. The threshold amplitude, above which snippets of waveform that followed any amplitude fluctuation were recorded, could be set independently for each channel of the MEA. The shapes of spikes that were being recorded could be viewed both in BrainWare and on an oscilloscope and the production of spikes could also be followed aurally through a loudspeaker.

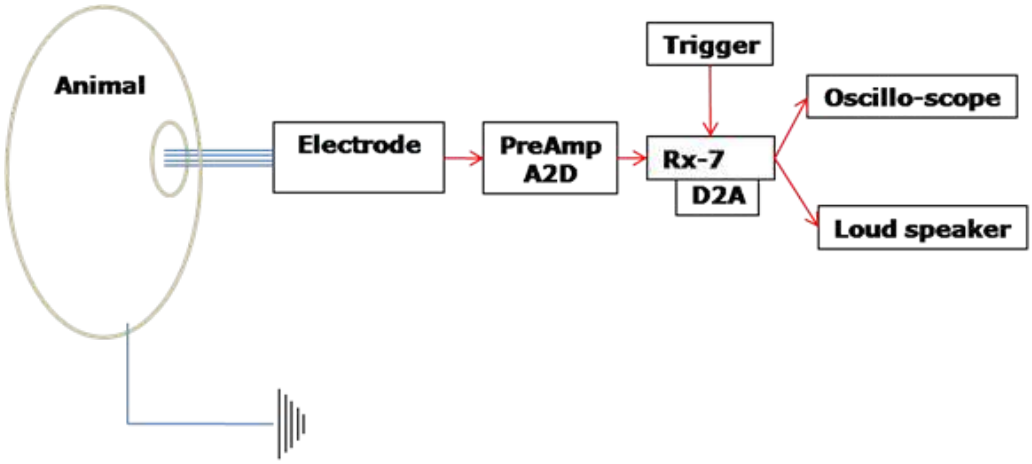


Figure 3-4. Illustration of the neurophysiological recording setup.

3.6. Searching for units

100ms broadband noise bursts (2ms cosine squared gates), attenuated by 10 dB (giving a level of around 80 dB SPL) were used as a search stimulus. Once a response was located, single 50ms tones (2ms cosine squared gates) were presented at a rate of 1/second at a range of frequencies and levels in order to determine the receptive field of a unit. The stimuli for an experiment were then presented, as described in each chapter.

3.7. Spike sorting

The waveforms recorded on each electrode could belong to a single neuron or to a collection of neurons. In order to separate the waveforms and attribute them to independent neural units, spike sorting was performed. The output files from BrainWare were converted using MatLab into a format that could be read into Plexon spike sorting software. Files that were recorded at the same location were concatenated so that spikes were sorted in the same way for each file. Spike sorting was performed using the Plexon Offline Sorter.

This program represented the spikes as waveforms and timestamps in either a 2- or 3-dimensional matrix (see figure 3-5 for examples of the 2D representation). Initially, waveforms were aligned at their first minimum peak to account for small temporal variations. Principal Component Analysis (PCA) was then applied the waveforms.

PCA is a statistical technique that can be used to transform a number of possibly correlated variables into a smaller number of uncorrelated variables: the principal components. Plexon calculates the first three principal components. The first principal component (PC1) will essentially be a generated waveform shape that explains the most variance in the data. The second principal component (PC2) will be a generated waveform that explains the majority of the residual variance after PC1 and the third principal component (PC3) will explain the majority of the residual variance after PC2. The degree to which a principal component fits an individual waveform (W) can be used to calculate a weight value so that $W = aPC1 + bPC2 + cPC3$, where a, b and c are the weights or levels of fit. These weights can then be plotted against each other. It is assumed that spikes from the same neuron have similar waveform shapes so that the level of fit or weight of each of the principal components will be similar. Thus, waveforms from the same unit should tend to form distinct clusters when viewed across these dimensions.

The weights of other features of the waveforms could also be used to discriminate them, for example the peak-valley amplitude or the height of the waveform at any user-defined point. Figure 3-5a shows an example where

two units have been retrieved from the spike sorting. The waveforms of the two units (one in yellow and one in green) are shown as insets in the left panel and both panels show different components plotted against each other. The left panel shows the weights of PC1 and PC3 plotted together: two distinct clusters can be seen relating to the two units. Components could also be viewed across timestamp (the right panel in figure 3-5a) which proved useful when there was a trend for the components to change over time. For example, in figure 3-5b, the spike waveform shown in the inset is variable. However, in this case it seems that PC1 is changing gradually over time as shown in the main panel of figure 3-5b, so it is likely that the waveforms come from the same unit. The unit shown in figure 3-5b is a single unit. In most cases single units were flagged at the recording stage as they were distinguished by having singular spikes with large signal-to-noise ratios. Units that were sorted using Plexon, such as those in figure 3-5a, were termed sorted units. The spikes recorded from sorted units are likely to come from a single neuron and so the sorted units were analysed as single units. Figure 3-5c shows an example of a multi-unit, where no obvious clusters were observed across any of the components.

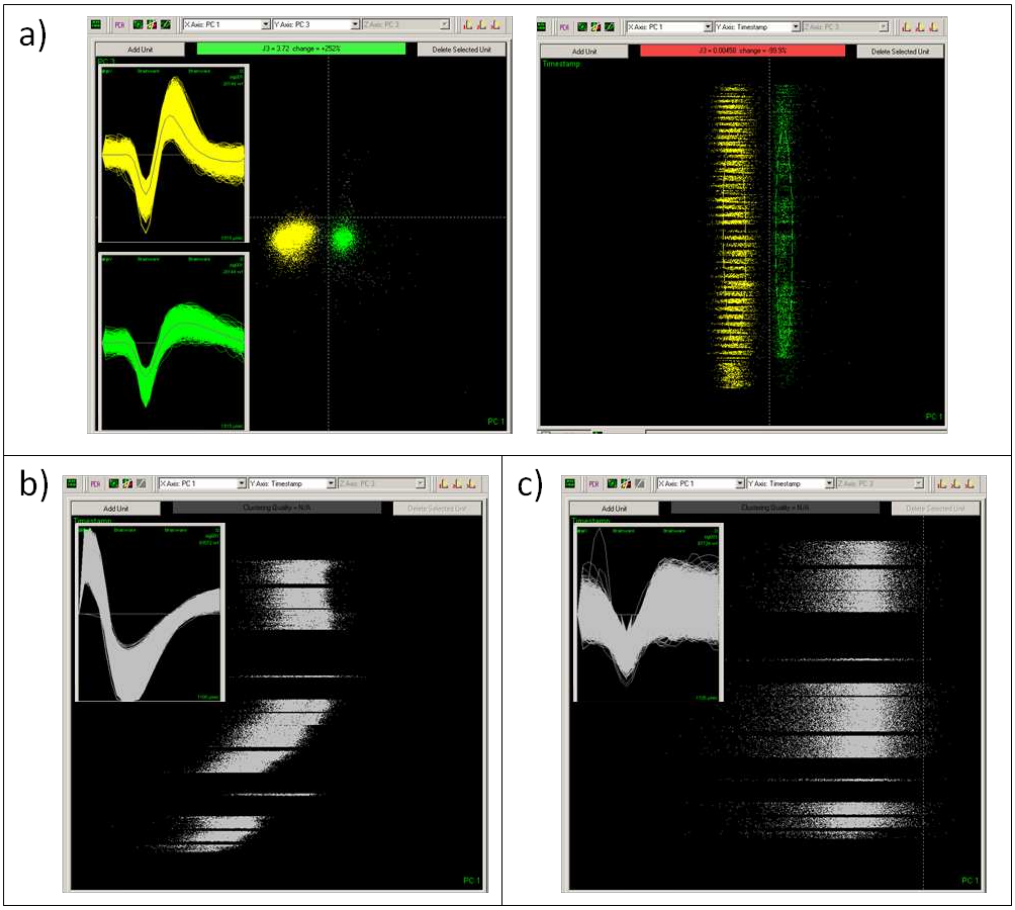


Figure 3-5. Example of units that have been sorted using Plexon. a) Two sorted units viewed across different components. The left panel shows PC1 against PC3 and the right panel shows PC1 against timestamp. The waveforms for each unit are displayed in the insets. b) A single unit example. PC1 is shown against timestamp in the main panel, showing a change in the weight associated with PC1 over time. The waveforms for this unit are displayed in the inset. c) Multi-unit example with PC1 shown against timestamp and the waveforms displayed in the inset.

3.8. The response time window

Most analyses were performed using custom-made programs in MatLab. SPSS 15 for Windows was used to perform statistical analyses. For each unit an analysis time window was set manually by examining a PSTH of the response to the first tone in all conditions (examples of which are

displayed in figure 3-6). To ensure that this time window was sufficient, a display routine was used that showed PSTHs for various conditions with the analysis time window indicated on the PSTHs. The time window could then be changed until I was satisfied that it captured the response to each tone. Spike count distributions for a given condition were computed by counting the number of spikes elicited by a tone within the time window for each repetition of a stimulus. Spike counts were computed by averaging the spike count distributions.

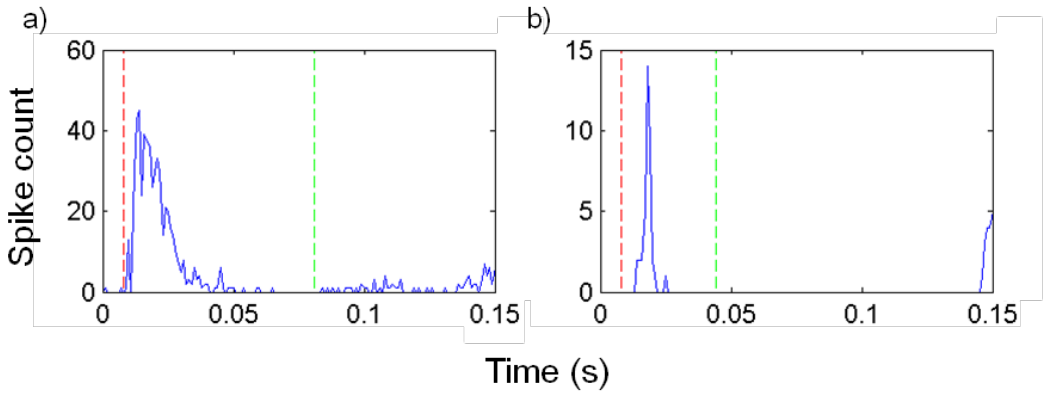


Figure 3-6. Summed PSTHs of the response to the first tone, showing how the analysis time window was set. The dashed red line indicates the start and the dashed green line indicates the end of the analysis time window. a) and b) show examples from two different units.

4. The frequency specificity of forward suppression

4.1. Introduction

This chapter is concerned with the question of whether forward suppression can be considered frequency-specific in the same way that long-term adaptation is (Ulanovsky et al., 2003; Ulanovsky et al., 2004). Specifically, the main hypothesis that is being tested is that varying the frequency of the probe tone away from CF in a two-tone paradigm will affect the pattern of suppression that is observed. This change may be manifested as a shift in the locus of suppression or the effects may be more subtle, manifested as changes in the area or shape of the suppressed region. However, the prediction is that suppressed tuning curves will shift away from the unit CF in the direction of the probe tone frequency. Further, I will investigate whether any changes in suppression are observed when the temporal gap between conditioner tones and probe tone is varied.

4.2. Methods

4.2.1. Stimuli

Two or more tones were selected from the receptive field to be used as probe tones (blue dots in figure 4-1a). To investigate the effect of prior stimulation on the response to the probe tones a conditioner tone was presented before it. These conditioner tones were parametrically varied over a range of frequencies and levels. In many cases, conditioning was investigated when there was no temporal gap between the conditioner and

probe tones. In other cases the inter-stimulus interval (ISI) was also varied. The range and spacing of conditioner tone frequencies and levels differed in each recording due to differences in receptive fields. Conditioner tone frequency steps were 1/4, 1/5 or 1/6 octaves and the level steps were 10 or 15 dB SPL (usually from 10 to 100 dB SPL attenuation). Either 2 or 3 probe tones were used in each recording and conditions were presented pseudo-randomly. The stimulus set was repeated between 3 and 10 times. A single set of forward suppression stimuli took around one hour to complete. Both the probe tone and the conditioner tones were gated 50ms tones (2ms cosine squared gates). Although there is evidence that tone interactions can occur at ISIs of longer than 1 second (Hocherman and Gilat, 1981; Werner-Reiss et al., 2006), in the interests of time, the repetition time for each condition was 1 second. Thus, it is possible that the system is being considered in a habituated state.

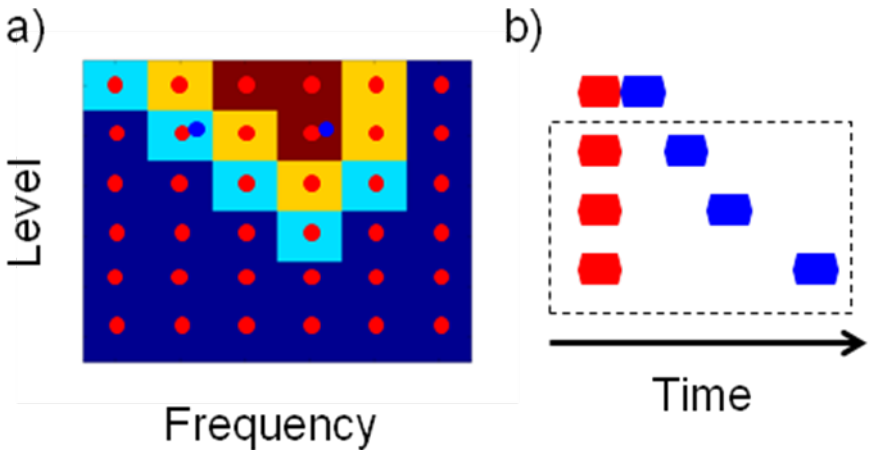


Figure 4-1. Schematic of the stimulus paradigm. a) In a given condition the conditioner tone (red circles) was selected randomly from a range of frequencies and levels and the probe tone (blue circles) was randomly selected from 2 or 3 tones that were selected from the receptive field. b) In every unit the effect of the conditioner tone (red hexagon) on a probe tone (blue hexagon) that followed directly after (ISI = 0 ms) was tested. In a subset of units the ISI was also randomly varied to investigate the recovery from suppression (those conditions within the dashed box).

4.2.2. Analysis

4.2.2.1. Receptive field and suppressed receptive field generation

A receptive field (RF) was created from the responses to the conditioner tones by displaying the spike counts as a function of conditioner tone level and frequency. Suppressed receptive fields (SRF) for each probe tone condition were then created by considering the response to each probe tone with respect to the level and frequency of the conditioner tone that had preceded it.

To prepare the RF and SRF for the use of an automatic threshold algorithm which extracted threshold tuning curves, two steps were taken. Firstly, an extreme value filter was used to account for any values that were much smaller or larger than the neighbouring values. It acted by considering the frequency/level cells in the RF in 3 x 3 windows. If the central value was the smallest or largest of the 9 values then it was set to be that of the next smallest or next largest value respectively. The 3 x 3 window then moved across the matrix of conditioner tone frequency-level combinations and applied to every block (overlapping). This step was necessary to remove extreme values for the smoothing step described below because extreme values had a large effect on the RFs given the low number of stimulus repetitions. Note that although this filter was applied to both the RF and SRF, the effect on the RF was minimal since excitatory responses were generally more reliable and the RF was calculated from the response to the conditioner tone across multiple probe conditions.

In the second step, the RF and SRF were smoothed using a pyramidal 3 x 3 window (the product of two triangular windows along the two axes, see Moshitch et al., 2006).

4.2.2.2. The frequency threshold curve

Two types of frequency threshold curve (FTC) were derived from both the RFs and the SRFs using an automatic algorithm (Sutter and Schreiner, 1991; Moshitch et al., 2006). The FTC for the RF will be referred to as the FTC and the FTC for the SRF will be referred to as the suppressed frequency tuning

curve (SFTC). For the RF, a frequency/level bin was considered to have a response to a tone if its spike rate was larger than the spontaneous response + criterion value * (maximal response – spontaneous response), where the spontaneous response was an average of all conditioner tone responses at the lowest (i.e. undriven) presented level (Sutter and Schreiner, 1991). For the SRF, a frequency/level bin was considered to have been suppressed if its spike rate was lower than the probe tone response – criterion value * (probe tone response – spontaneous response), where the probe tone response was the mean response to the probe tone when it was preceded by conditioners presented at the lowest intensity. The two types of FTC differed in the criterion values used: 0.4 for a threshold FTC (or SFTC) and 0.9 for a best FTC (or SFTC). The threshold FTC shows at which levels and frequencies the neuron is most sensitive either to excitation (for the RF) or to suppression (for the SRF). The best FTC gives an indication of the frequencies and levels at which the strongest excitation or suppression is observed.

In essence the algorithm used to create the FTC traced the region of the RF or SRF where the response crossed the criterion threshold outlined above. However, due to the SRFs generally being noisier than the RF, the algorithm had to be made more robust for the SRFs. For the 0.9 FTC the algorithm worked as outlined above. For the 0.4 FTC, any regions of activity that were isolated from the main region were ignored and the FTC was fitted with an 8-degree polynomial curve. For the 0.9 SFTC again any isolated regions were ignored but a polynomial fit was not used. The 0.4 SFTC was

constructed in a similar way to the 0.4 FTC except that the algorithm only selected the largest uninterrupted supra-threshold area in the SRF, ignoring supra-threshold areas that were isolated. For each frequency column, the threshold was calculated using linear interpolation between the spike rate at the lowest suprathreshold level and the level bin below it.

To ensure that the automatic algorithm was characterising the RFs and SRFs adequately, FTCs and SFTCs were checked visually on a unit-by-unit basis. The method that has been chosen for each type of RF (outlined above) reflects the most adequate approach, as assessed visually, after experimentation with a number of different strategies. In addition, several automatic criteria, for example little or no spike count response to the probe tone, were utilised to remove RFs and SRFs that were too noisy to characterise (mainly due to low spike counts).

4.2.2.3. Facilitation

Although facilitation was observed in some units, the prevalence of facilitatory effects was low compared to the prevalence of suppression. Also, it was difficult to characterise facilitation using the FTC method outlined above. So, in this chapter I will concentrate on the suppressive effects of conditioner tones.

4.2.2.4. Characterizing the frequency threshold curves

4.2.2.4.1. Threshold and best and characteristic frequencies

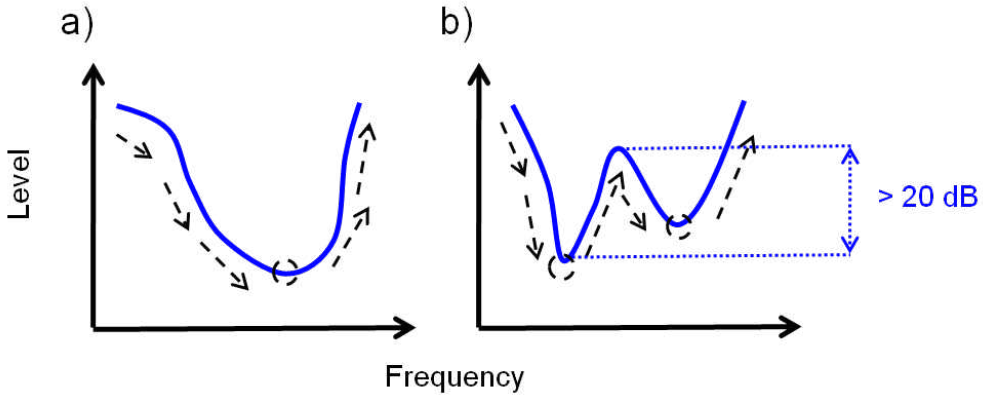


Figure 4-2. How the automatic algorithm worked. a) The algorithm moved along the threshold curve until it reached the minimum point. Minimum points were flagged and then collected into a single 'peak' if they were all within 20 dB (or 10 dB for the best FTC). b) If two minima were separated by more than 20 dB (or 10 dB for the best FTC) then they were grouped into separate 'peaks'. Solid blue lines show the FTCs. The dashed arrows show the direction that the algorithm followed and dashed circles show the minima associated with separate 'peaks'.

An automatic algorithm was used to trace the FTCs and to separate them into peaks (as shown in figure 4-2). The algorithm worked by moving along the threshold curve from low to high frequencies, picking out minima in the curve. Threshold minima were then grouped together into a 'peak' if all points between them were within 20 dB (or 10 dB for the 0.9 condition) of the lowest threshold (after Moshitch et al., 2006). The characteristic frequency (CF) and the best frequency (BF) were defined as the frequency at which the lowest threshold occurred for the FTC with a 0.4 and 0.9 criterion respectively. In a similar manner, the characteristic suppressed frequency

(CSF) and the best suppressed frequency (BSF) were defined as the frequency at which the lowest threshold occurred for the SFTC with a 0.4 and 0.9 criterion respectively. Inspection of the data revealed that RFs and SRFs generally had single 'peaks' so that in cases where more than two 'peaks' were selected by the threshold algorithm, only the lowest threshold 'peak' was used.

4.2.2.4.2. Bandwidth

The upper and lower frequency limits at 10 dB and 40 dB above the minimum threshold were found and used to define the 10 dB and 40 dB bandwidth for both the FTCs and the SFTCs.

4.2.2.4.3. Area

The suppressed area was calculated by integrating the SFTC. The area of suppression thus represents the range of conditioner tone frequencies and levels that caused suppression.

4.2.2.4.4. Probe Following Index

To characterise the locus of suppression relative to the CF (or BF) and the probe tone frequency, a probe following index (PFI) was calculated. The PFI is essentially the angle that a given data point makes with the axes when the distance from the CSF to the CF is plotted as a function of the distance of the probe from CF. Specifically, the PFI for the CSF is given by:

$$PFI = \frac{4}{\pi}\arctan\left(\frac{CF - CSF}{CF - probe}\right)$$

Values close to 1 correspond to points with angles of around 45° that lie around the probe-following gradient of 1 (indicated by a dashed blue line in figure 4-4a). A value of 0 indicates an angle of 0° where the CSF is at the CF (indicated by a dashed red line in figure 4-4a). For the BSF, the CF was replaced by the BF in the equation.

4.2.2.4.5. Maximum suppression time

To define when suppression was considered to have occurred, the SRF was collapsed across level to give a mean spike count for each frequency. If the mean spike count at any frequency was less than 40% of the response to the probe tone alone then suppression was considered to have occurred. For a given unit and probe tone frequency, the maximum suppression time was then taken as the last ISI in which suppression occurred.

Note on nomenclature: A criterion value of 0.4 (40%) suppression can be considered as conservative compared to previous criteria for suppression (Calford and Semple, 1995; Brosch and Schreiner, 1997; Scholl et al., 2008). This value was used because it allowed suppression to be characterised in most units as compared to lower criterion values (assessed visually as described above). To compare equivalent response characteristics the criterion for the FTC was also set to be 0.4. The absolute probe tone level was found to have little correlation with any of the dependent variables that were investigated in this chapter. Correlations were observed, however, when the probe tone level was expressed in terms of the RF of the unit. Specifically, in this chapter, the probe tone level will be expressed as the threshold at the

probe tone frequency in the FTC minus the absolute tone level. As the 0.4 criterion leads to relatively high thresholds, the level with respect to the threshold may be negative in some cases.

4.3. Results

A total of 161 units were recorded (56 single units and 105 multi-unit), yielding 276 SRFs for an ISI of 0 ms (90 from single units and 186 from multi-units). Figure 4-3 shows an example of the effect of varying the probe tone frequency on the suppression pattern that is observed. The RF of the unit is displayed in figure 4-3a. The FTCs at 0.4 and 0.9 criteria are displayed as a solid and dashed white line respectively. The CF and the BF for this unit are around 20 kHz. In this example, two probe tone frequencies were presented (indicated by yellow stars): one at 20 kHz, close to the CF and one at 12 kHz, away from the CF. The SRF for the 20 kHz probe tone is displayed in figure 4-3c, along with the SFTCs at 0.4 and 0.9 criteria. The SRF has a similar shape to the RF and both the CSF and BSF are at the probe tone frequency. When the probe tone frequency was away from CF (as shown in figure 4-3b), there were two main effects. Firstly the SRF was larger in area, having a lower threshold and a wider bandwidth. Secondly, both the CSF and the BSF were closer to the probe tone frequency than to the CF or the BF.

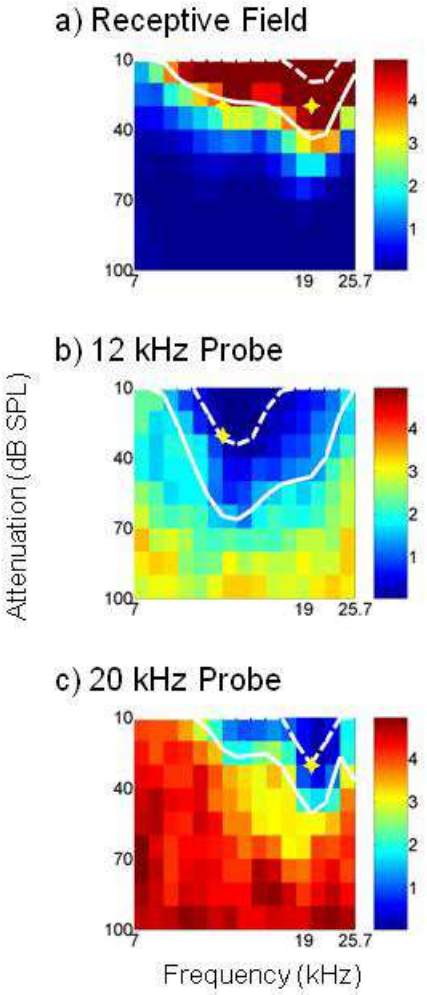


Figure 4-3. Single unit example of the effect of varying the probe tone frequency on the pattern of suppression. The probe tones are indicated by yellow stars. The 0.4 and 0.9 SFTCs are indicated by solid and dashed white lines respectively. Spike counts indicated by the key are spikes per presentation. a) The RF of this unit – the response to tones presented at the frequencies and attenuations indicated on the axes. b and c) The response to the 12 kHz and 20 kHz probe tones when preceded by tones at the frequencies and attenuations indicated on the axes.

These two effects of presenting a probe tone away from CF on the suppression pattern will be considered separately in the next two sections.

4.3.1. The locus of forward suppression

4.3.1.1. The locus of immediate forward suppression is influenced by the probe frequency

Figure 4-4a displays the distance from the CF to the CSF in octaves as a function of the distance from the probe tone to the CF. Dashed lines indicate the trend that would be expected if the CSF was always at the frequency of the CF (red line) or at the probe tone frequency (blue line). The distance from the CSF to the CF was not found to be significantly different for the SU and MU populations (one-way ANOVA, $p = 0.49$). A linear regression of the single and multi-unit data is displayed as a solid black line ($m = 0.67$, $R = 0.67$). A gradient of 0.67 indicates that the CSF was, in general, closer to the probe tone frequency than to the CF. To investigate this further, the probe tone-following index (PFI) for each data point was calculated and histograms of the SU and MU data are displayed in figure 4-4b. The majority of CSFs cluster broadly around the PFI value of 1, again indicating that suppression generally follows the probe tone. A proportion of CSF values are also at a PFI of 0. These points can be seen in figure 4-4a, along the abscissa, indicated to the right of the figure as ‘CSF at CF’.

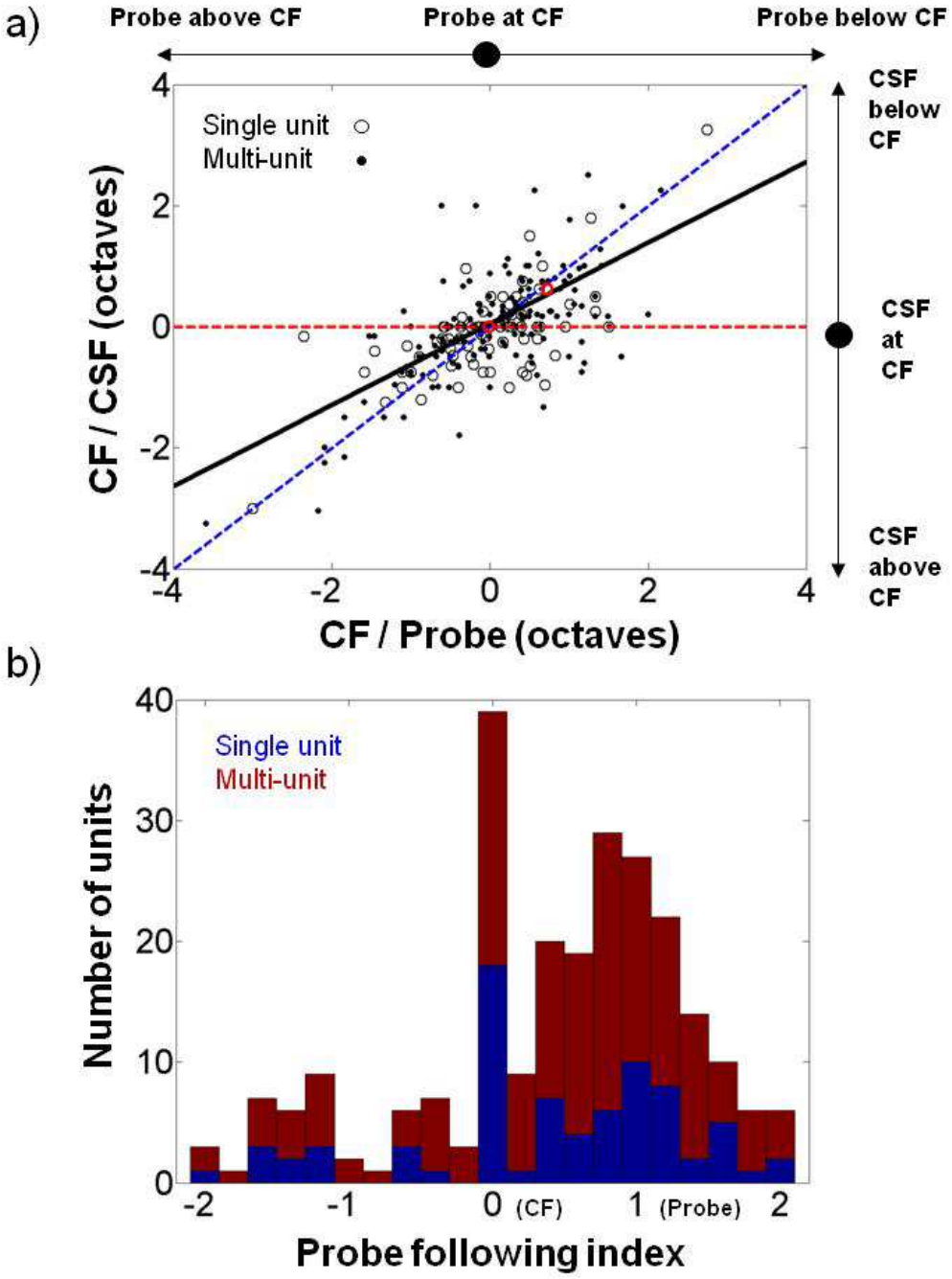


Figure 4-4. The locus of suppression from the 0.4 SFTC. a) The distance of the CSF from the CF as a function of the distance of the probe tone from the CF. Dashed lines indicate the trends that would be expected if the CSF was always at CF (red) or always at the probe tone frequency (blue). The solid black line shows a linear regression of all of the data. The 2 probe tone conditions from the example unit in figure 4-3 are indicated by red circles. b) The PFI for each of the data points in a).

This apparent evidence of CF-following could in fact be an artefact of sampling. The frequency steps used for the RF and SRF, from which the CF and CSF were measured, varied (1/4, 1/5 or 1/6 octave) from recording to recording. Thus frequency differences between the CF and CSF could be multiples of 1/4, 1/5 or 1/6 octaves. Regardless of the step size a CF-CSF difference of 0 was possible in every recording. Thus a PFI of exactly zero was more likely than any other value. A simulation confirmed this possibility. In figure 4-5a, evenly distributed random numbers of up to 1/8 octave have been added to all of the CF and CSF values and the PFI has been recalculated. The CF-following peak disappears (indicated by the arrow) while the probe-following tendency remains. The CF-following peak returns if the values with added noise are turned back into discrete numbers, as indicated by the arrow in figure 4-5b.

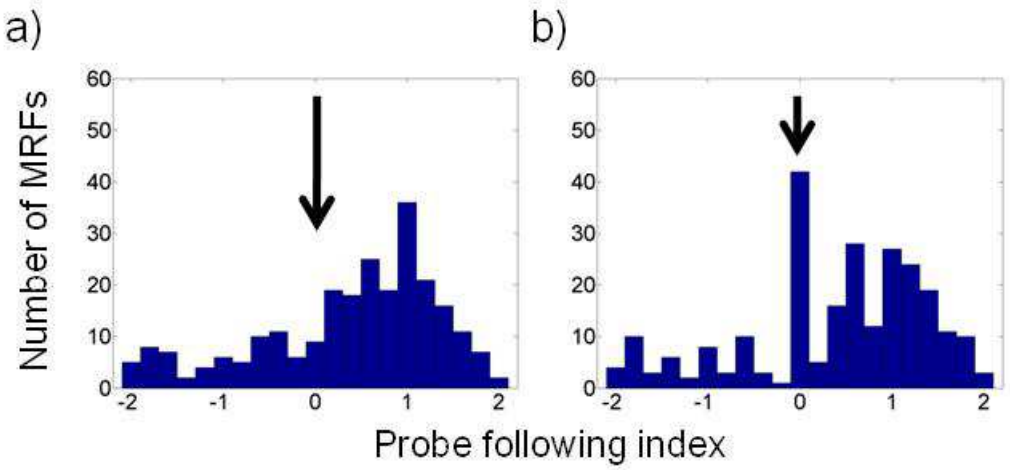


Figure 4-5. Simulation demonstrates that CF-following may be an artefact of sampling. a) When noise is added to the data, the CF-following peak disappears as indicated by the arrow. b) The CF-following peak returns if the noisy data is transformed back into discrete values as indicated by the arrow.

Figure 4-6 is similar to figure 4-4, except that the BF and BSF are displayed. Again, the distance between the BSF and the BF for the SU and MU populations was not found to be significantly different (one-way ANOVA, $p = 0.52$). The solid black line in figure 4-6a indicates a linear regression of the data ($m = 0.92$, $R = 0.87$). More of the variance is accounted for by the fit than for the CSF, as indicated by a higher value of R, and a gradient of 0.92 implies that the BSF was, in the majority of SRFs, closer to the probe tone frequency than to the BF. Figure 4-6b displays a histogram of the PFI for the BSF.

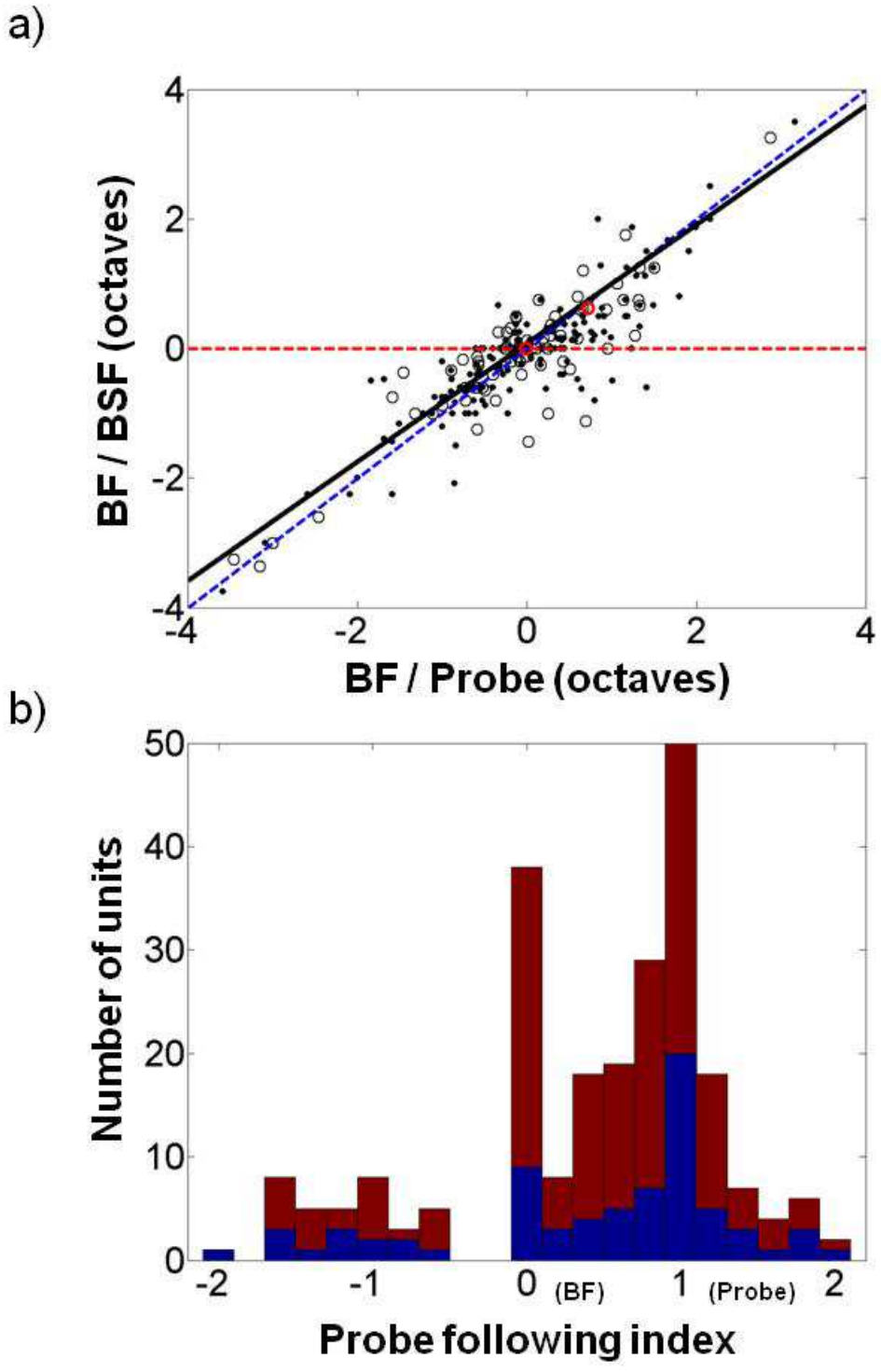


Figure 4-6. The locus of suppression from the 0.9 SFTC. Same conventions as used in figure 4-4.

Further evidence of the influence of the probe tone frequency on the locus of suppression comes from ANOVAs applied to general linear models

with the probe tone frequency and either the CF or BF as predictors and either the CSF or the BSF as the dependent variable. Significant interactions exist between both of the predictors and the dependent variable, but the amount of variance explained by the probe tone frequency is higher for both the CSF (partial eta^2 = 0.36 for the probe tone and 0.12 for the CF) and the BSF (partial eta^2 = 0.64 for the probe tone and 0.12 for the BF).

4.3.1.2. The probe frequency affects the locus of forward suppression as the ISI is increased

I have shown that suppression is influenced by the frequency of the probe tone when the probe tone is presented directly after the conditioners. In this section I address the question of whether this frequency specificity is observed when the temporal gap between conditioner tone and probe tone is increased.

Figure 4-7 shows a SU example of the suppression observed at two different probe tone frequencies over time. Figure 4-7i shows the RF for this neuron and figure 4-7ii, iii, iv and v show the SRFs for the two probe tone conditions with ISIs of 0, 50, 100 and 200 ms respectively. The same conventions are used in figure 4-7 as in figure 4-3, so the probe tones are indicated by stars (either yellow or black for clarity) and the 0.4 and 0.9 SFTCs are indicated by solid and dashed white lines respectively. The probe tone at 2 kHz is closer to the BF (3 kHz) and the CF (2.4 kHz) of the neuron than the probe tone at 5 kHz. As has been noted previously (Brosch and Schreiner, 1997), the suppressed area is large initially and then recedes as the temporal

gap between conditioner tone and probe tone is increased. Evidence of frequency specificity can be observed in the 0.9 SFTC for the 0 ms gap condition, with the BSF correlating closely to the probe tone frequency, but not in the 0.4 SFTC. Increasing the temporal gap between conditioner tone and probe tone caused the SFTC to retract around the frequency of the probe tone, and thus the suppression became more specific to the probe tone frequency.

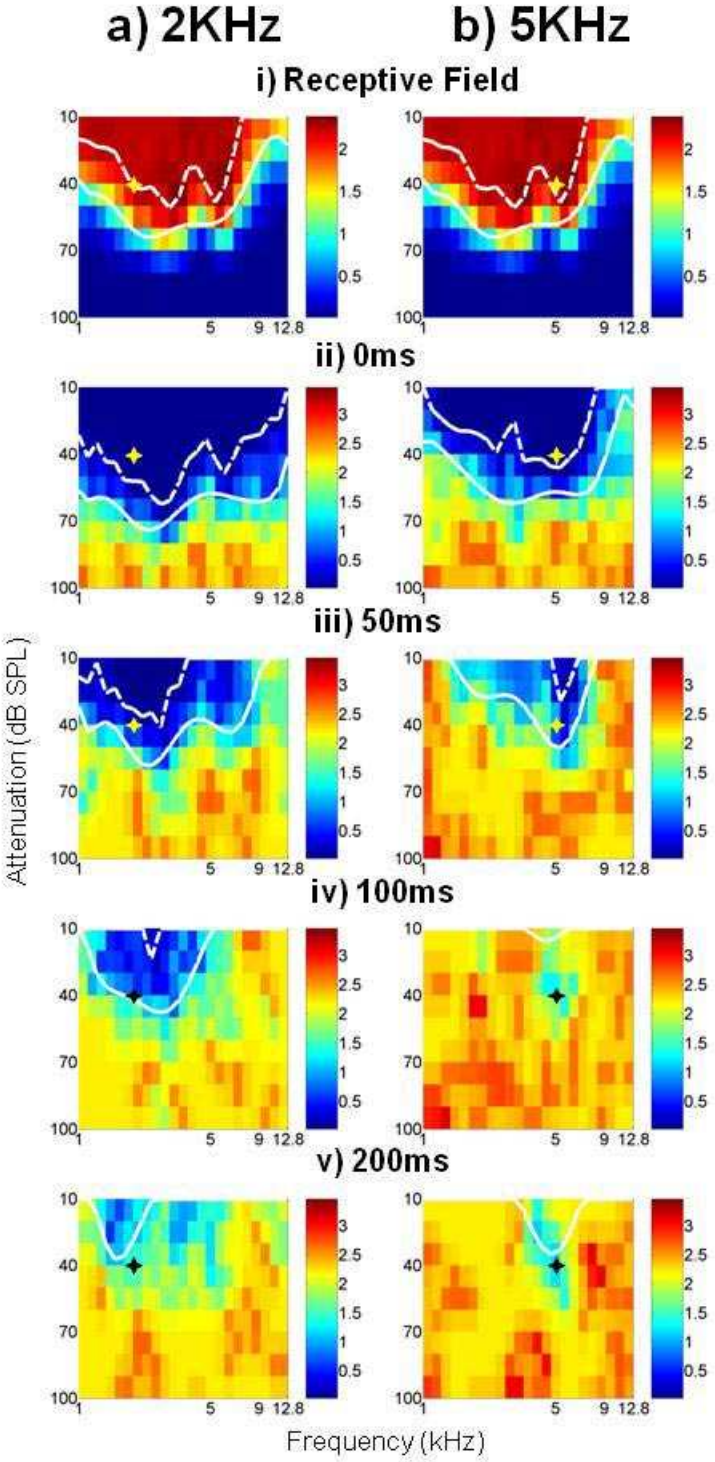


Figure 4-7. Single unit example of the time course of suppression. Same conventions as used in figure 4-3. Stars represent the probe tones, displayed in either yellow or black for clarity. The ISI is indicated above each pair of RFs.

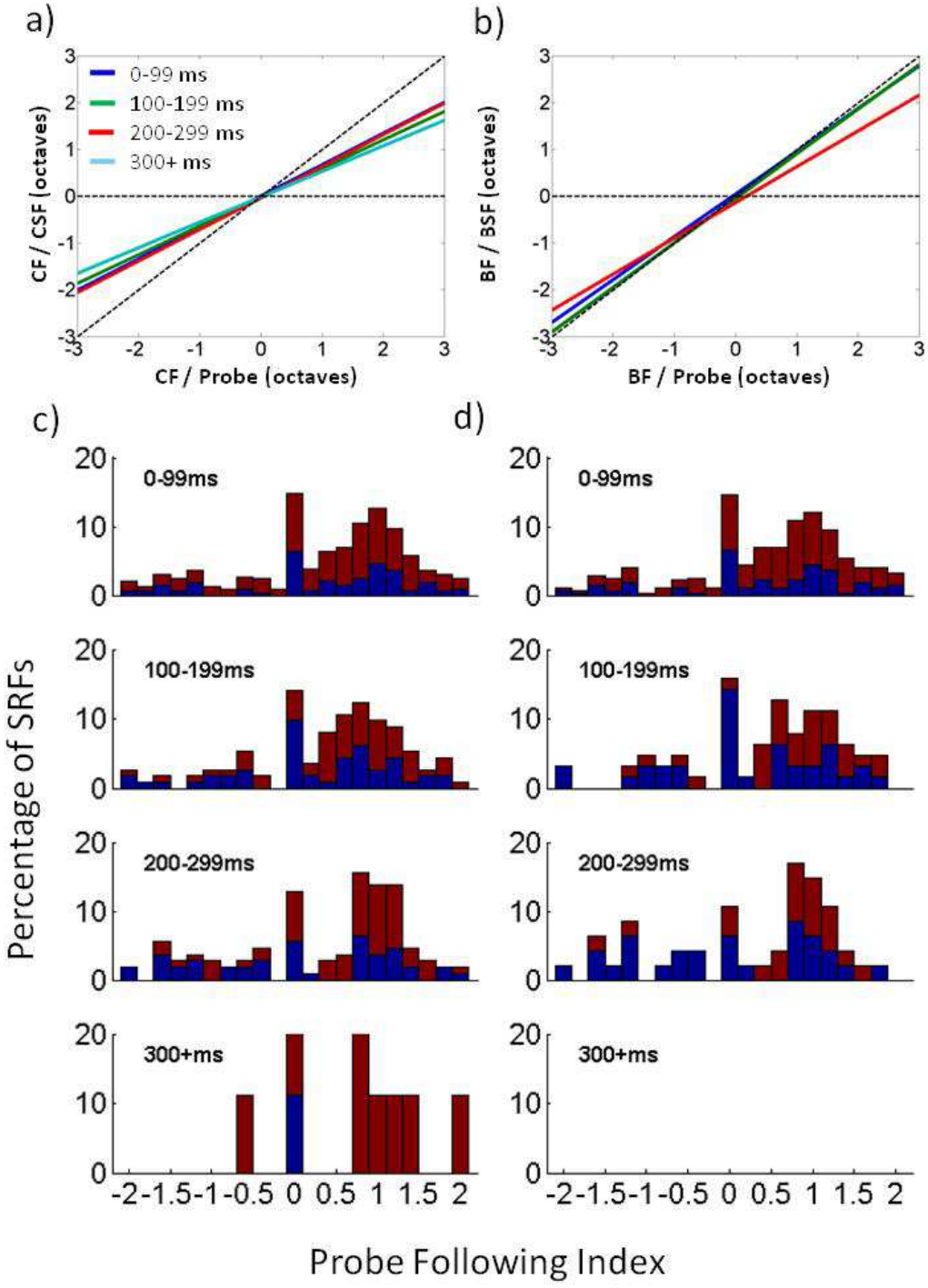


Figure 4-8. The locus of suppression as the temporal gap between conditioner tone and probe tone is increased. a and b) The linear regressions for the CSF and BSF data grouped in time, as indicated in the legend. c and d) Histograms of the PFI for data points from a and b).

Figure 4-8 shows how the locus of suppression changes over time for the population of units (the fit parameters are displayed in table 4-1). Although in the example unit displayed in figure 4-7 the locus of suppression moved closer to the probe frequency over time, there was no significant relationship between time and the distance from either the CSF to the CF or the BSF to the BF (linear mixed model, p>>0.05). So, as indicated in figure 4-8a and b, the relationship between the probe tone frequency and both the CSF and the BSF remained relatively constant as the temporal gap between conditioner tone and probe tone was increased. This can also be observed if the PFI measure is considered (refer to figure 4-8c and d).

Time group (ms)	**0 – 99**	**100 -199**	**200 – 299**	**300+**
CSF Gradient (R value)	0.67 (0.68)	0.61 (0.53)	0.67 (0.69)	0.55 (0.82)
BSF Gradient (R value)	0.91 (0.84)	0.95 (0.82)	0.77 (0.70)	N / A

Table 4-1. Gradients and R values of linear regression fits between the probe tone frequency and the CSF or BSF, at each time gap.

4.3.2. The threshold, bandwidth and area of suppression

Having shown that the locus of forward suppression does depend on the frequency of the probe tone, this section addresses the effects of varying the probe tone frequency on the threshold and bandwidth of suppression. As noted in the example unit displayed in figure 4-3, the range of conditioner tones that caused suppression was larger for the off-CF probe tone than for

the on-CF probe tone. Figure 4-9a shows the SRF area as a function of the distance of the probe tone from BF. As the distance between the probe tone and the BF increased the area also increased (linear regression, m = 36, R = 0.37, *p<<0.01*). Another measure of the position of the probe tone within the RF is the probe tone level. This measure exhibits a negative correlation with the SRF area, as shown in figure 4-9b (linear regression, m = -1.3, R = 0.32, *p<<0.01*). So, probe tones that are deeper into the RF relative to threshold are suppressed by a smaller number of tones than those that are nearer to the threshold. A general linear model with BF/probe tone and threshold as predictors showed significant contributions from both measures with regression coefficients of 27 and -0.87 respectively.

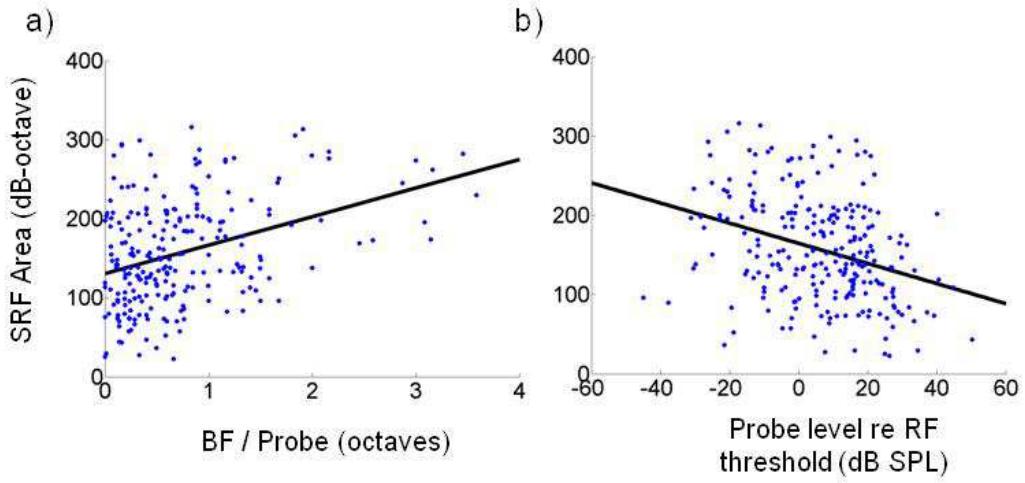


Figure 4-9. Measures of probe tone position against area. a) The area of the SRF as a function of distance between the BF and probe tone frequency in octaves. b) The area of the SRF as a function of the probe tone level relative to the RF threshold.

Given that the position of the probe tone within the RF predicts the suppressed area, a further question would be how the position of the probe

tone affects specific aspects of the suppressed area, for example the threshold and bandwidth of the SFTC. To investigate this, the data was entered into a multivariate general linear model with the threshold at the CSF and the bandwidth at 40 dB as dependent variables and the distance of the probe tone from BF and the probe tone level as predictors. The only significant relationship that was found was that the threshold increased as the probe tone level increased. Therefore, it is probable that the increase in area as the probe level decreases was due to the threshold at CSF decreasing rather than an increase in the bandwidth.

Figure 4-10 shows how measures of the suppression area change over time when the data is grouped as in figure 4-8. Figure 4-10a shows that the normalised mean SRF area decreases as the ISI increases. This decrease in area coincides with a mean increase in threshold relative to the threshold at time zero (figure 4-10b) and a decrease in the mean absolute bandwidth at both 10 and 40 dB SL (figure 4-10c and d).

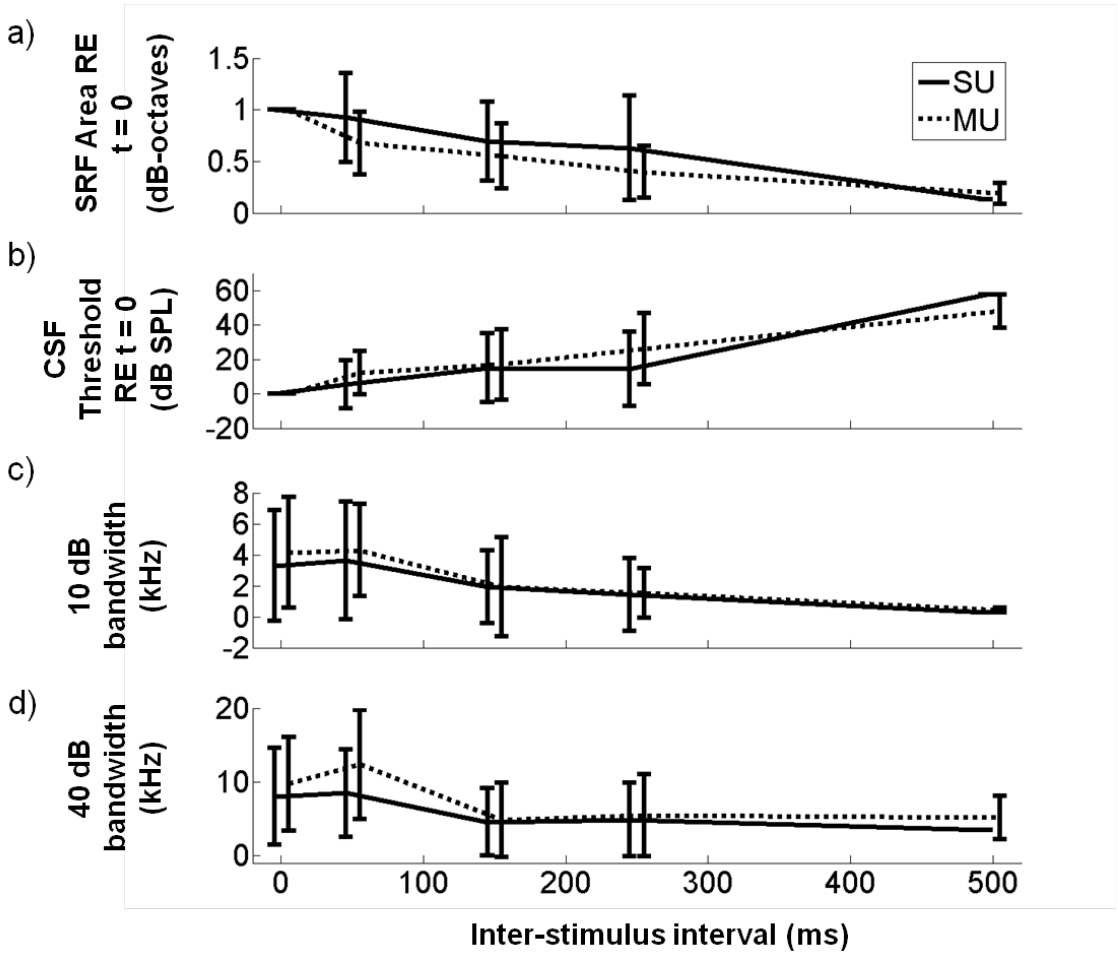


Figure 4-10. How the extent of suppression changes as a function of ISI. a) The area of the suppressed receptive field relative to the suppressed area when the ISI was 0 ms. b) The threshold of suppression relative to the suppressed threshold when the ISI was 0 ms. c) The 10 dB bandwidth of suppression. d) The 40 dB bandwidth of suppression. Error bars show standard deviations.

Figure 4-10 shows that the suppressed area decreases with the temporal gap between conditioner tone and probe tone. This has been shown previously (Brosch and Schreiner, 1997) and so in the context of this study it is interesting to address any differences that are due to the different probe tone conditions that were used. Firstly, the effect of the distance between the probe tone and the BF and the probe tone level on the suppressed area was

assessed for each time group using the same approach of a general linear model as before. Table 4-2 displays the coefficients of regression for each of the time groups. The probe tone level was a significant predictor of the area for all of the time groups, with the coefficient becoming more negative over time as the area decreased. The positive correlation between the distance of the probe tone from the BF and the suppressed area was only significant up to around 200 ms.

Time group	Distance between the probe frequency and the BF	Probe level relative to the threshold on the FTC	Intercept	N
0-99 ms	30.33 (**)	-0.72 (**)	123.05 (**)	328
100-199 ms	18.66 (*)	-1.35 (**)	87.88 (**)	109
200-299 ms	6.43	-1.61 (**)	79.12 (**)	106
300+ ms	21.19	-2.71 (*)	44.89 (*)	9

Table 4-2. Regression coefficients for area across time, with the probe tone level and the distance between the probe tone and the BF as predictors. * *p<0.05*. ** *p<0.01*.

A multivariate general linear model applied to the data from each time group showed a significant positive correlation ($p<<0.01$) between the probe tone level and the threshold for all but the 300+ ms time group. This may be because the number of observations for this group was low (N = 9). This

implies that an increase in the probe tone level leads directly to an increase in the threshold of suppression up to around 300 ms.

4.3.3. The maximum time of suppression

As has been shown, the suppressed area is correlated with the position of the probe tone in the RF and this correlation seems to be present when the temporal gap between conditioner tone and probe tone is varied. In this section I investigate how the position of the probe tone affects the maximum time at which suppression is observed. One hypothesis concerning the area of suppression would be that the larger the area, the longer it would take for recovery to occur. Further, because of the relationships that have been noted between the area and the probe tone level and frequency, it may be expected that these would have an effect on the maximum time of suppression.

Figure 4-11 shows the maximum time of suppression, grouped as before, as a function of the probe tone level and frequency and the SRF area at time zero. Of the small number of units tested at longer time gaps, none displayed suppression (at the criterion level) at a maximum time of 300 ms or over. Figure 4-11a shows the number of SRFs that have a maximum suppression time in each of the 3 time groups as a function of the initial SRF area. There doesn't seem to be any systematic change in the maximum suppression time across SRFs with different initial areas, and this is confirmed if one considers the mean initial SRF area for each group, displayed in figure 4-12a. A more striking relationship is observed if one considers the maximum

suppression time as a function of the probe tone level (figure 4-11b and figure 4-12b). As the temporal gap increased, the SRFs that still displayed suppression were generally those where the probe tone level was lower. If one considers the other measure of the probe tone position in the RF, the distance from the probe tone to the BF, there is a shift present, though it is more subtle than the shift with probe tone level (figure 4-11c and figure 4-12c).

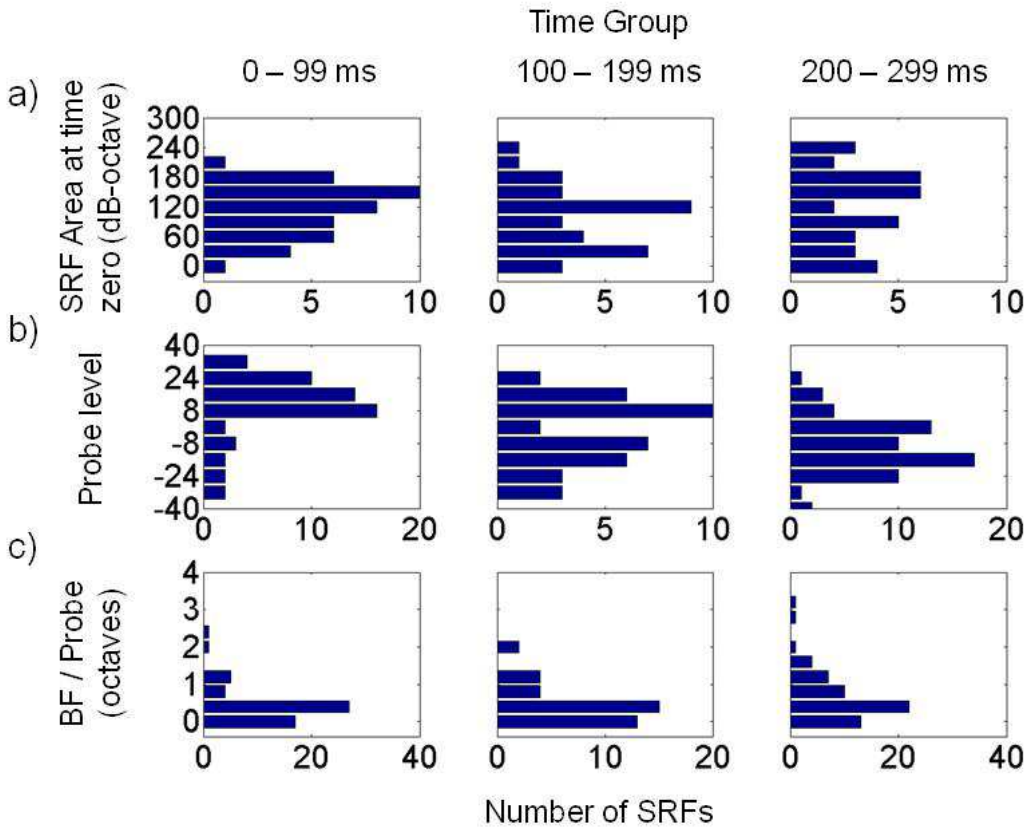


Figure 4-11. Histograms of the maximum time of suppression grouped and displayed as a function of a) the initial SRF area, b) the probe tone level relative to the RF threshold and c) the distance of the probe tone from BF in octaves.

ANOVAs conducted on general linear models with the maximum time of suppression as the dependent variable confirm the above statements.

Firstly, there is no significant relationship between the initial SRF area and the maximum time of suppression ($p >> 0.05$). A second model with probe tone level and frequency as the predictors shows a significant negative correlation between the maximum time of suppression and probe tone level ($p<0.01$, partial $eta^2 = 0.23$), but no significant contribution from the probe tone frequency.

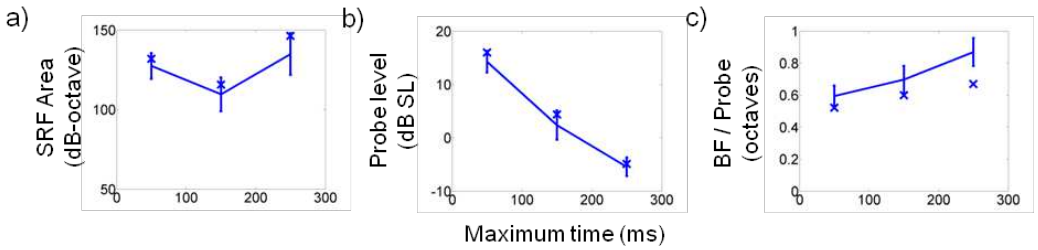


Figure 4-12. The mean values for each maximum time group for a) the initial SRF area, b) the probe tone level relative to the RF threshold and c) the distance of the probe tone from BF. The median values show similar trends to the mean values and are displayed as crosses. Error bars are standard error.

4.3.4. Example of units

In this section I present examples of units that either show or do not show the trends described within this chapter. Figure 4-13 shows the FTCs and SFTCs for 6 units in which there is data for more than one probe condition. The single unit in figure 4-13a displays two facets of the suppression described in this chapter. Firstly, the CSF is close to both probe tone frequencies, far from the CF in both cases. Secondly, the area of suppression is greater for the probe tone shown in red, which has a lower level relative to the FTC. The effects of varying the probe tone frequency in

the single unit shown in figure 4-13b are more subtle. The CSF for both probe tones is at CF. The only effect of increasing the probe tone frequency is to extend the suppressed area on the high frequency side of the SFTC. The

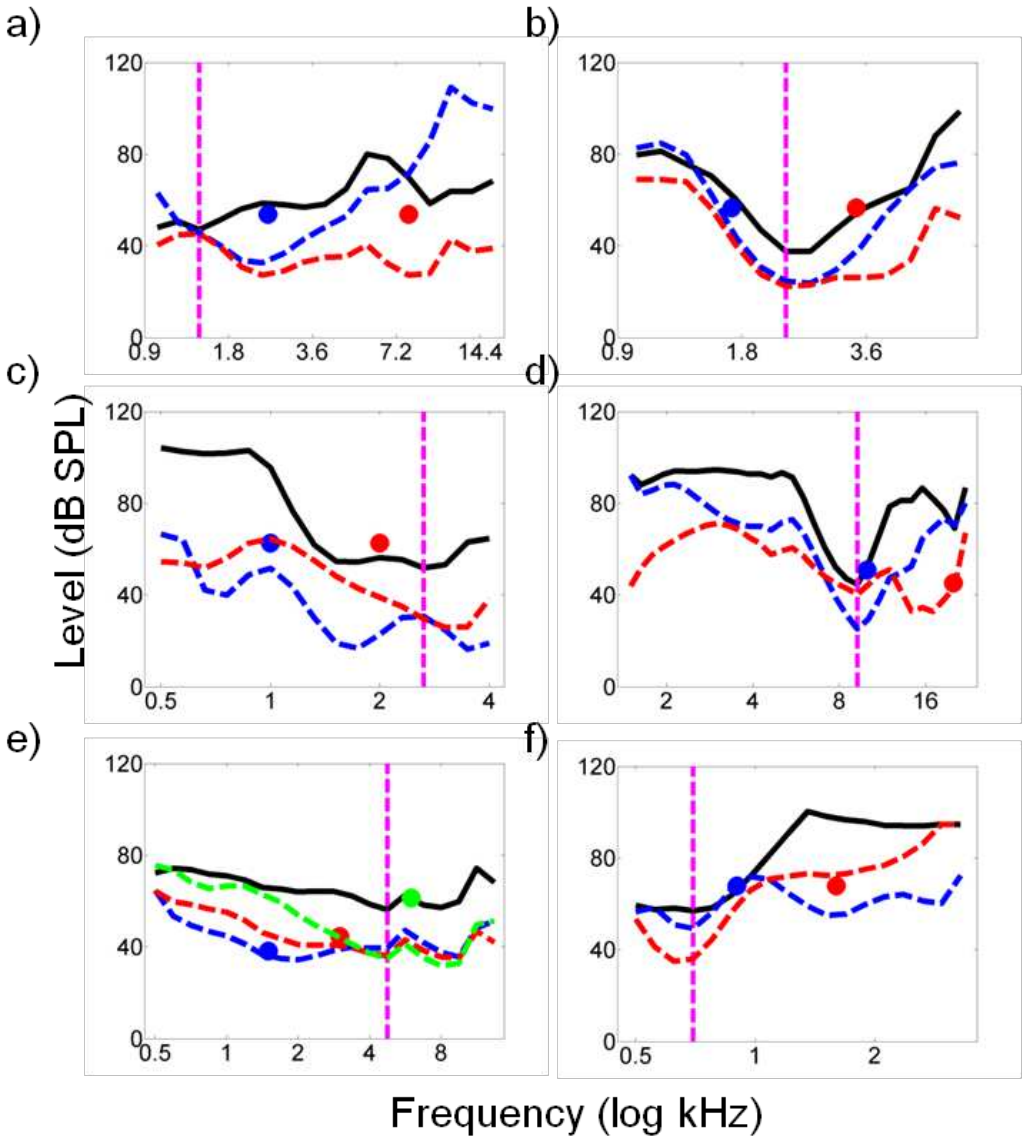


Figure 4-13. Examples of FTCs and SFTCs from 6 units. a-d) Single units. e-f) Multi-units. The FTC for each unit is displayed as a solid black line. The probe tones for each unit are indicated by coloured circles and the SFTCs associated with each probe tone are displayed as dashed lines with the same colour as the probe tone. The CF is indicated by a vertical pink line.

effects of varying the probe tone frequency on the CSF are also subtle in the single unit shown in figure 4-13c, however the suppressed area is again larger for the probe tone with the lowest level relative to the FTC (the blue probe tone). For the single unit in figure 4-13d one probe tone is near to the CF while the other is off-CF. The CSF follows the frequency of the probe tone and there is some evidence of a larger suppressed area for the lower level red probe tone.

The 3 SFTCs for the multi-unit shown in figure 4-13e display a systematic decrease in suppressed area as the probe tone level moves closer to the FTC. The SFTCs are relatively flat but there is some evidence of the CSF increasing as the probe tone frequency increases. The multi-unit in figure 4-13f displays a rare tendency for maximal suppression to occur at the CF regardless of the probe tone frequency. There is also little difference in the area of suppression for each probe tone even though the red probe tone is lower in level relative to the FTC.

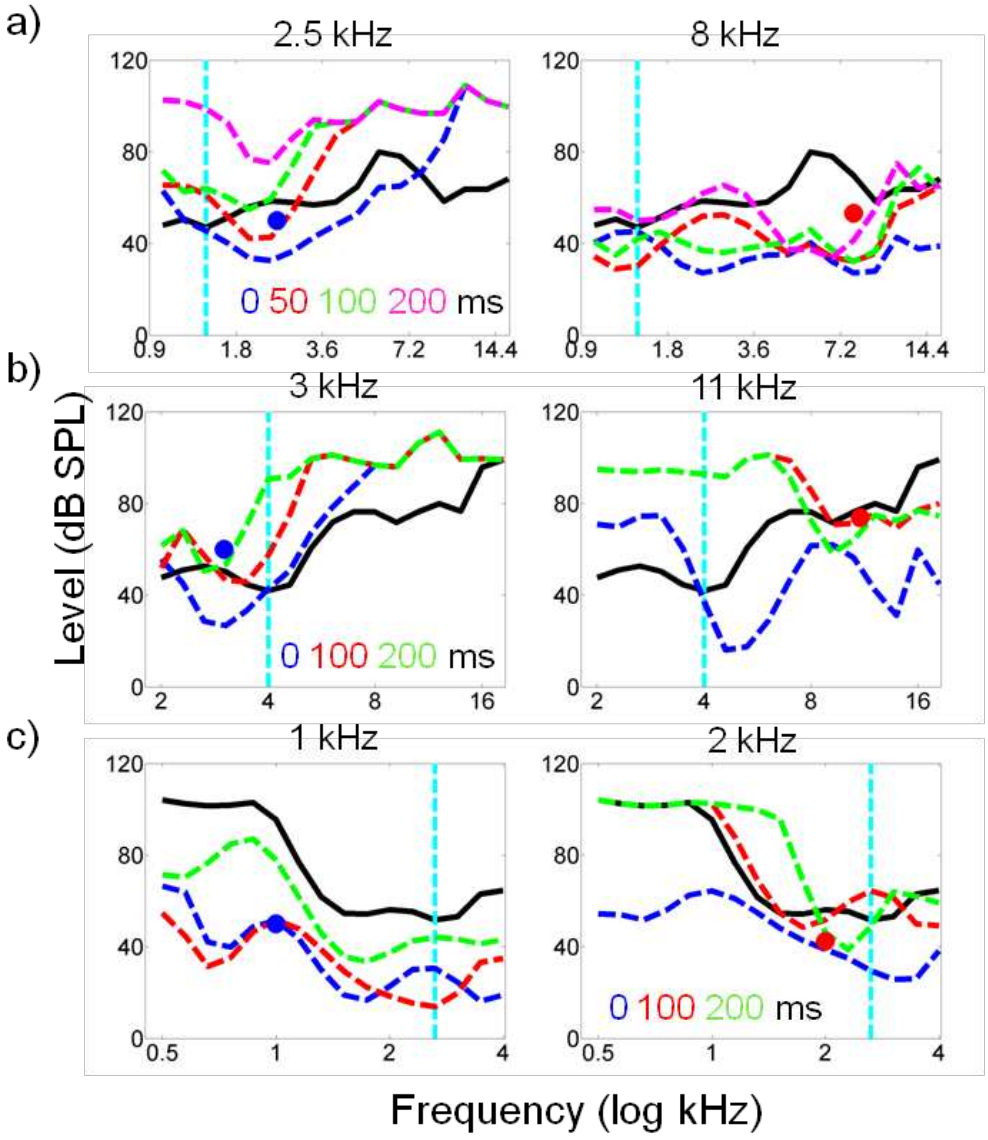


Figure 4-14. Examples of FTCs and SFTCs of single units as the ISI is increased. The FTC is shown as a solid black line. The probe tones are indicated by filled circles and at the top of each panel. The ISI for each SFTC is indicated in the panel for one probe tone frequency for each unit. The CF is indicated by a vertical light blue line.

Figure 4-14 shows how the SFTCs for different probe tones change as the ISI is increased in three single units. For the unit in figure 4-14a the SFTCs contract as the ISI is increased, with frequencies away from the probe tone frequency recovering faster. The SFTC contraction is faster for the 2.5 kHz

probe tone, the level of which is marginally closer to the FTC. This implies that suppression will last longer for the lower level probe tone, as indicated above in the population analysis of the maximum time of suppression. For the single unit displayed in figure 4-14b, the CSF is around the CF when the ISI is 0 ms for both probe tones. However, as the ISI is increased the CSF moves towards the probe tone in the condition where the probe tone frequency is 11 kHz. Thus suppression becomes more specific to the frequency of the probe tone, as indicated in the population analysis described above. The single unit in figure 4-14c is another example where the CSF is around the CF rather than the probe tone frequency, and in this case it does not move close to the 1 kHz probe tone as the ISI is increased. There may be a subtle influence of probe tone frequency on the CSF as it seems to be at a lower frequency for the low frequency probe tone. Again the area of suppression is larger and contracts more slowly for the 1 kHz probe tone which has a lower level relative to the FTC.

4.3.5. Is suppression correlated to the conditioner tone response?

It has been shown in the primary auditory cortex that there is a high correlation between the amount of suppression and the response to the conditioner tone when the probe tone is one of the PBCs in a level-response area (Zhang et al., 2005). However, it has also been noted in IC that neurons can display suppression that is more related to stimulus history than to response history (Malone and Semple, 2001). A salient question is whether

there is still a correlation between the conditioner tone response and the amount of suppression when the probe tone is not one of the preferred frequency-level combinations (at a high level at CF). To answer this question, for each probe condition in 74 units, the Spearman's Rho correlation coefficient was calculated based on the conditioner and probe tone responses at all frequency-level combinations.

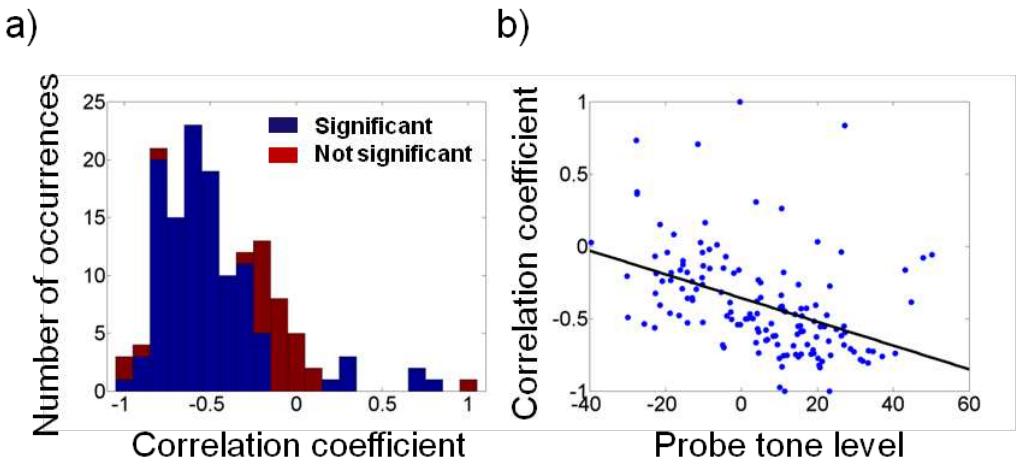


Figure 4-15. Correlations between the response to the conditioner tone and the amount of suppression for each probe tone condition. a) Histogram of the correlation coefficients for each probe tone condition. b) The relationship between the correlation coefficient and the probe tone level. The solid black line shows a linear regression of the data. Both significant and not significant coefficients are displayed (discussed in text).

Figure 4-15a shows a histogram of the correlation coefficients for the 74 units. The purpose of this analysis was to establish whether units tended to show a high correlation or not, so the non-significant values have been left in and are shown as maroon bars. There is a spread of coefficients which is generally larger than those reported previously, when the probe tone is one of the PBC (Zhang et al., 2005). Some units showed a strong relationship

between conditioner response and suppression, but a substantial proportion did not. Figure 4-15b shows that there is a negative correlation between the probe level and the correlation coefficient (m=-0.008, R=0.43, P<<0.01). This indicates that there was more likely to be a correlation between spike counts to the conditioner and probe tones for higher level probe tones. A possible explanation for this is that when the probe tone is at a preferred frequency-level combination the suppression is based on response history, whereas it becomes more based on stimulus history as the probe tone moves away from preferred stimulus combinations. However, this hypothesis would require further study.

4.4. Discussion

4.4.1. Main findings

- The maximum suppression was most often at the frequency of the probe tone rather than at the CF or BF
- This relationship held as the ISI was increased up to ISIs of 300 ms or over
- The range of conditioner tones that suppressed the probe tone increased with the distance of the probe tone from BF and decreased as the probe tone level increased (though see the caveat noted below)
- There is a significant correlation between the probe tone level and the threshold of suppression which may underlie the effects on the range of conditioner tones that lead to suppression

- This relationship is significant up to ISIs of around 300 ms and possibly longer; the paucity of data for ISIs above 300 ms made it unlikely to find any significance

- There is a significant negative correlation between the probe tone level and the maximum time of suppression, indicating that suppression lasted longer for lower level probe tones

4.4.2. The larger extent of suppression of tones to which neurons respond weakly

There was a correlation between the probe tone level and the range of tones that act to suppress it. One caveat of this analysis is that the absolute decrease in spike count that is defined as suppression will be lower for probe tones that elicit a lower spike count. Thus it is more likely for probe tones with lower spike counts (for example, further from CF or with a lower level relative to the FTC) to be defined as suppressed. So, the conclusion that low level probe tones are suppressed by a larger number of conditioner tones may be a reflection of the increased propensity for those tones to be defined as suppressed. It is not known exactly what aspects of neural suppression are important in the processing of tones; for example are absolute decreases in spike count or percentage decreases in spike count of more importance? In fact, the dependence of the suppressed area on the probe level is also evident when an absolute decrease in spike count is considered. The threshold algorithm was performed on a sample of units (59 units yielding 81 SRFs) with probe tone responses between 3 and 7 spikes. With suppression defined as a

drop of 1.5 spikes or more relative to the control probe tone response there was a significant correlation between probe tone level and suppressed area that was similar to that shown in figure 4-9b (m = -1.6, R = 0.36, $p<0.01$). Thus, probe tones with lower levels are suppressed by a wider range of conditioner tones when either the percentage *or* absolute decrease in spikes is considered.

4.4.3. Forward suppression and its relation to streaming

Forward suppression has been implicated as the basis for the effects observed in physiological streaming experiments, though previously it was not possible to directly infer the effect of a non-CF tone ('B' tone) on a CF tone ('A' tone). The results presented in this chapter allow a direct inference to be made. Firstly, the 'B' tones will be suppressed to a larger extent by tones with similar frequencies, so as the frequency difference is increased this effect would decrease. Secondly, lower level probe tones (relative to the RF threshold) are subject to suppression from a larger range of conditioner tone levels and frequencies and for a longer time. As the FD increases the relative level of the 'B' tone will decrease, so we would expect that the range of tones that cause suppression would increase. One issue that has only been partially resolved is the differential suppression of the 'B' tone as the PR is increased: the decrement in the response to the 'B' tone is greater than the decrement in the response to the 'A' tone as the PR is increased. The results from this chapter suggest that this differential suppression occurs because the 'B' tone response is more likely to be suppressed than the 'A' tone response, both by

the response to previous ‘A’ tones and ‘B’ tones (as suppression lasts longer for weaker tones).

4.4.4. The relation of forward suppression to adaptation

Ulanovsky et al. (2003, 2004) found that adaptation was frequency specific to long duration tones (230ms), well separated in time (SOA was usually 736ms) in long probabilistic sequences. Their analysis suggests that a substantial proportion of the frequency specific adaptation was a local effect between tone pairs. The data presented in this chapter establishes that this effect is still robust at even shorter timescales (when there is no gap between the conditioner and probe tones). One potential difference is that we saw no evidence of ‘hyperacuity’ in the tuning of suppression. However, to investigate this, the paradigm that was used here would need to be reversed, with a single conditioner tone followed by tones at a range of frequencies and levels. This is investigated further in chapter 6.

4.4.5. Experimental design and future questions

The paradigm used in this study sought to investigate the effects of varying the probe tone frequency on the suppression pattern that was observed in A1. It is essentially an adaptation of the paradigm that had been used previously to investigate both cortical (Calford and Semple, 1995; Brosch and Schreiner, 1997) and subcortical (Shore, 1995) sequential two-tone interactions. In lower auditory nuclei, the effects of prior stimulation tend to disappear relatively quickly (Harris and Dallos, 1979; Shore, 1995; Finlayson, 1999), while in A1, effects can be seen even after 1 second (Hocherman and

Gilat, 1981; Werner-Reiss et al., 2006). This creates a problem when constructing a paradigm for the investigation of cortical responses: the time between conditions must be as long as possible so that the neuron has had time to recover from previous stimulation. In this study and others (Brosch and Schreiner, 1997) that used a two-tone paradigm, a duty cycle of 1 second was used with the caveat that the results may reflect the characteristics of a habituated system. Even with a duty cycle of only 1 second, there remains a problem with time. The amount of time that a single unit can be recorded from is finite, on average between 30 and 60 minutes. So, there is a compromise between the number of conditions that can be presented and the number of repetitions that can be performed. The key to obtaining reliable response measurements is to attempt to limit the number of conditions and have as many repetitions as possible. In the two-tone paradigm the number of conditions is governed by three factors: the number of conditioner tone attenuations that are presented, the number of conditioner tone frequencies that are presented and the number of probe tones that are presented. In an example experiment the total number of conditions would be around 1080 (9 attenuations from 10-100 dB SPL in 10 dB steps, 15 frequencies covering 2.5 octaves in 0.16 octave steps, 4 time gaps and 2 probe tone conditions). This represents a large number of conditions and in some experiments the single unit that was being recorded from disappeared before enough repetitions could be performed. This led to noisy RFs and SRFs in some cases, especially with data collected when the temporal gap was varied.

So, while the main effects of this study, the locus and area of suppression, were well characterised, it was more difficult to look deeper into questions regarding the separate roles of threshold and bandwidth over time or the maximum time of suppression (because only 1 - 4 time gaps could realistically be presented). The problem of limited time has been addressed previously for large datasets by simply presenting the dataset once and then filtering the SFTCs to remove noise (Brosch and Schreiner, 1997). A similar approach was used here but in order to investigate specific questions like those mentioned above it may be more logical to reduce the number of conditions used and thus increase the number of repetitions. For example, to just investigate the maximum time of suppression for on- and off-CF probe tones, the number of time gaps could be increased if either the number of frequencies or the number of attenuations were reduced. While the power to look at the effects on the whole RF may be diminished, the noise would be reduced to the extent that this specific question could be answered directly.

5. Responses to tone sequences in the anaesthetised guinea pig

5.1. Introduction

As discussed in chapters 1 and 2, streaming (and perceptual organisation in general) can be considered to consist of two processes: primitive and schema-based (Bregman, 1990). The primitive elements are hard-wired and unlearned and from a neurophysiological perspective simply consist of the neural responses. Schema-based elements are learned and are presumably manifested as modulations of the neural activity by higher-level processes. A further distinction between the two facets of streaming is that the primitive elements are considered to be pre-attentive, while the schema-based processes are heavily influenced by attention (Carlyon, 2004). Physiological studies into streaming have mainly been carried out in awake animals with varying degrees of control over attention. It is not clear from these experiments the extent to which attention is affecting the neural responses that are reported. This chapter is aimed at assessing whether the physiological characteristics of streaming are similar in anaesthetized animals, where attention is implicitly absent, to those reported in awake animals. If the characteristics were similar then this would allow further investigation of streaming in the anaesthetized animal model. Streaming has not been exhibited behaviourally in guinea pigs but it seems reasonable to suppose that the primitive mechanisms that Bregman (1990) highlighted are similar in

mammals that have similar peripheral auditory structures. In fact, neural responses in the anaesthetised guinea pig ventral cochlear nucleus have been shown to account for the build-up of streaming (Pressnitzer et al., 2008). This study represents the first investigation of streaming in A1 of the anaesthetised guinea pig.

Where streaming has been investigated physiologically, studies have focussed on one of two aspects of the psychophysical responses to streaming stimuli: the build-up of streaming (Micheyl et al., 2005; Pressnitzer et al., 2008) or the change in response that is associated with varying the FD and PR of the stimulus (Fishman et al., 2001; Bee and Klump, 2004; Fishman et al., 2004; Bee and Klump, 2005). The neurophysiological build-up of streaming has been characterized in single units in the auditory cortex of awake macaque (Micheyl et al., 2005), in multi-units in Field L (the avian analogue of the mammalian A1) in the awake starling (Micheyl et al., 2008) and in single units in the ventral cochlear nucleus (VCN) of the anaesthetised guinea pig (Pressnitzer et al., 2008). In these studies, the neurophysiological build-up of streaming was directly compared to the psychophysical build-up using a model based on signal detection theory (Micheyl et al., 2005). In this chapter, a similar model based on signal detection theory will be applied to the responses to streaming stimuli in A1 of the anaesthetised guinea pig to investigate whether these responses can account for the build-up observed psychophysically.

Investigations into the effect of varying the FD and PR of streaming stimulus sequences have shown that the mean neural responses over the course of the tone sequence can be correlated with the percept that would be elicited by the sequences (whether the sequence would be segregated or integrated perceptually). This correlation is due to the differential mean responses to the 'A' and 'B' tone in the ABAB (or ABA_ABA) sequence under different stimulus conditions. An increase in the FD leads to an increase in the probability that the sequence will be segregated perceptually. Physiologically, an increase in the FD leads to a decrease in the response to the 'B' tone in neurons that respond preferentially to the 'A' tone. Similarly, one would expect that at another cortical position where neurons preferentially respond to the 'B' tone, the response to the 'A' tone would decrease with increasing FD. Thus, there is a spatial separation of coding that could correlate to the segregation of the tone sequence stimulus into two streams. This theory of neurophysiological streaming based on tonotopy extends to the effects of varying the PR on the streaming perception. So, while the response to both the 'A' and 'B' tones decreases with an increase in the PR, the decrement in response to the 'B' tone is greater such that the relative response of the 'B' tone to the 'A' tone decreases as the PR is increased (differential suppression of the 'B' tone). This correlates with an increased probability of a segregated percept as the PR is increased.

The effect of varying the FD and PR has only been investigated systematically with either multi-unit activity (Bee and Klump, 2004) or

ensemble neural activity (multi-unit activity and current-source density) (Fishman et al., 2004). One aim of this chapter is to investigate whether the effects of varying the FD and PR in single neurons of the anaesthetised animal are similar to those reported in multi-units in the awake animal. For example, given that the frequency tuning of neurons in the primary auditory cortex is generally similar (Sutter, 2000; Moshitch et al., 2006) under anaesthetized and awake conditions, it may be expected that the decrease in the response to the 'B' tone with increases in the frequency difference would also be observed in the anaesthetized preparation. One caveat is that the tuning bandwidth tends on average to be greater in anaesthetized A1 responses (Schreiner and Sutter, 1992; Qin et al., 2003). It also may be expected that increasing the presentation rate will lead to a differential suppression of the 'B' tone response, as forward suppression has been shown to be similar in awake and anaesthetised preparations (Brosch et al., 1999; Brosch and Scheich, 2008). However, in anaesthetized preparations cortical neurons have been shown to struggle to phase-lock to stimuli presented at rates above between 10-20 Hz. It might then be expected that phase-locked responses to streaming stimuli in single neurons of the anaesthetized animal would not be observed at 20 and 40 Hz, in contrast to the responses observed in neural ensemble activity in awake animals (Fishman et al., 2004). So, a further aim of this chapter is to investigate how this decrease in phase-locking ability will impact on the differential suppression of the 'B' tone.

The signal detection model has only previously been applied to the build-up of streaming; it has not been applied to the steady-state responses (the mean responses to tones in the sequence after the initial adaptation of activity) as the FD and PR are varied. So, the final aim of this chapter was to investigate whether the signal detection model could be used to relate the changes in guinea pig neurophysiological responses related to varying the PR and FD with the human perceptual changes that are observed.

5.2. Methods

5.2.1. Stimuli

The stimulus that was used in this study was the ABAB tone sequence (50 ms tones with 4.5 ms linear ramps) used previously to investigate the effect of varying the FD and PR (Fishman et al., 2004). A level was chosen at which the neuron under investigation reliably responded to and at which there was a gradient of response from the BF to the edge of the RF. The 'A' tone was then set to a frequency at or near to the BF at that level and the 'B' tones were spaced at equal intervals away from the 'A' tone in either ascending or descending frequency steps (generally at 0.125, 0.25 or 0.5 octave steps depending on the bandwidth of the RF at that level). For a given condition the stimulus was selected randomly from the range of 'B' tone frequencies and a range of presentation rates (2, 4, 8, 12 or 16 Hz). This range was chosen after pilot data showed that many cortical neurons failed to lock to the onset of tones for a PR of 20 Hz (as was used in Fishman et al., 2004). Due to the limited amount of time available for single unit recording, the full

range of PRs was not tested for each unit. Also, due to the use of MEAs, several units could be recorded simultaneously. Although care was taken to align the electrode along an isofrequency contour so that neurons with similar tuning properties were recorded from, in many cases for secondarily recorded units the 'A' tone was not at the BF. For the purposes of population analysis of streaming, only those units for which the 'A' tone elicited the maximal spike count were used.

5.2.2. Data analysis techniques

5.2.2.1. Period histograms

As will be shown in section 5.3.2, the responses to tones within the ABAB sequence decrease or adapt as the sequence continues. This initial adaptation may be involved in the build-up of streaming (Micheyl et al., 2007). To simplify the assessment of changes in the responses of neurons to the streaming stimuli at different PRs and FDs, and to align the results with the study on which the current stimuli are based (Fishman et al., 2004), the adaptation will be considered separately to the overall response of the neuron. The steady-state responses (which occur after the adaptation) will be examined by constructing period histograms (see below) of the responses after 2 seconds of stimulus presentation. The response to each tone pair was calculated by counting the number of spikes within a time period corresponding to the presentation of two tones. For example, for the 2 Hz condition, the spikes were counted over one second periods, while for the 4 Hz condition the spikes were counted over 0.5 second periods. The responses

to each tone pair were then averaged for each condition. Statistics, such as the ‘A’ and ‘B’ response spike counts, were calculated by counting the number of spikes within the analysis time window (described in the methods chapter) in the period histograms. In order to display period histograms for different presentation rates on the same axis, the time axis will be shown in phase with a period of pi. However, statistics were derived from period histograms with the time axis expressed in seconds. Where population period histograms are displayed, in order to normalise for different response latencies, the period begins at the start of the response as defined from the analysis time window.

5.2.2.2. The vector strength and Rayleigh statistic

To characterise the extent to which unit responses were locked to the tone onset, the vector strength (Goldberg and Brown, 1969) for each condition was calculated. Initially, period histograms with a period of one tone were constructed for each condition. To investigate tone-locking in the steady-state response, the period histograms started after 2 seconds of stimulus presentation. Then, a fast-Fourier transform (FFT) was performed on the period histograms for each condition in order to view the data in the frequency domain. The vector strength was then calculated by dividing the 1 Hz component of the FFT with the 0 Hz component (the total number of spikes). The vector strength represents the fraction of elements in a signal that occur at the same time. So, it is applied in this context with the assumption that neurons that lock to the tone onsets will have spikes

occurring close in time and will thus have high a vector strength value (figure 5-1a). Conversely, those neurons that do not lock to the tone onset will have spikes that occur randomly and so will have a low vector strength value (figure 5-1b).

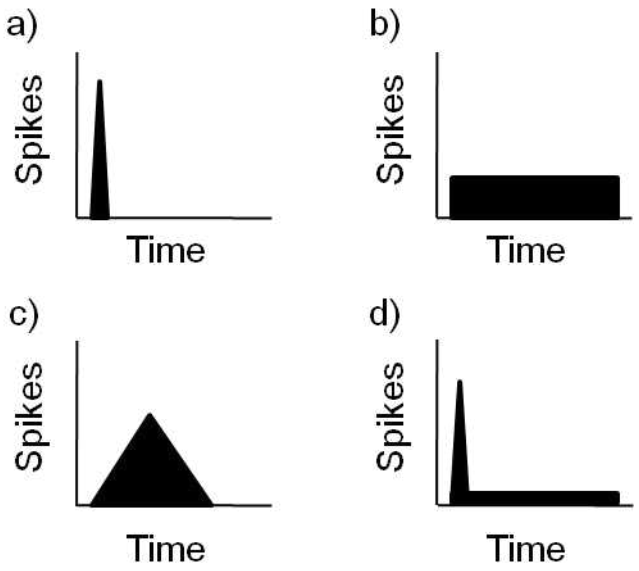


Figure 5-1. Example period histograms.

One assumption when using the vector strength as a measure of tone-locking is that neurons lock to the *onset* of tones. Neurons with sustained responses where spikes are spread over time, such as the example in figure 5-1c, will have lower vector strengths. Also, any spontaneous activity, which increases the total number of spikes, will tend to bring the vector strength value down (for example, figure 5-1d).

To assess the statistical significance of the vector strength, the number of observations (or spikes) used to calculate it must be taken into account. The Rayleigh test of uniformity can be used to show that a period histogram is

a sample of an oriented distribution rather than a uniform distribution, using the factor $2nR^2$, where n is the total number of spikes and R is the vector strength. If the value of $2nR^2$ is larger than 13.8 then the probability that the distribution is uniform is less than 0.001 (Buunen and Rhode, 1978; Mardia and Jupp, 2000). In this chapter, when vector strength is used to constrain the data, only those conditions with a Rayleigh statistic larger than 13.8 will be considered.

5.2.2.3. The adaptation of spike count to successive tones

To investigate how the response to the tone sequences changed over time the mean spike count for each tone in the sequence was calculated. This was achieved for each unit by counting the number of spikes within the analysis time window in response to each of the tones in the sequence. So, at 2 Hz, the response to the first 'A' tone would be counted within the analysis time window after 0 ms, the response to the first 'B' tone would be calculated within the analysis time window after 500 ms, the response to the second 'A' tone would be counted within the analysis time window after 1000 ms, and so on. The spike counts for each unit were then normalised for the number of stimulus presentations in each condition.

5.2.2.4. Grouping into 5 frequency groups for population analysis

When recording from a unit, the frequency differences were set so that they effectively spanned the edge of the RF, falling in response as the

distance from the BF increased. Due to RF differences it was not always possible to use the same frequency differences and indeed the same frequency steps in each unit. In order to analyse the population data it was necessary to group the frequency differences. After assessing at which frequency differences the majority of the data were collected, the data were grouped into quarter octave bins from 0 to 1 octave frequency difference. The data at only one frequency difference from each unit could be added into a frequency difference group. Also, frequency differences from below the BF were grouped with those from above the BF. This seemed reasonable based on visual analysis of the data and previous reports that there was no difference between streaming results collected above and below BF (Fishman, 2004).

5.2.3. Signal detection model

One way of quantifying the response to tone sequences is to count the average number of spikes in response to a tone within the sequence. While spike counts can provide important information about the operation of neurons, relating them to psychophysical data can be problematic. One solution to this is to apply signal detection theory to the responses. So, instead of considering the absolute spike counts we consider the *distributions* of spike counts across each presentation of a given stimulus condition. Due to the inherent variability in neural responses, neurons do not respond with exactly the same number of spikes on each presentation of an identical stimulus. An example of an idealised spike count distribution is displayed in

figure 5-2a as black circles, along with a normal fit to the distribution (dashed blue line). The proportion of trials over which the spike count will exceed a fixed threshold (displayed as the grey area in figure 5-2b) will vary probabilistically. This provides the basis for establishing the probability that a neuron responded to the 'B' tone for a given condition and thus the probability of a single stream response.

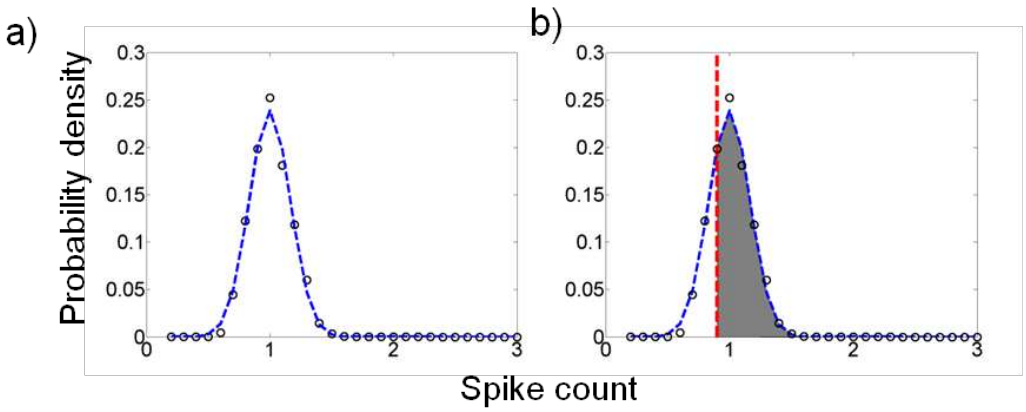


Figure 5-2. Example of an idealised spike count distribution. a) Spike count distributions (black circles) fitted with a normal distribution (dashed blue line). b) Applying a threshold (dashed red line) to the fitted spike count distribution allows an estimation of the probability that a spike occurs. This would be calculated by dividing the grey area above threshold by the total area under the fitted curve.

The spike count distributions for single units are relatively noisy, due to the limited number of presentations. In order to assess the probability of a one stream response for the population, the spike count distributions for each tone were convolved to create a population spike count distribution. As the number of units included was different for each condition the distributions were normalised by dividing by the number of included units.

Each distribution was then fit with a normal distribution, as shown by the dashed blue line in figure 5-2a. This gave two parameters for each condition – the mean and the standard deviation of the spike count distribution.

The signal detection model was used to investigate both the build-up of streaming and the steady-state responses. For the build-up of streaming, the spike count distributions were derived from the response to each tone in the sequence in each condition. The threshold was then set to provide the closest fit, as assessed visually, between the neurometric functions (the probability of the response to the 'B' tone, derived as described above) and the psychometric functions. The threshold was the only free parameter in the model. Although it was set to maximise the fit between the two functions, it was the same for each condition. In order to smooth the neurometric functions both the mean and the standard deviation from the normal distributions were fitted as a function of time for each condition. The data were fitted with double exponential, single exponential, power law and linear functions. The sum of the squared error was calculated for each fit and used to choose which function was most appropriate to fit the data with (see section 5.3.3.2 for a further description).

To investigate the steady-state response, the spike count distributions were derived from the responses to all tones between 2 and 7 seconds for a given condition. As the model has not been used before to account for the steady-state response there is no psychophysical data with which it can be

compared directly. Instead, I investigate the limits over which the threshold can be set given certain criteria. This will be elaborated in section 5.3.3.1.

5.3. Results

5.3.1. The steady-state response to tone sequences

5.3.1.1. Single unit examples

Figure 5-3 shows the receptive fields for 2 example single units with the frequency and attenuation of the 'A' tone highlighted by a black rectangle and the range used for the 'B' tones highlighted by a white rectangle. In the unit in figure 5-3a the 'A' tone is close to, but not at, the BF of the neuron. The unit in figure 5-3b has a low firing rate. Nevertheless, the 'A' tone is in the region of the receptive field that the neuron preferentially responds to.

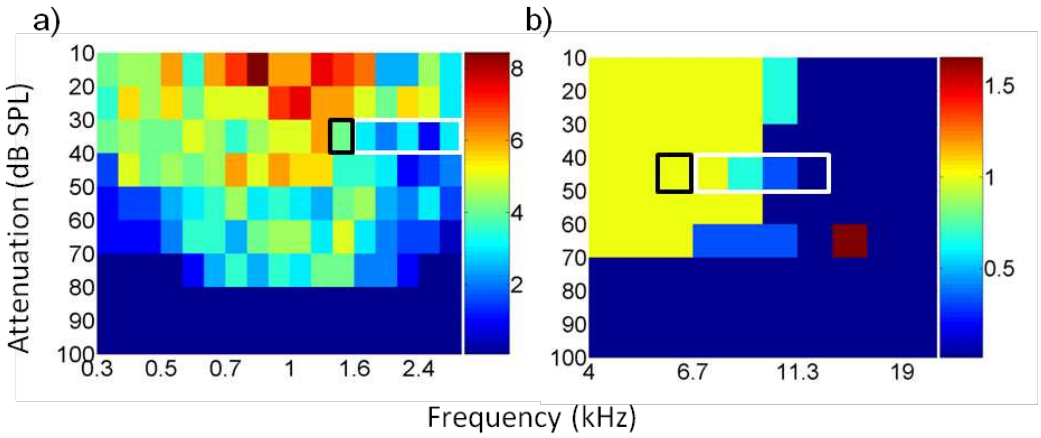


Figure 5-3. Receptive fields for 2 single units. The number of spikes per presentation is shown; warmer colours indicate a higher spike count. The position of the 'A' tone is indicated by a black rectangle and the range of 'B' tones is indicated by a white rectangle.

Period histograms for the example single units are displayed in figure 5-4 and figure 5-5, with the frequency difference increasing from left to right and the presentation rate increasing from top to bottom. The period histograms display several of the features noted in awake preparations (Fishman et al., 2001; Fishman et al., 2004). Firstly, at all presentation rates, as the frequency difference is increased the response to the 'B' tone decreases. Secondly, as the presentation rate is increased there is a general decrease in the response to both the 'A' and the 'B' tones. However, the response to the 'B' tone decreases to a larger extent than the response to the 'A' tone so that at the fastest PR (12 Hz in figure 5-4 and 16 Hz in figure 5-5), the neuron only shows an obvious response to the 'A' tone.

A third observation is that the response to the 'A' tone increases as the frequency difference increases. This is can be observed at presentation rates of 4, 8 and 12 Hz in figure 5-4 and 8, 12 and 16 Hz in figure 5-5. In fact, if one refers to the 8 Hz condition in each example unit, it seems that the 'A' tone response increases with an increase in the frequency difference and then begins to decrease again as the frequency difference is increased further.

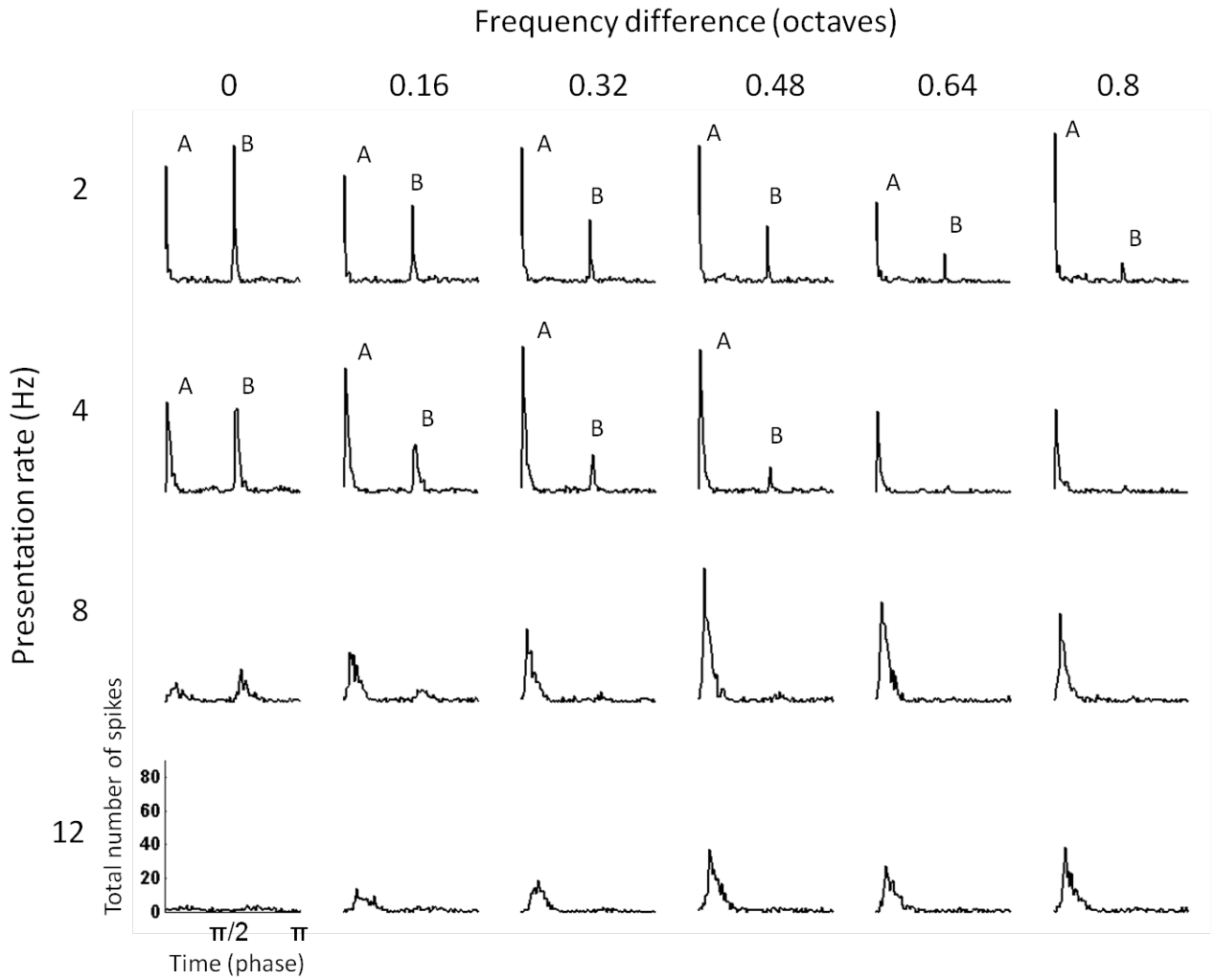


Figure 5-4. Period histograms for a single unit, the RF of which is displayed in figure 5-3a. The frequency difference is indicated above each column and the presentation rate is indicated beside each row. The axis for each period histogram is the same and is displayed for the 12 Hz / 0 octave condition.

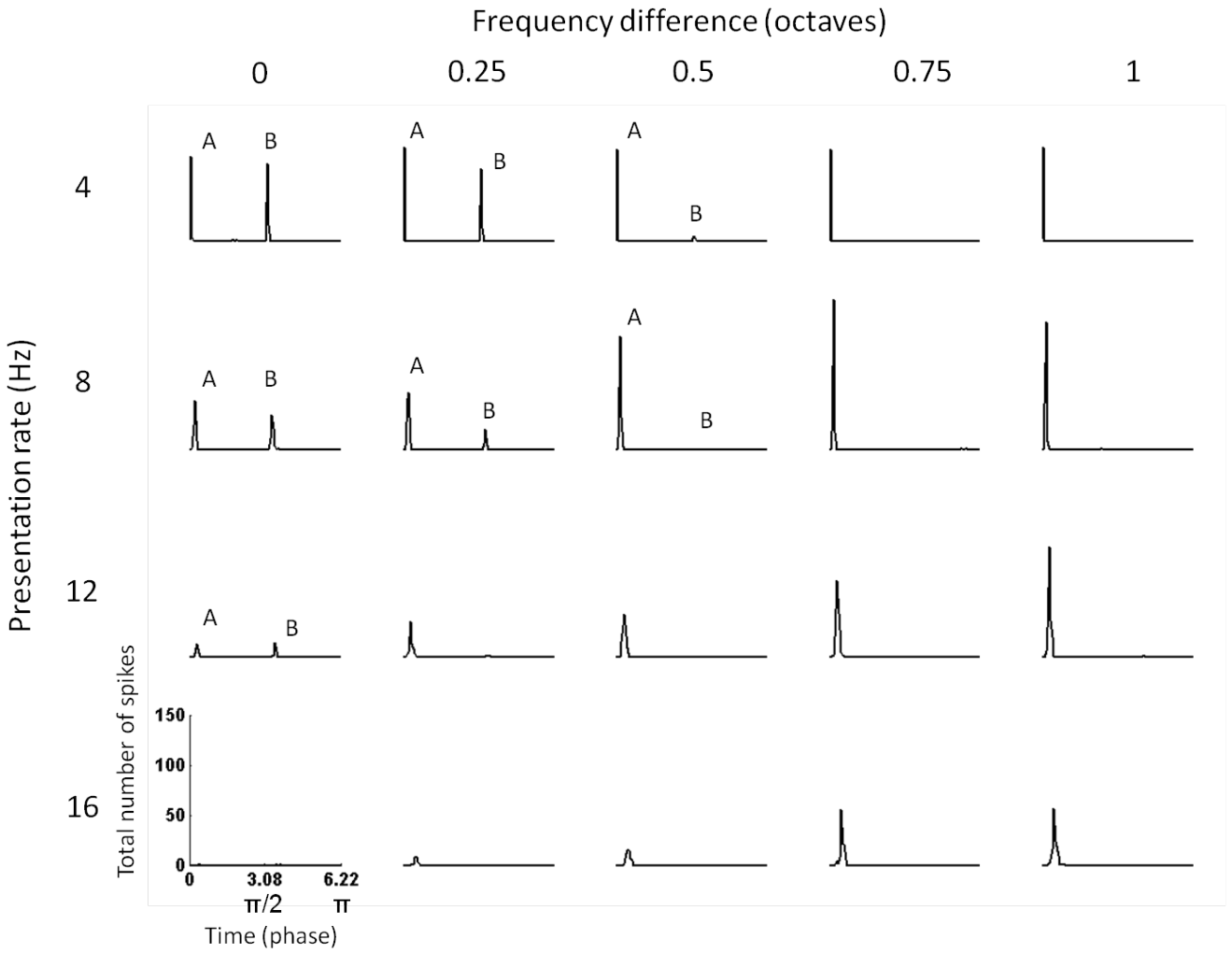


Figure 5-5. Period histograms for a single unit, the RF of which is displayed in figure 5-3b. Same conventions as figure 5-4.

In Fishman et al. (2004), the ratio of the 'B' tone response amplitude and the 'A' tone response amplitude was used to quantify PSTH results for the population data. In figure 5-6, the spike count response ratio is shown as a

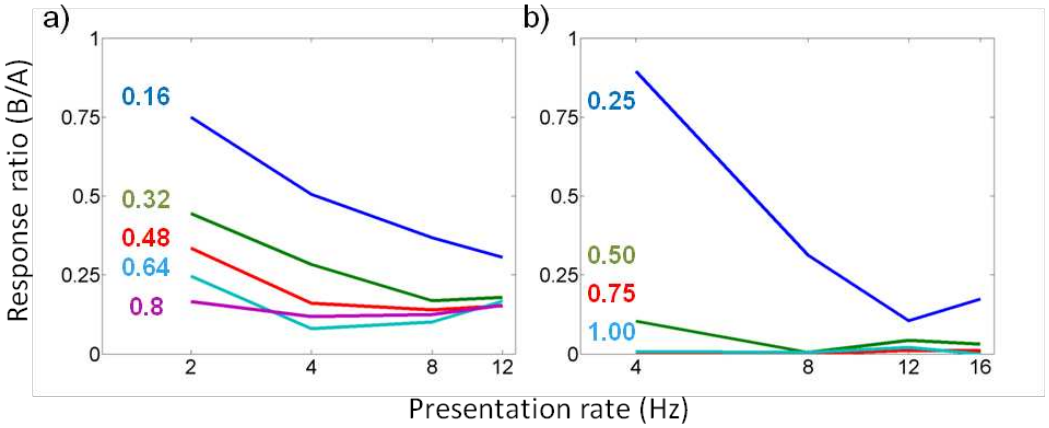


Figure 5-6. Response ratio ('B' response / 'A' response) as a function of presentation rate with frequency difference as the parameter for a) the SU shown in figure 5-4 and b) the SU shown in figure 5-5. Frequency difference values are colour-coded and are expressed in octaves.

function of presentation rate for both unit examples. Similar trends can be observed in both anaesthetised guinea pig single units as were observed by Fishman et al. (2004) in neural ensemble activity in an awake macaque. In general, the response ratio decreased as the presentation rate increased. However, the ratio increased for the faster PR conditions. This may be due to a decrease in the signal to noise ratio as the response to each tone decreases with PR, and is considered in more detail in the population responses. The response ratio also decreased as the frequency difference was increased, as displayed in figure 5-7.

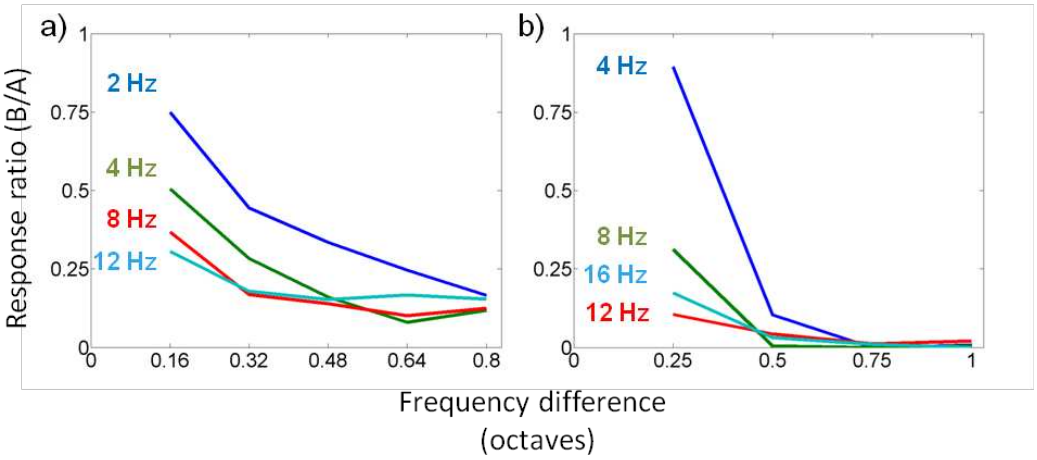


Figure 5-7. Response ratio ('B' response / 'A' response) as a function of frequency difference with presentation rate as the parameter for a) the SU shown in figure 5-4 and b) the SU shown in figure 5-5.

In summary, I have shown evidence that the neural responses of single units of an anaesthetised guinea pig can be consistent with steady state behavioural responses to the streaming stimulus. Also, the neural responses are consistent with the behavioural responses over a wide range of stimulus conditions. This provides further support for the notion that consciousness is not necessary for the underlying neural elements of streaming to be present. In the following section I present evidence that this is also the case when the population of units is considered.

5.3.1.2. Population data

In total, 106 units were recorded from guinea pig A1. However, due to the nature of multi-unit recording allowing simultaneous recording of more than one unit with different RFs, not all units responded maximally to the 'A' tone. For the analysis of streaming, only those units that displayed maximal

firing to the 'A' tone at a 4 Hz presentation rate were included (4 Hz was chosen because it was a condition that was presented in every unit). Table 5-1 displays the total number of units recorded from and the number of units used in the streaming analysis (in brackets) for the guinea pig population.

Single unit	**Multi-unit**	**TOTAL**
55 (30)	51 (30)	106 (60)

Table 5-1. Total number of units separated by unit type. The number of units used in the streaming population analysis are indicated in brackets.

5.3.1.2.1. Steady-state responses

The average period histograms for the population are displayed in figure 5-8. They display similar trends across frequency difference and presentation rate that were observed in the single unit. However, one noticeable aspect of the population period histograms is that the baseline activity tends to increase relative to the response as the PR increases, such that it is difficult to tell the response from the baseline activity. One possible reason for this, that was mentioned in the introduction, is that as the rate is increased the number of units which phase-lock to the stimulus decrease. Thus, as the PR is increased, the population period histogram reflects a decrease in the relative contribution of units with phase-locked responses and an increase in the proportion of units that contribute either nothing (stop responding) or noise that is not associated with the stimulus. To test this

hypothesis, only units that reliably locked to the stimulus were selected, given that it is unlikely that units that did not lock to the stimulus would contribute to the coherent or segregated percept. As described in section 5.2.2.2, the Rayleigh statistic of each unit was calculated based on the period histograms for each condition and it was then used to remove conditions where the vector strength was not statistically reliable.

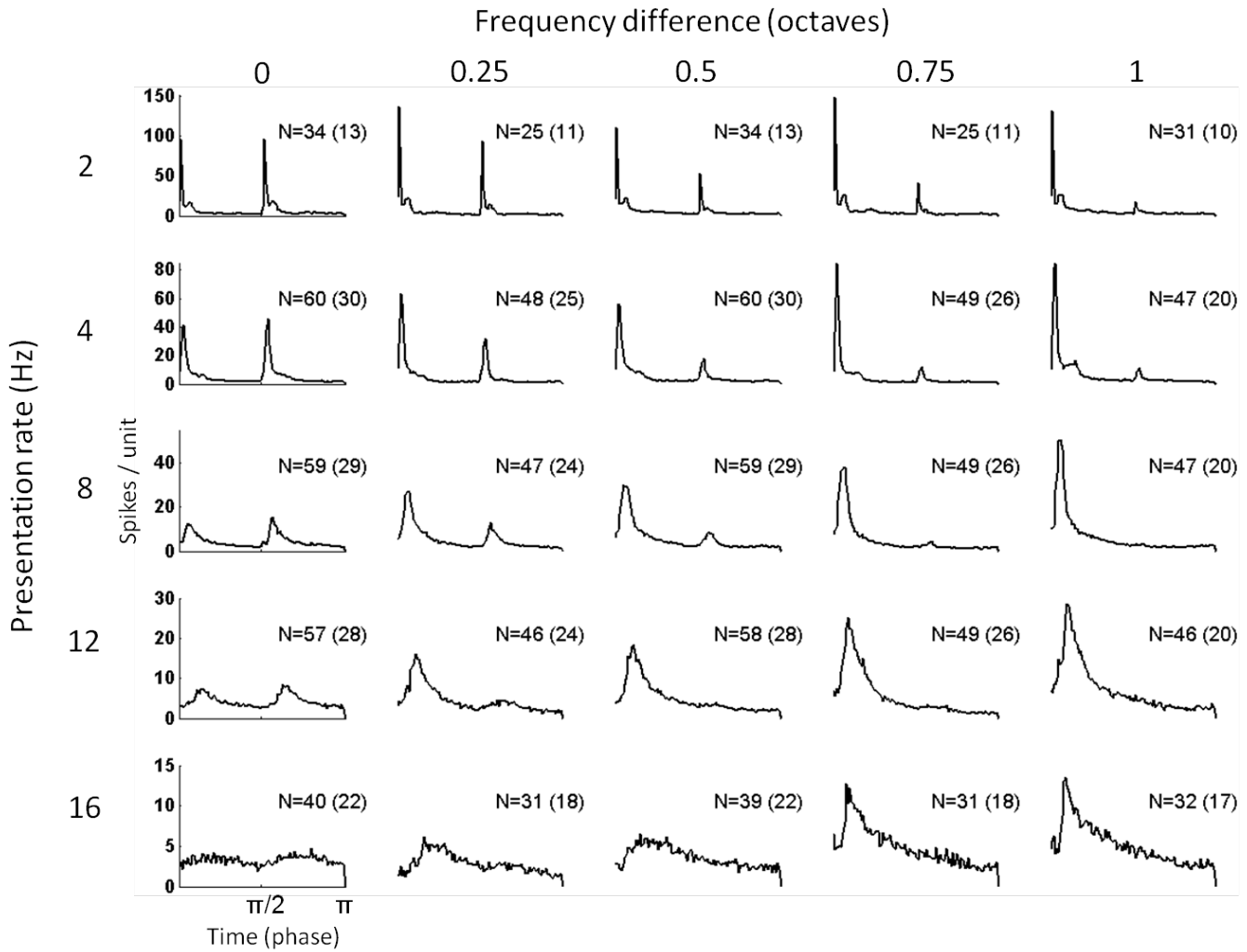


Figure 5-8. Mean population period histograms. The frequency difference is indicated above each column and the presentation rate is indicated beside each row. Note that different ordinate axes are used for each PR. The number of units that were averaged (N) is indicated for each condition, with the number of single units in brackets.

Figure 5-10 shows a comparison between the mean population period histograms for the raw data (shown with a solid blue line) and when only those units with significant vector strengths above 0.4 have been included (dashed red line). A vector strength of 0.4 was chosen after visual inspection of the period histograms and the statistics that resulted from setting the criterion at several values. The distribution of vector strengths is shown in figure 5-9. A compromise between removing the noise and reducing the number of units had to be met, with 0.4 giving a reasonable result. The period histograms show that the noise that was apparent previously has been reduced. This is especially noticeable in the 16 Hz conditions where distinct responses to the tones can be observed after the removal of the noise. Thus, by removing those units that do not phase-lock to the stimuli, the evidence of responses to the tones is increased, leading to a greater resemblance between the population response and the psychophysics.

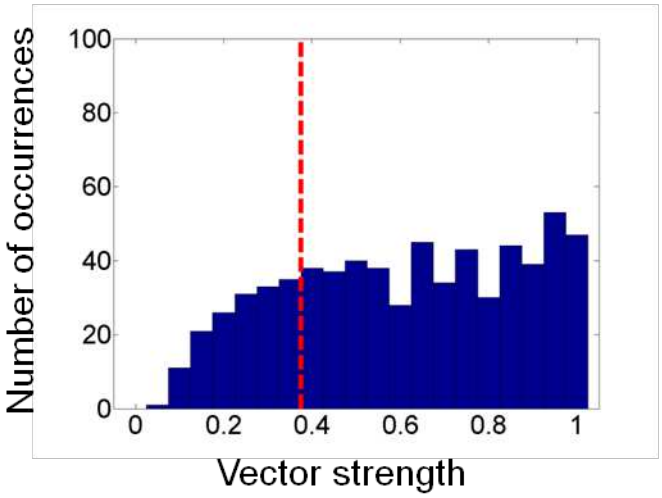
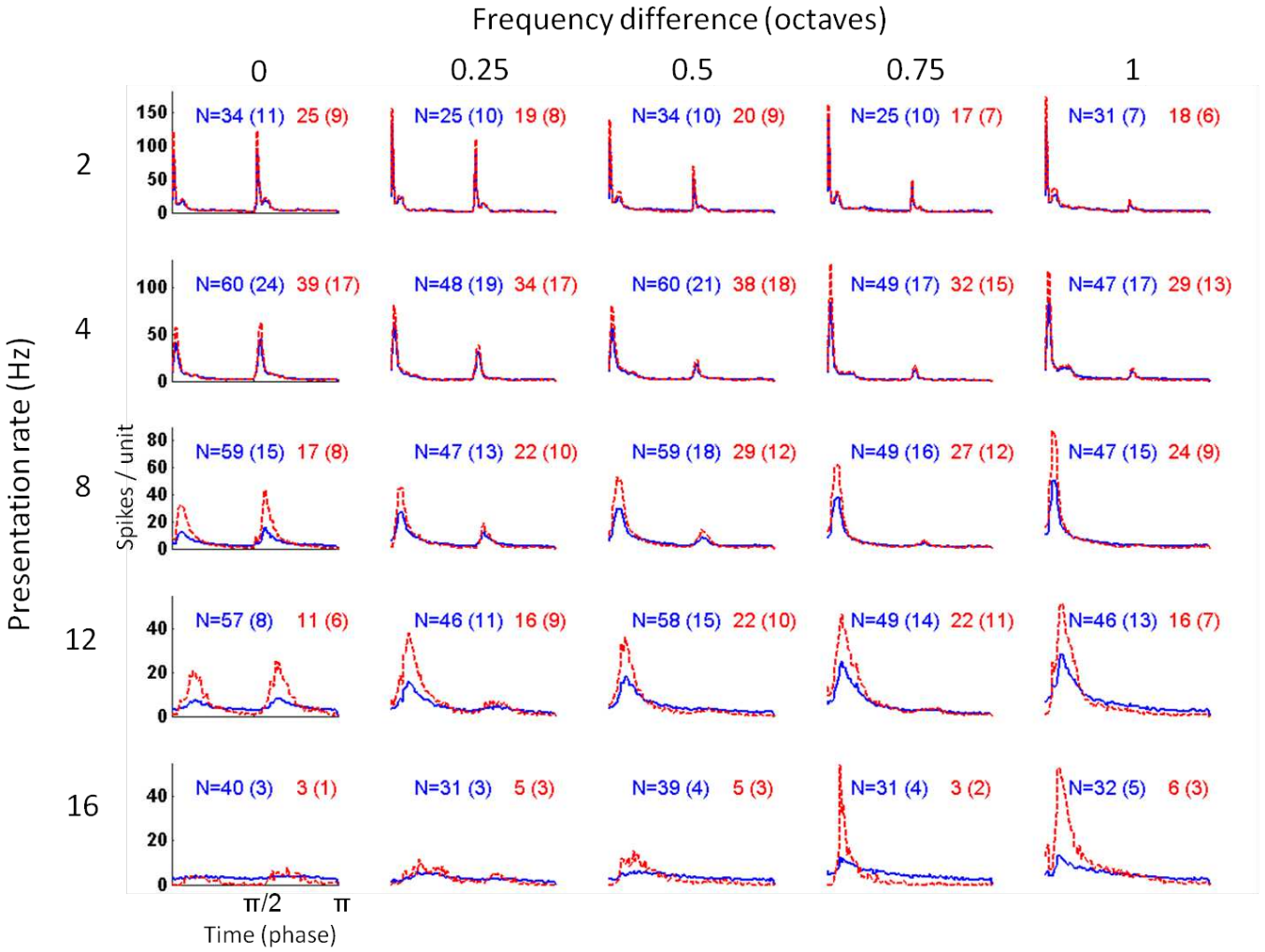


Figure 5-9. Distribution of vector strengths for conditions where the vector strength was significant. The red dashed line indicates a vector strength of 0.4.

Figure 5-10. Mean population period histograms before (solid blue) and after (dashed red) individual period histograms with vector strengths of less than 0.4 have been removed. The frequency difference is indicated above each column and the presentation rate is indicated beside each row. Note that different ordinate axes are used for each PR. The number of units (N) that were averaged is indicated both before (blue) and after (red) data removal, with the number of single units indicated in brackets.



The response ratios associated with the histograms in figure 5-10 after removal of those with a significant vector strength lower than 0.4 are displayed in figure 5-11. Figure 5-11a shows the mean of the response ratios from individual units, with the standard error of the mean indicated by the

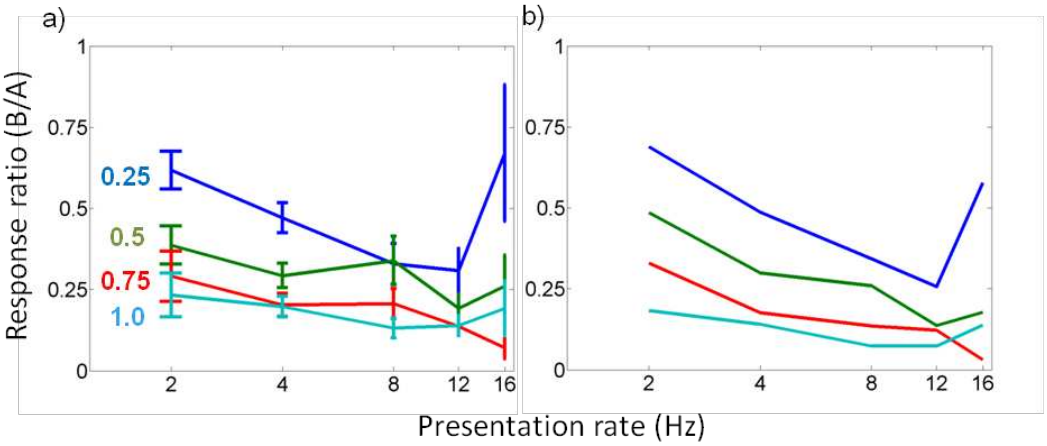


Figure 5-11. Mean population response ratio plotted as a function of PR and with frequency difference as the parameter a) The mean response ratio calculated by averaging the ratios for each unit (error bars represent standard error of the mean). b) The mean response ratio calculated by dividing the mean population response to the 'B' tone by the mean population response to the 'A' tone.

error bars. For some data points the standard error of the mean is large, indicating that the response ratios from individual units tended to be noisy. Figure 5-11b shows the population ratio when it has been calculated as the mean population response to the 'B' tone divided by the mean population response to the 'A' tone. The ratios show similar trends in the population data as they do in the single unit data: a decrease in value as either the frequency difference or presentation rate is increased. This is true up until 16 Hz, where the ratios are generally larger than expected. This seems to be due both to

poor phase-locking and the response to each tone becoming more similar at this rate.

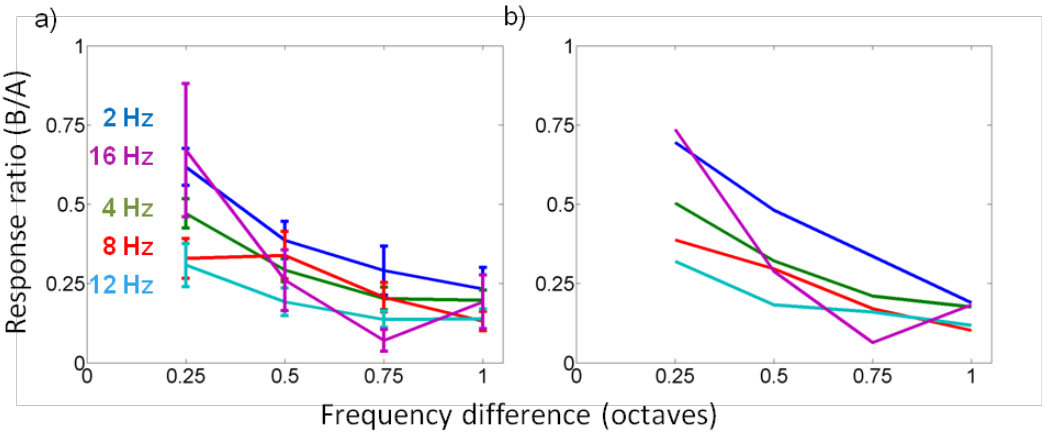


Figure 5-12. a) and b) Same data as in figure 5-11 but plotted as a function of frequency difference and with PR as the parameter.

Figure 5-12 shows the same data as in figure 5-11 but plotted as a function of the frequency difference. Again, the mean response ratios from individual units are displayed in figure 5-12a, while the response ratios calculated by dividing the mean 'B' tone responses by the mean 'A' tone responses are plotted in figure 5-12b. There is a general decrease in the response ratio as the frequency difference is increased, in line with the results from the single unit example.

The increase in the response to the 'A' tone as the frequency difference is increased can be seen in the guinea pig population data as well as in the guinea pig single unit example (the population averages are shown in figure 5-13a). It can be observed at all PRs but increases in magnitude as the

PR is increased (as indicated by the percentage increases displayed in figure 5-13b).

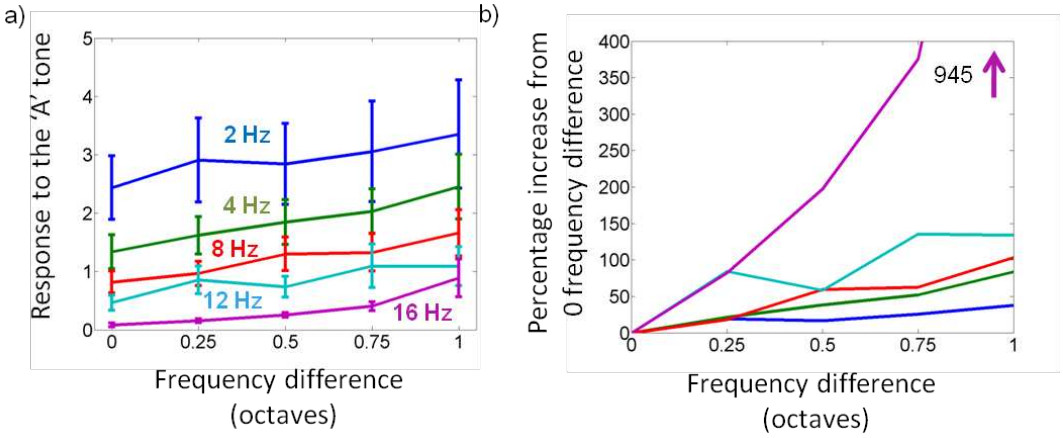


Figure 5-13. The mean population response to the 'A' tone as a function of frequency difference and with PR as the parameter (error bars represent standard error of the mean). a) The absolute spike counts and b) the percentage increase from the 0 octave frequency difference condition.

5.3.2. Adaptation of spike count to successive tones

The above analysis focussed on the 'steady-state' or adapted responses to the stimuli, with a similar analysis to that used previously in awake preparations, albeit using spike counts instead of peak amplitude (Fishman et al., 2004). Another aspect of the response to streaming stimuli is how neural responses adapt towards the 'steady-state'. The mean spike counts for the population as the sequence unfolds are displayed as circles (blue for the 'A' tone and red for the 'B' tone) in figure 5-14. In general, responses decreased or adapted over the course of the tone sequence, though this is less obvious at a PR of 2 Hz. To characterise the adaptation, exponential functions were fitted to the data, displayed as dashed lines in

figure 5-14. For clarity, the fits have been artificially increased by 3 spikes / presentation.

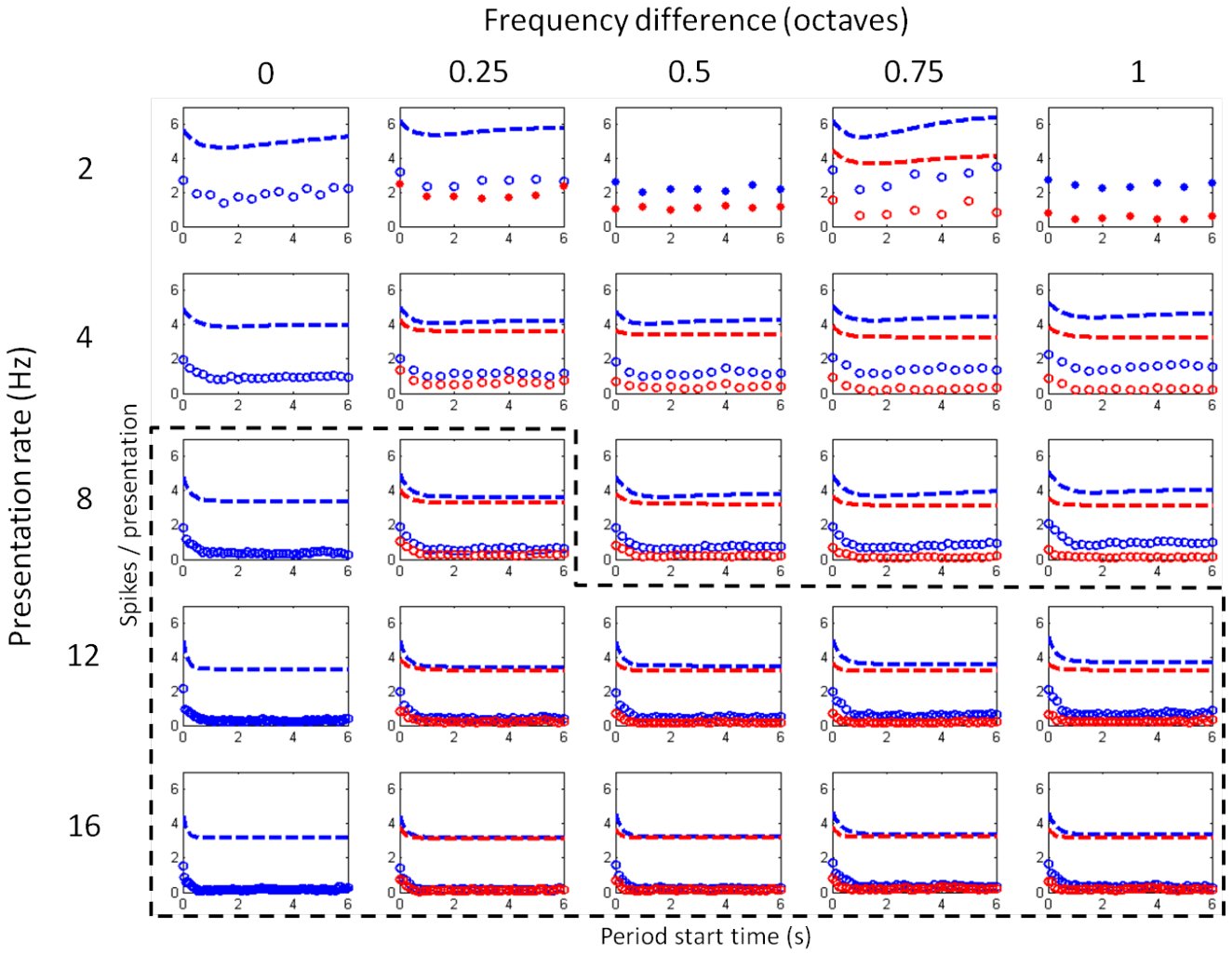


Figure 5-14. Mean population spike counts in response to tone sequences (circles) with exponential curves fitted (dashed lines). The spike counts and fits for the 'A' tone are shown in blue, while those for the 'B' tone are shown in red. In the panels surrounded by the dashed line both the 'A' and 'B' tone data were sufficiently fit with single exponential curves. In the other panels data were fit with a double exponential. The fitted curves have been shifted up by 3 spikes / presentation for clarity.

The adaptation time constants are displayed in table 5-2 for the 'A' tones and in table 5-3 for the 'B' tones. In general, adaptation was well characterised using a single exponential fit of the data (conditions shown as a shaded grey area in table 5-2 and table 5-3):

$$y = A + Be^{-t/\alpha}$$ **Equation 5-1**

However, in some conditions there seemed to be a recovery from adaptation which was best characterised with the sum of two exponentials, an initial decaying exponential followed by a growing exponential (examples of this can be seen in the 2 Hz condition of figure 5-14):

$$y = A + Be^{-t/\alpha} + De^{-t/\beta}$$ **Equation 5-2**

In other cases, little adaptation was displayed and so no function was fitted as the fitted parameters would not be meaningful.

In general, for the 'A' tone, at the faster PRs (12 and 16 Hz) neurons displayed single exponential decays in firing rate while at the slower PRs they showed little adaptation or an initial adaptation followed by recovery. For conditions where a single exponential fit was used, the 'A' tone time constant shows a general increase as the frequency difference is increased. The time constants also show a decrease as the PR is increased, as would be expected. For those conditions that show recovery for the 4 Hz PR, there is a general increase in both time constants as the frequency difference is increased. For the 2 Hz PR, where the data have been fitted with a double exponential, the

alpha time constant displays an increase as the frequency difference is increased.

Presentation rate (Hz)	Frequency difference (octaves)				
	0	**0.25**	**0.5**	**0.75**	**1**
2	0.65	1.00	N / A	1.00	N / A
	5.39	1.04	N / A	1.65	N / A
4	0.76	0.69	0.64	0.93	1.04
	0.91	0.75	0.89	1.01	1.13
8	0.30	0.36	0.77	0.52	0.81
	N / A	N / A	0.93	7.34	0.89
12	0.18	0.28	0.27	0.34	0.44
	N / A	N / A	N / A	N / A	N / A
16	0.15	0.24	0.24	0.35	0.32
	N / A	N / A	N / A	N / A	N / A

Table 5-2. The time constants of adaptation in seconds for the 'A' tone calculated from the mean population spike counts. Where two values are given (in conditions in white cells), the upper and lower values represent the alpha and beta time constants of the double exponential fit respectively (from equation 5-2).

Presentation rate (Hz)	Frequency difference (octaves) 0	0.25	0.5	0.75	1
2	0.65	N / A	N / A	1.00	N / A
	5.39	N / A	N / A	1.70	N / A
4	0.76	0.29	0.35	0.36	0.48
	0.91	N / A	N / A	N / A	N / A
8	0.30	0.36	0.41	0.36	0.29
	N / A	N / A	N / A	N / A	N / A
12	0.18	0.37	0.28	0.25	0.25
	N / A	N / A	N / A	N / A	N / A
16	0.15	0.23	0.20	0.20	0.20
	N / A	N / A	N / A	N / A	N / A

Table 5-3. The time constants of adaptation in seconds for the 'B' tone calculated from the mean population spike counts. Same conventions as table 5-2.

The majority of the 'B' adaptation data can be sufficiently fitted with a single exponential; there is only 1 condition (other than the 0 octave A-A conditions) where recovery may be occurring. Across frequency difference, time constants don't seem to vary in a systematic way, increasing with

frequency difference for the 4 Hz condition and decreasing with increases in FD for the 8 Hz, 12 Hz and 16 Hz conditions. As with the 'A' tone adaptation, there is a decrease in the time constant as the PR is increased.

The *A* component of the exponential fits is displayed in figure 5-15. This component signifies the spike count at which the adaptation levels out. For the 'A' tones (shown in figure 5-15a) the *A* component tended to increase as the frequency difference increased, irrespective of whether the data were fit with a single or double exponential. This component also decreased as the PR was increased. These trends are similar to those described for the response to the 'A' tone in the period histograms. The *A* component of the 'B' tone adaptation generally decreased as the FD and PR were increased, as shown in figure 5-15b.

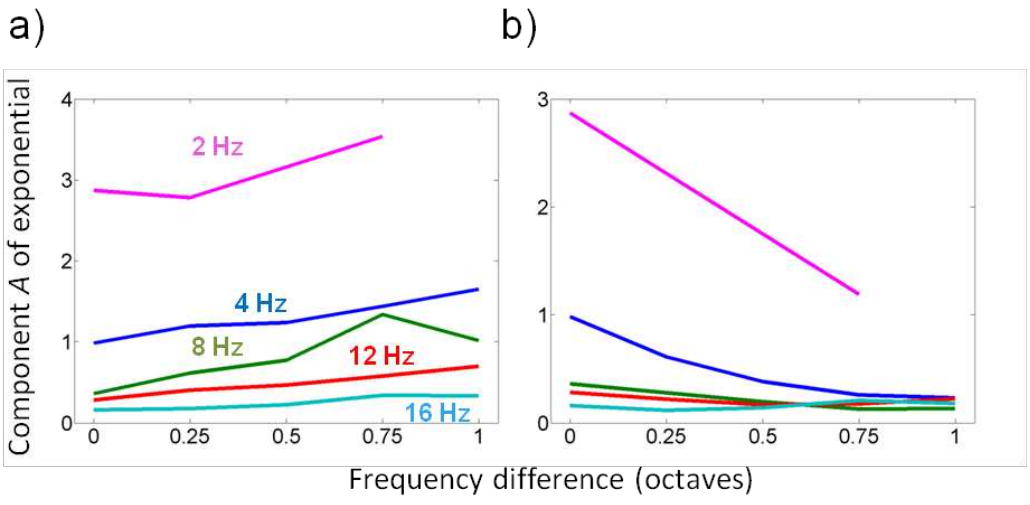


Figure 5-15. Component *A* of the exponential fits to the spike count data for a) the 'A' tones and b) the 'B' tones. Component *A* signifies the spike count at which the adaptation reached its steady-state level.

In summary, adaptation was generally over by 2 seconds and at slower rates there was evidence of a recovery, especially of the 'A' tone. Adaptation occurred faster and became more complete as the PR was increased for both the 'A' and 'B' tones. As observed in the period histograms, the steady-state 'A' tone response increased as the FD increased. Further, adaptation of the 'A' tones generally became slower as the frequency difference was increased. One question would concern whether the increase in the *A* component was due to the recovery of the 'A' tone response in certain conditions. However, while recovery may have augmented the increase at lower PRs it was not necessary for the increase to be observed, for example, in the 12 and 16 Hz conditions.

The differences in adaptation displayed by the 'A' and 'B' tones in each condition could be correlated with the perceptual build-up of streaming. For example, the rate of the build-up of streaming will depend on the relative rates of adaptation between the 'A' and 'B' tones. The mean spike counts described above are difficult to relate directly back to perceptual streaming, however this issue will be returned to in section 5.3.3.2.

5.3.3. Signal detection model

In order to compare the neural representation with the perception of streaming I now utilise a model based on signal detection theory (Micheyl et al., 2005). The aim is to consider whether spike count distributions from units in the anaesthetised guinea pig can be used to account for both the perceptual build-up and the perceptual boundaries of streaming. I begin by

describing the use of the model to account for steady-state responses, because it is simpler to describe how the model works in this case. I then describe how the model can be used to see if the data can account for the build-up of streaming.

5.3.3.1. The steady state response

The signal detection model was applied to the steady state responses by creating spike count distributions from the responses between 2 and 7 seconds of recording and then fitting these with normal distributions. Figure 5-16 shows the 'B' tone population spike count distributions at each PR for the steady-state responses. Only data with a significant vector strength of 0.4 or above has been included in the generation of population spike counts.

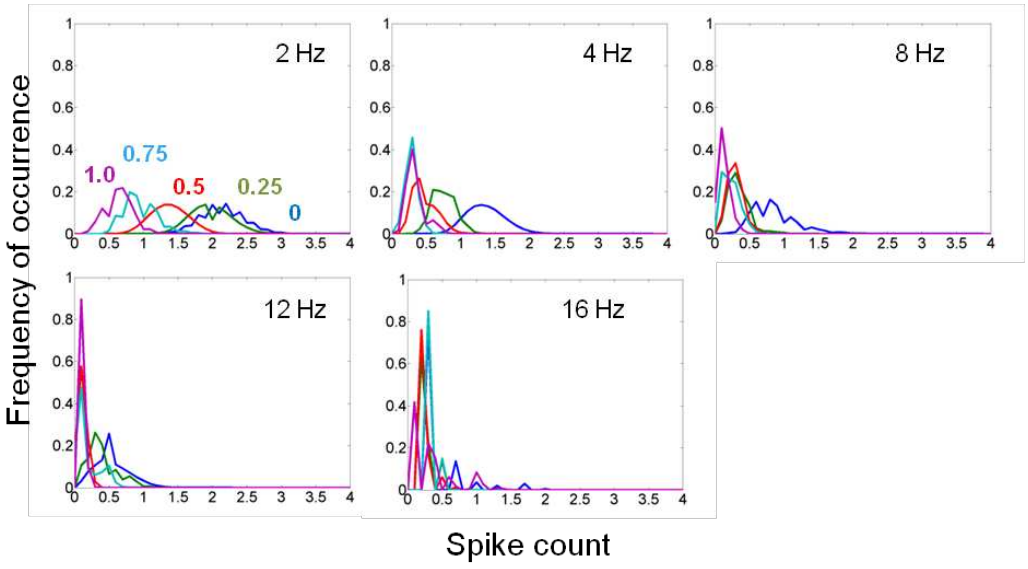


Figure 5-16. Population spike count distributions for the steady-state responses to the 'B' tones. Only data with a significant vector strength of 0.4 or above has been included in the creation of the spike count distributions. The colour code for frequency differences is indicated in the first panel.

Figure 5-16 helps to demonstrate how the signal detection model works. One can set a threshold along the spike count dimension that is fixed for each condition. The number of spikes that are to the right of the threshold (and will thus be detected) will differ for each condition. By dividing the number of spikes that are detected by the total number of spikes the probability of detection can be estimated. The probability of segregation is then the inverse of the probability of detection of the 'B' tones.

Figure 5-17 shows the predicted human psychophysical percept at each of the FDs and PRs used in this study. One approach to investigating how well the model could account for the data was to explore the range of thresholds that were plausible for each condition. For example, it can be seen from figure 5-16 that a single threshold across PRs would not allow the model to capture the psychophysical boundaries of streaming; any threshold above around 0.6 would result in segregation at all FDs for the 8, 12 and 16 Hz PRs. To investigate the range of thresholds that would capture the psychophysical boundaries of streaming it was necessary to constrain the model.

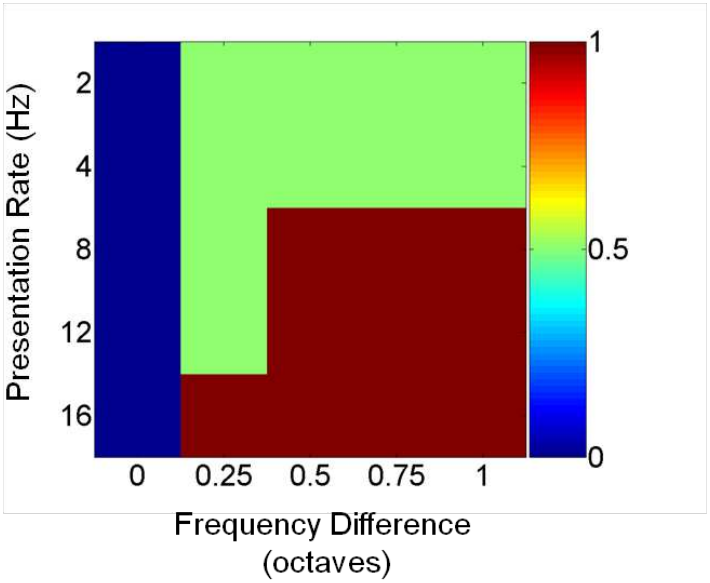


Figure 5-17. Human perceptual streaming regions for the PRs and FDs used in this study. Values of 0, 0.5 and 1 represent the coherent, ambiguous and segregated regions respectively. Adapted from McAdams and Bregman (1979).

To constrain the maximum threshold for each condition, I focussed on the area of figure 5-17 where the percept had to be coherent. The constraint dictated that for any PR the probability of segregation for the 0 octave FD condition must be less than 0.05. This would mean that on less than 1 in 20 presentations of the stimuli contained within this perceptual region a segregated response would be given. To find the maximum threshold, the threshold was ramped from 0 upwards, until the constraint was breeched. This gave a maximum threshold at each PR as shown in table 5-4.

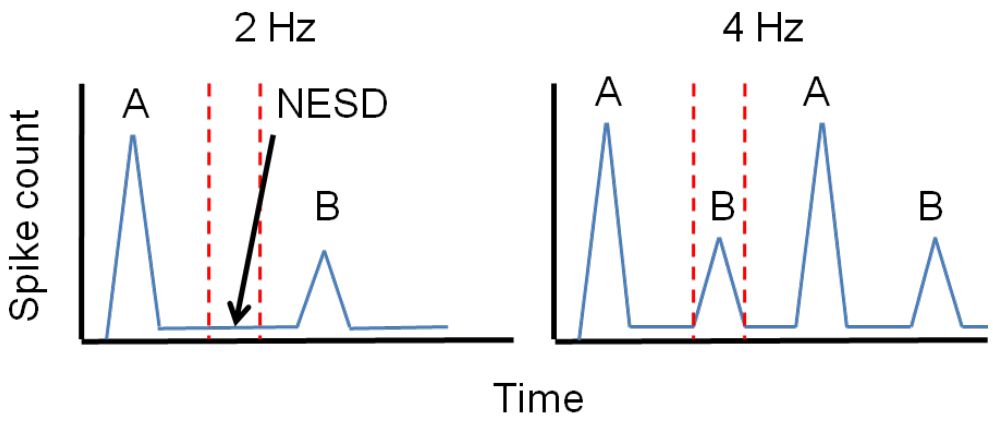


Figure 5-18. Schematic of the calculation of the no-tone equivalent spike count distribution (NESD). The NESD was calculated from the period of silence that followed the 'A' tone in the condition with the same FD but at half the PR. As there were no 1 or 6 Hz PR conditions, the 2 Hz NESD was used for the 2 Hz condition and the 4 Hz NESD was used for the 12 Hz condition.

I have shown that the maximum threshold can be calculated using the criterion that the zero FD condition must always be coherent. It is reasonable to suggest that the minimum threshold should be set such that it is above the activity that would occur in the same location in the absence of a tone, which will be termed the no-tone equivalent spike count distribution (NESD). For example, if the response distribution for a tone is similar to or below the NESD then it is unlikely that the tone will be represented. NESDs for a given condition were derived from the same FD condition at half the PR, as shown in figure 5-18. Thus, the period of the NESD should reflect that which would occur if the tone were not presented. The minimum thresholds for each condition were calculated as the lowest spike count which is greater than 95% of the NESD and are shown in table 5-4.

Presentation rate (Hz)	Frequency difference (octaves) 0	0.25	0.5	0.75	1
2	0.5 – 1.7	0.5 – 1.7	0.2 – 1.7	0.1 – 1.7	0.4 – 1.7
4	0.5 – 1.0	0.5 – 1.0	0.2 – 1.0	0.1 – 1.0	0.4 – 1.0
8	0.2 – 0.5	0.2 – 0.5	0.2 – 0.5	0.2 – 0.5	0.1 – 0.5
12	0.2 – 0.3	0.2 -0.3	0.2 – 0.3	0.2 – 0.3	0.1 – 0.3
16	0.6 – 0.4	0.2 – 0.4	0.6 – 0.4	1.0 – 0.4	0.6 – 0.4

Table 5-4. Minimum and maximum thresholds the derived as explained in the text. Grey cells indicate conditions in which the minimum threshold is greater than the maximum.

Figure 5-19 shows the NESDs and the response spike count distributions for the 'A' and 'B' tones, with the minimum and maximum thresholds indicated by black and pink dashed lines respectively. As a 1 Hz condition was not tested, the NESD used here is that for the 2 Hz condition. In this case, the NESD will be overestimated. A 6 Hz PR was not tested so the NESD for the 4 Hz condition has been used for the 12 Hz response. In this case, the NESD will be slightly underestimated. It can be seen that in some conditions at the faster PRs the NESD and response spike distributions are nearly identical such that it would be difficult to disambiguate a spike response from the NESD. Due to this, at the 16 Hz PR some conditions have

minimum thresholds that are higher than the maximum thresholds (indicated as grey cells in table 5-4).

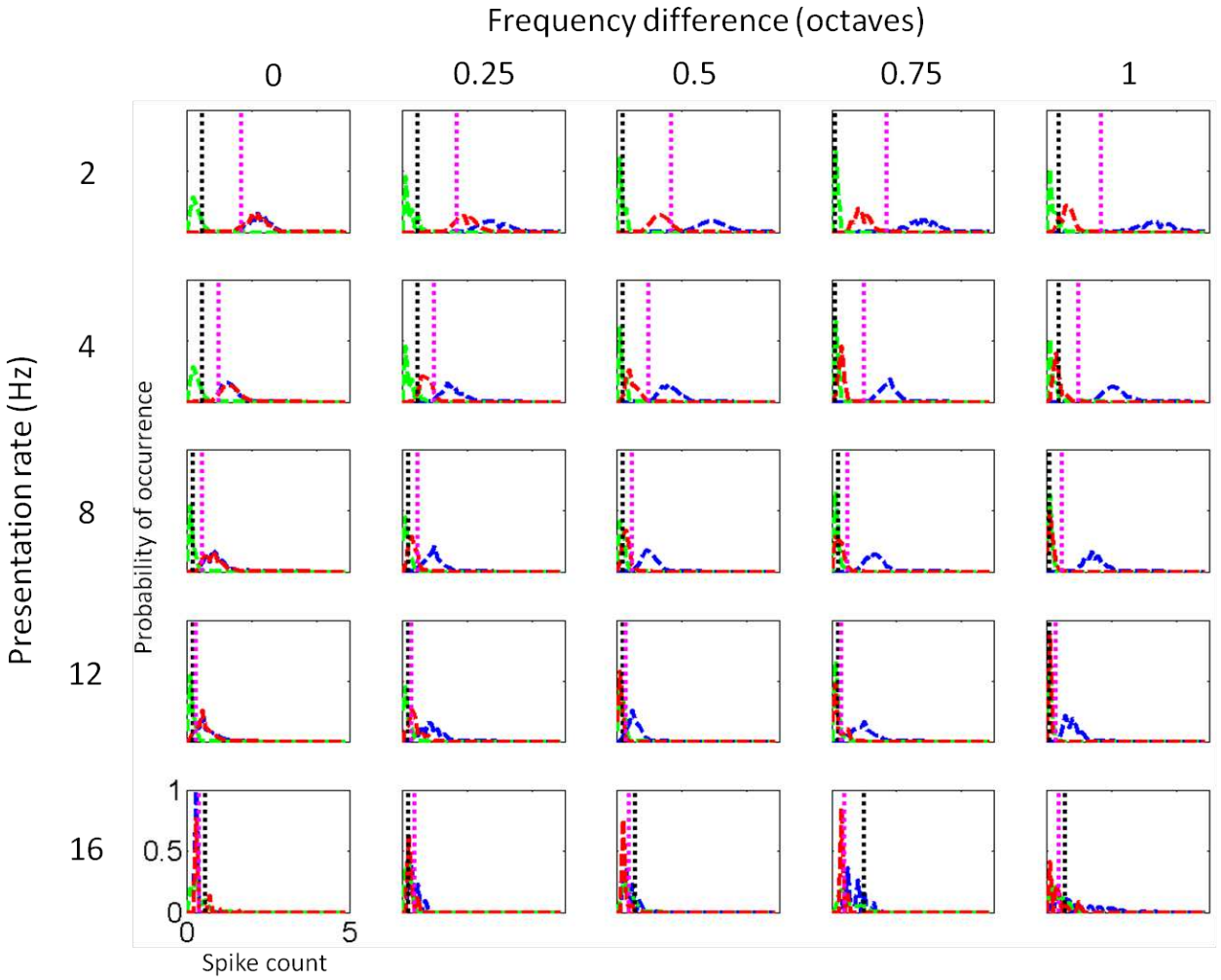


Figure 5-19. Mean population NESDs (green dashed line) and spike count distributions for the response to the 'A' tone (blue dashed line) and the response to the 'B' tone (red dashed line) for each FD and PR. The minimum and maximum thresholds, derived as described in the text are shown as black and pink dotted lines respectively.

In order to compare the physiological data to the psychophysical data, each condition was defined as either coherent, segregated or ambiguous using the following criteria. If the probability of the response to a ‘B’ tone was greater than 0.95 when the threshold was set to both the minimum and maximum value then the condition was classed as coherent. If the probability of the response to a ‘B’ tone was less than 0.95 when the threshold was set to both the minimum and maximum value then the condition was classed as segregated. Conditions that were neither classed as segregated or coherent were classed as ambiguous.

Figure 5-20b shows the result of this classification along with the perceptual boundaries for comparison (in figure 5-20a). In general the signal detection model captures the perceptual boundaries well, though there are conditions where it fails. As would be expected given the criterion by which the minimum threshold was derived, the coherent region is well characterised, apart from in the 16 Hz condition (indicated by a blue star in figure 5-20b) where the percept is ambiguous. This is not surprising given that the response spike count distributions overlap with the NESDs to such an extent at this PR. In some conditions the neural data indicate a segregated percept when the psychophysical data indicate an ambiguous percept (indicated by red stars). This again is likely due to low spike counts in these conditions that are similar to the NESDs. Thus the threshold can not be set low enough for the ‘B’ tone to be represented, given the criterion imposed. In some conditions the neural data indicate an ambiguous percept when the

psychophysical data indicate a segregated percept (indicated by black stars). In these cases it seems that the NESDs are low enough for the 'B' tone to be represented, though the spike counts in response to the 'B' tone are still low.

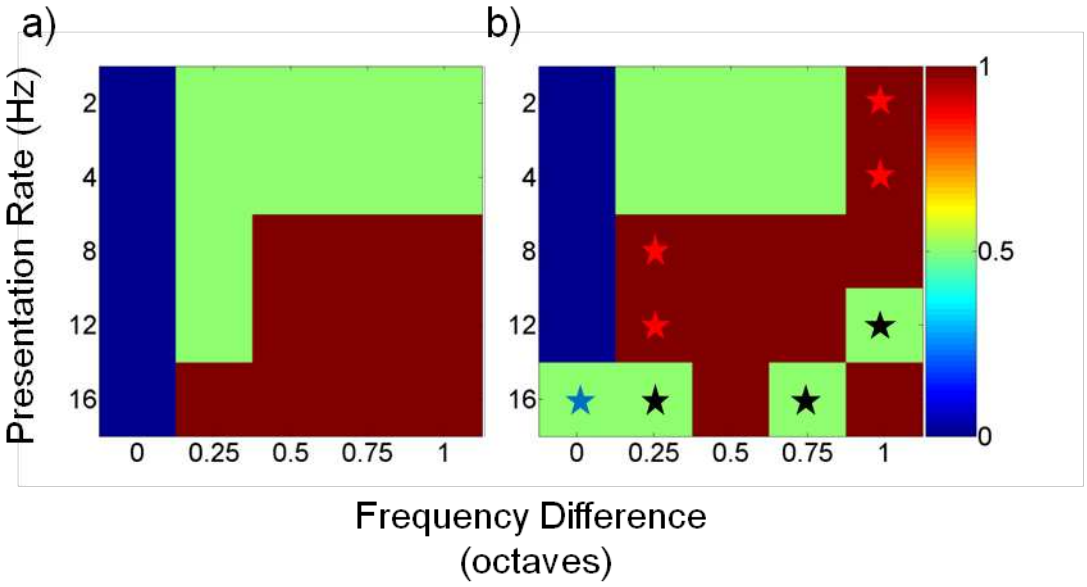


Figure 5-20. The perceptual boundaries derived a) psychophysically and b) using the signal detection model on the steady-state data. Stars indicate where there is a discrepancy between the psychophysical and neural data. Red stars indicate where the neural data indicate a segregated percept and the psychophysical data indicate an ambiguous percept. Black stars indicate where the neural data indicate an ambiguous percept and the psychophysical data indicate a segregated percept. Blue stars indicate where the neural data indicate an ambiguous percept and the psychophysical data indicate a coherent percept. The psychophysical data in a) is adapted from McAdams and Bregman (1979).

5.3.3.2. The build-up of streaming

I now move on to discuss the use of the signal detection model to investigate the build-up of streaming. In this case, the spike count distributions were created for each tone in the sequence for each condition.

The model was initially run using the raw spike count distributions, the results of which can be seen compared to human psychometric curves in figure 5-21. The stimulus used for the psychometric curves was different to the stimulus used here. Firstly, it was in the form ABA_ABA_ instead of ABABABAB. Secondly, the tones were 125 ms in duration with silent gaps of 50 ms between each tone and a silent gap of 175 ms between each triplet (Micheyl et al., 2005). This makes direct comparison between the datasets difficult, but, for interest, the neurometric functions for the 8 Hz condition are displayed in figure 5-21.

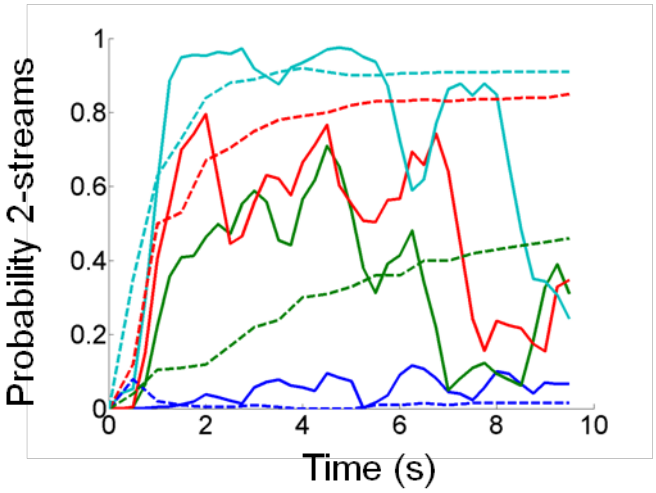


Figure 5-21. Comparison between human psychometric functions and neurometric functions from the population data. Dashed lines show the psychometric curves and solid lines show the neurometric curves for the 8 Hz PR with a threshold of 0.3. Frequency differences of 1, 3, 6 and 9 semitones are shown by dark blue, green, red and light blue lines respectively. Neurometric curves have been smoothed with a moving average linear filter (3 term, each term was 1/3). Psychometric data estimated for the slower PR data from Micheyl et al. (2005).

There are signs of build-up at the 8 Hz PR and there is a separation between the neurometric curves at different FDs that can also be seen in the psychometric curves. However, even though the neurometric curves have been smoothed they are still jagged compared to the psychometric curves.

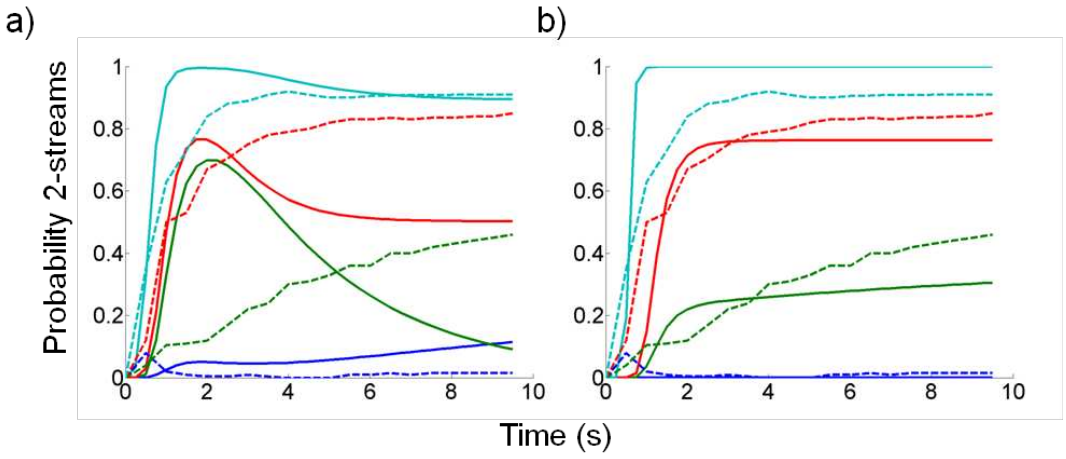


Figure 5-22. Predictions of neurometric functions for the 8 Hz PR data derived by fitting normal spike count distribution parameters (mean and standard deviation) as a function of time with a) the best-fitting functions (mainly double exponential) and b) the best fitting functions after the double exponential. The standard deviation parameter was divided by 4 to simulate the combination of spike count information in a population of 4 statistically independent neural sites. Colour code as in figure 5-21.

Also, the neurometric curves tend to decrease after around 6 seconds, a feature not observed in the psychometric curves. In order to smooth the recovery functions both the mean and standard deviation from the normal distributions were fitted as a function of time for each condition as described in section 5.2.3. Figure 5-22a shows how the neurometric fits compare to the psychophysical data. In general, double exponential functions gave the least sum of squared errors and thus the best fit to the data. However, the

recovery of responses which can be seen as a decrease in the neurometric functions in figure 5-22a is incompatible with the psychometric functions, which only show an increase with time.

If the function with the next best fit to the data is used (after the double exponential) then the neurometric curves start to resemble the psychometric curves to a greater extent, at least qualitatively, as shown in figure 5-22b. The neurometric curves have a sigmoidal shape due to the fast time constants of adaptation associated with the functions: there is an abrupt shift in the probability of the 'B' tone being represented as the threshold is crossed by the function. Also, the neurometric curve for the 9 semitone condition shows a build-up that is too fast and too complete to account for the psychophysical data.

To summarise, there is some evidence of the build-up of streaming in the data from anaesthetised guinea pigs, however, the neurometric curves are jagged compared to the psychometric curves. Fitting functions to the curves helped to smooth the data but only certain components of the perceptual build-up could be captured. Double exponential functions provided the best fit for the data but the recovery was incompatible with the psychometric functions. Fitting with the next best function (and thus ignoring the recovery of responses) showed that the psychometric data could be accounted for with the neurometric curves with some caveats. Firstly, the neurometric curves were sigmoidal due to the fast time constants of adaptation causing an abrupt change in the probability of the 'B' tone being

represented. Secondly, the build-up of streaming was too quick and complete for the 9 semitone condition as compared to the psychophysical data.

5.4. Discussion

5.4.1. Main findings

- The steady-state responses to ABAB tone sequences in cortical single units of the anaesthetised guinea pig were similar to neural ensemble responses in awake animals (Fishman et al., 2001; Fishman et al., 2004)

- As the frequency difference between the 'A' and 'B' tones was increased the response to the 'B' tone decreased and the response to the 'A' tone increased

- As the presentation rate of the sequence was increased the response to both the 'A' and 'B' tones decreased, but the response to the 'B' tone decreased to a greater extent

- Similar trends were observed in the mean population responses, however, at faster presentation rates the number of neurons with responses that were locked to the onset of the tones in the sequence decreased: clear trends were only obvious after removal of units that did not phase-lock

- Spike count responses adapted as the sequences unfolded and was generally over by 2 seconds

- At slower presentation rates there was evidence of a recovery of response after the initial adaptation, especially for the ‘A’ tone responses

- Adaptation occurred faster and was more complete as the presentation rate was increased

- Adaptation of the ‘A’ tone responses became generally slower as the frequency difference was increased

- A signal detection model applied to the spike count distributions for the steady-state responses captured some of the features of the perceptual responses to streaming stimuli

- The build-up of streaming could be accounted for to some extent using a signal detection model applied to the neural data over time

- However, the recovery from adaptation mentioned above had to be ignored for the neural data to fit the psychophysical build-up data

5.4.2. Comparison to previous studies

The following is a comparison of the results from this chapter with streaming responses that have been reported previously, mainly in awake animals. Although the results presented in this chapter were generally consistent with those from awake animals, there were some discrepancies which I will conclude are likely due to a faster and more complete adaptation of spike count responses in the anaesthetised preparation.

5.4.2.1. The steady-state responses

5.4.2.1.1. The spike count responses

The steady-state results presented in this chapter display similarities to those from awake preparations (Fishman et al., 2001; Bee and Klump, 2004; Fishman et al., 2004; Bee and Klump, 2005). The steady-state responses can be explained somewhat as an extension of the effects of forward suppression (Fishman et al., 2001). The increase in the 'A' tone response with increasing FD could be due to a release of suppression by the 'B' tones. This follows from evidence that the suppression of BF responses decreases as the frequency of the conditioner tone moves away from BF (e.g. Brosch and Schreiner, 1997). Similarly, the differential suppression of the 'B' tone as the PR is increased could relate to the fact that forward suppression is stronger when the spectral and temporal gap between the conditioner and probe tones is smaller (Brosch and Schreiner, 1997). This notion is supported by a study that shows that the absolute response decrement of the 'B' tone is larger when the FD is smaller and increases with PR (Bee and Klump, 2004).

At the fastest PR tested (16 Hz), the number of neurons that showed statistically significant phase-locking to the tone sequences was small. Even after removal of units that did not phase-lock sufficiently, the response ratio at the 16 Hz PR does not show a systematic decrease as the FD is increased (as it does at all other PRs) and is larger than would be expected at most FDs. This probably occurs because the responses at 16 Hz have adapted to near-zero, especially at the smaller FDs (consider the 0, 0.25 and 0.5 FDs for the

red period histograms in figure 5-10). The response to the 'A' and 'B' tones are both affected such that the ratio between the two becomes noisy. It is perhaps not surprising that only a small number of units locked to the tone sequence as the PR was increased, as a previous study has shown that synchronous firing in cortical neurons ceases at around 20 Hz for amplitude modulated tones (Creutzfeldt et al., 1980). It is likely that synchronous responses were recorded at up to 40 Hz in awake preparations for two reasons (Fishman et al., 2004). Firstly, the ability of A1 neurons to respond synchronously to repetitive stimuli is susceptible to anaesthesia (a direct comparison of awake vs. ketamine-xylazine anaesthesia can be found in Rennaker et al., 2007). Secondly, the neural ensemble activity reported by Fishman et al. (2001, 2004), represents the responses of a large number of units and would thus be less sensitive to unit-by-unit variations in synchronisation ability and less vulnerable to spike count variability.

5.4.2.1.2. The signal detection model

Generally, steady-state neural responses could be used within the framework of signal detection theory to account for the streaming percepts from psychophysical experiments. In half of the conditions where there was a discrepancy between the neural and psychophysical responses, the neural percept was segregated when the psychophysical percept was ambiguous. These discrepancies are likely due to low response spike count distributions that are similar to those for the spontaneous activity. It is possible that the larger spike counts encountered in an awake preparation (Rennaker et al.,

2007; Micheyl et al., 2008) would allow for a better representation of the perceptual results.

5.4.2.2. Adaptation and the build-up of streaming

5.4.2.2.1. Spike count responses

For those conditions where the spike counts were best fit with exponential functions (the shaded cells in table 5-2 and table 5-3), the trends in the time constants of adaptation can be explained using a framework of forward suppression. If one considers that decreases in PR result in an increase in stimulus onset asynchrony, then the increase in spike counts as PR decreased can be considered as a release from forward suppression. Similarly, as the PR was decreased, the time constants for both the 'A' and 'B' tones increased (adaptation decreased). This suggests that the adaptation and forward suppression are related and may have a common underlying mechanism, for example, synaptic depression (see Chapter 4). Further evidence for this hypothesis can be seen in the increase in time constants for the 'A' tone as the FD is increased. Again this can be related to the release from forward suppression as the FD is increased. Time constants for the 'B' tones generally increase from the 0 octave FD but there is no systematic increase in the time constants with FD.

A feature of the adaptation to tone sequences in this study that has not been reported previously in awake studies is the recovery of responses in some of the conditions. Generally recovery of responses occurred at the

lower PR conditions for the ‘A’ tones and, over the 6 seconds of analysis, was well fit by a double exponential function. It is possible that the ‘recovery’ is a manifestation of an oscillatory response. Indeed, it is possible to see subtle oscillations in some of the spike counts, for example, in the 4 Hz PR / 1 octave FD condition of figure 5-14, notice how the response initially decreases, then recovers and then decreases again at the end. Oscillatory responses have been observed previously in the auditory cortex, in the response to both tones (Sally and Kelly, 1988) and click trains (Eggermont, 1992; Rennaker et al., 2007). However, oscillations are thought to be a consequence of anaesthesia as they have not been observed in awake preparations (Rennaker et al., 2007). Recovery of responses have not been observed in tone sequences in awake animals (Micheyl et al., 2005; Micheyl et al., 2008), however the differences between this study and those of Micheyl et al. (2005, 2008) extend beyond the state of the animal, for example the stimulus used was different. Thus, the recovery of responses may be an interesting topic for further study, preferably with longer tone sequences to disambiguate between pure recovery and oscillatory behaviour. Speculatively, an oscillatory change in the response to the ‘A’ tone over the course of a tone sequence could suggest a mechanism by which bistability occurs. For example, if the ‘A’ tone response were to move closer to the ‘B’ tone response there would be a higher likelihood that the responses would be integrated into a single stream percept and *vice versa*. Interestingly, recovery was only observed in those conditions under which the percept is ambiguous and thus given to bistability.

5.4.2.3. Adaptation and the build-up of streaming

The recovery discussed above presents a problem when attempting to use the data to account for the build-up of streaming. The psychometric functions show a systematic increase with time whereas the neurometric functions created by fitting a double exponential function to the responses begin to decrease after around 1-2 seconds. It seems that this trend is reflecting the underlying neural responses (consider the similarity between the raw data in figure 5-21 and the double exponential functions in figure 5-22) and so the conclusion is that the neural responses cannot be used to account for the build-up of streaming.

For interest, and with the assumption that the recovery of responses may be related to the anaesthetic, the neural data was fitted with either single exponential or power law functions (whichever provided the better fit to the data). The fit between these neurometric functions and the psychometric functions was considerably better indicating that the time constants of adaptation might be similar enough to those in awake animals for the build-up to be accounted for. However, there are still discrepancies between the neural and psychophysical data which mainly relate to the adaptation of neural responses occurring too quickly. For example, the neurometric curve for the 9 semitone condition in figure 5-22b is steeper than the psychometric curve, indicating that adaptation was too fast in this case. Also, nearly all of the neurometric curves are sigmoidal probably because there is an abrupt shift in the probability of the 'B' tone being represented as

the threshold is crossed by the function. With a slower adaptation this would be less likely to be a problem.

The differences in the rate of adaptation between the anaesthetised and awake animals may be due to stimulus differences. As mentioned in section 5.3.3.2, the stimulus employed in this study is different to that which was used in both the psychophysical and the CN experiments. Alternatively the differences could simply be due to anaesthetic as it has been reported that adaptation occurs at a faster rate in ketamine-anaesthetised preparations than in awake preparations (Rennaker et al., 2007). However, the same anaesthetic and animal were used in the study by Pressnitzer et al. (2008) so the differences could be due to a combination of both anaesthetic and stimulus differences.

6. The effect of multiple conditioner tones on suppression

6.1. Introduction

I have shown in chapter 4 that forward suppression is specific to the frequency of the probe tone. I have also shown that the anaesthetised guinea pig can be considered a model for the study of auditory streaming, with the caveat that the build-up of streaming could not be accounted for using the signal detection model. One question that is yet to be addressed is how the frequency specific effects observed in forward suppression develop into the responses that are observed in streaming. The studies presented in this chapter are split into two parts, with both aiming to address the question of how forward suppression builds up in a sequence of tones.

The first study is concerned with how suppression builds up in the frequency dimension. The effect of varying the number of isofrequency conditioner tones on the frequency response (FR) of a neuron is investigated. Due to differences in the FRs recorded from neurons at different electrode positions, the conditioner tone would not always be at the best frequency of a neuron. Thus, it is possible to investigate the effects of adaptation of both on-BF and off-BF conditioner tones on the FR of a neuron. I have shown that in an ABAB tone sequence the 'A' tone adapts differently to the 'B' tone and previous studies in awake animals suggest that this differential adaptation may be a substrate of the build-up of streaming (Micheyl et al., 2005; Micheyl

et al., 2008). However, the ABAB stimulus allows only a narrow view into what is actually happening within single neurons. Essentially, physiological studies have sought to find neural correlates of the perceptual features of streaming, thus they have not gone further than correlating the neural responses with the perceptual responses. It is not known, for example, how adaptation builds up within and between the 'A' and 'B' tones.

Forward suppression has been posited as underlying the neurophysiological responses to ABAB sequences (Fishman et al., 2004) and I have shown in chapter 4 that tones of different frequencies, for example on-BF and off-BF tones, are suppressed by different ranges of tones. However, there is a somewhat theoretical contradiction between the effects of forward suppression and the effects observed with ABAB sequences. Forward suppression is strongest when the conditioner and probe tones have similar frequencies and gradually gets weaker as the frequency difference increases. Conversely, the response to the 'B' tone in an ABAB sequence decreases more as the frequency difference increases. As mentioned in chapter 5, this is due to the 'B' tone effectively dropping out of the receptive field, but it still doesn't belie the fact that more forward suppression will be occurring when the frequency difference is 0 octaves than when the FD is 1 octave. So, forward suppression alone is insufficient to explain the effects observed with the ABAB sequences. It must develop in a way that results in an overall preferential response to the 'A' tone, even though most forward suppression will be directed towards the 'A' tone. By investigating the effects of the build-

up of adaptation for both on-BF and off-BF tones on the FR I hoped to gain an insight into the changes that occur.

One aspect of the manner in which forward suppression develops over the course of several tones is the extent to which frequency specificity either increases or decreases. In a similar paradigm to the one used in this study, Condon and Weinberger (1991) investigated the effect of repetitive stimulation with sequential isofrequency tones on the RF of guinea pig cortical neurons. In that study, 300 ms tones were presented at a slower rate (1.25 Hz) and for a longer time (around 5 – 7 minutes). Following repetitive stimulation, multi-unit cortical RFs exhibited highly frequency-specific suppression at the conditioner tone frequency with a bandwidth of less than 1/8 octave. Over a similar time course Ulanovsky et al. (2003) reported that stimulus-specific adaptation in single cortical units displayed a hyperacuity, occurring when two tones had a frequency difference of only 0.06 octaves. One question that I addressed in this chapter is whether such narrowly-tuned frequency-specificity is observed after the presentation of isofrequency tones at a much shorter timescale.

A further aim of this study was to investigate the extent to which the models outlined previously (Grill-Spector et al., 2006) apply to adaptation at the level of a single or small group of neurons. In addition, the nature of adaptation, as investigated in this way, would provide further information on the manner in which neural FRs are created by the inputs to a neuron.

The second study in this chapter investigates how units recover from adaptation and how this recovery is affected by increasing the number of conditioner tones that precede a probe tone. An aim of the second part of this study was to investigate how the recovery from adaptation builds up over the course of a series of tones. For example, is the effect of an additional conditioner tone to simply shift the recovery function along in time while it maintains its overall shape? Or do we see an overlap in recovery functions that causes changes in the shape? This will help to clarify the way in which forward suppression develops to give rise to the physiological results on streaming.

6.2. Methods

6.2.1. Stimuli

6.2.1.1. Experiment 1: The frequency response

Figure 6-1 shows a schematic of the stimulus used to measure the effect of multiple conditioners on the frequency response (FR). A level at which the neuron responded well (at around 30-40 dB above threshold) was chosen and a range of frequencies at this level were used as probe stimuli. The conditioner tone frequency was selected from the probe frequencies. Conditions were presented randomly and repeated between 5 and 20 times. Tones were 50 ms (4.5 ms linear gated) in duration, conditioner tones were presented at 4 Hz and the ISI between the last conditioner tone and the probe tone was 100 ms.

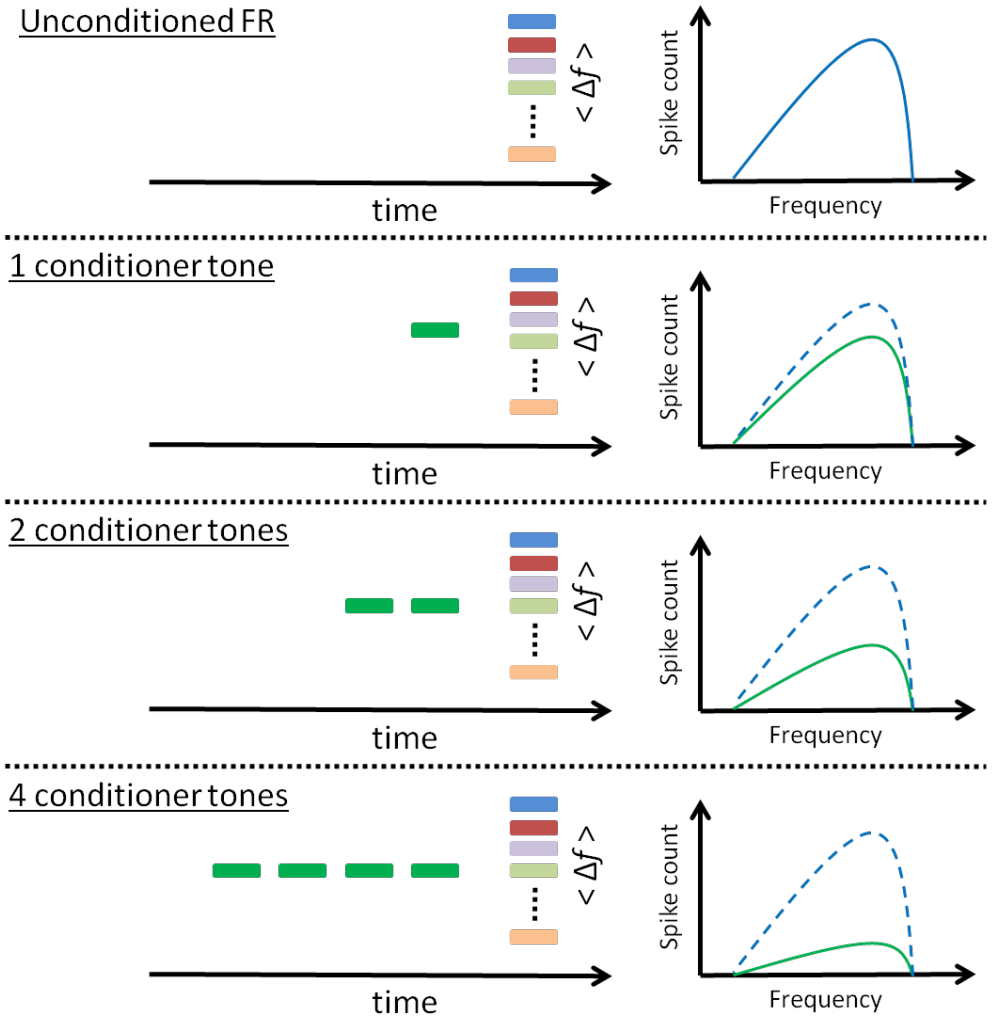


Figure 6-1. Schematic of the stimulus used to measure the effect of multiple conditioners on the FR. An unconditioned FR was constructed by presenting tones at various frequencies as shown in the upper panel. Conditioner effects were measured by presenting 1, 2 or 4 conditioner tones before the tones used to derive the FR, as shown in the lower panel.

6.2.1.2. Experiment 2: The recovery from adaptation

Figure 6-2 shows a schematic of the stimuli used to measure the effect of multiple conditioners on the recovery from adaptation. The conditioner and probe tones were at the same frequency, selected from the RF so as to elicit a reliable response. Conditions were presented randomly and repeated

between 4 and 20 times. Tones were 50 ms (4.5 ms linear gated) in duration and for a given unit the conditioner tones were presented at 4, 8 or 16 Hz. The ISI between the last conditioner tone and the probe tone was varied either on a linear or a log time scale. It was unclear how best to sample the recovery over time and so different ISIs were used for different units.

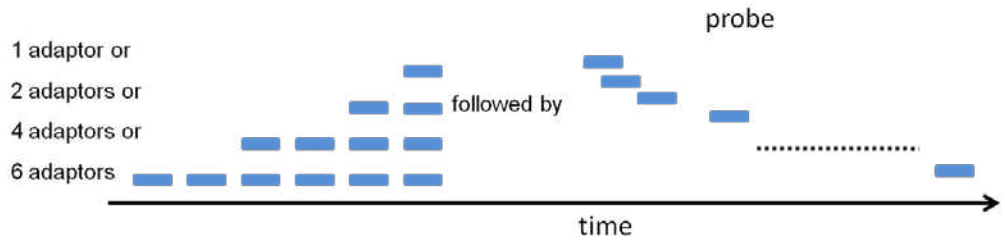


Figure 6-2. Schematic of the stimulus used to measure the effect of multiple conditioners on the recovery from adaptation. A condition consisted of 1, 2, 4 or 6 conditioner tones presented at 4, 8 or 16Hz followed after a gap of 1-640ms by a tone of the same frequency.

6.2.2. Analysis

6.2.2.1. The response time window

Response time windows were set as described in section 3.8. In some cases, the unit response had two components, for example at the onset and offset of the tone. For these units, two time windows were set, one for each component, and the components were analysed separately (see section 6.3.1.1.2 for a description of these units). These units were classified as onset/offset. Other units were classified as either sustained or onset, depending on a visual inspection of the temporal response.

For a given condition, spike count distributions were created by counting the number of spikes elicited within the time window for each presentation of the stimulus. A spike count distribution for the spontaneous activity was created by counting the number of spikes that occurred within silent periods before tone presentation. The time window over which the spontaneous activity was assessed was a scalar multiple of the response time window such that the distribution could be normalised by multiplying it by a scale factor. For experiment 1 the silent period preceded the FR tones in the no conditioner stimulus condition. For experiment 2 the silent period preceded the conditioner tone in the one conditioner stimulus condition.

6.2.2.2. The bootstrap method for comparing neural populations

A statistical bootstrap method was used to compare spike count distributions for different conditions in order to test whether they could arise from the same underlying distribution. Firstly, spike count distributions for both the unconditioned and adapted conditions were compared with the spontaneous spike count distribution. Secondly, unconditioned spike count distributions were compared to adapted spike count distributions to give a measure of when the adapted responses were not statistically different to the unconditioned responses. A statistical bootstrap method was used because the spike counts were not always normally distributed, thus statistical tests that assumed normality (for example a t-test) were unsuitable.

The bootstrap method worked as follows. The two distributions to be compared were concatenated to create a single list of spike counts. These spike counts were then randomly distributed into two lists of the same size as the input lists and the difference between the means of these new lists was calculated. This process was repeated 500 times. The percentile that the real mean difference was at was calculated by comparing it to the 500 randomly generated mean differences. A probability criterion of 0.05 was used in each distribution comparison.

6.2.2.3. Quantitative measures of the FR

For a frequency to be included in the unconditioned FR, the spike count distribution at that frequency had to be significantly different from the spontaneous spike count distribution. In the suppressed frequency response (SFR) only those frequencies that lay within the unconditioned FR were considered in any further analysis. The best suppressed frequency (BSF) was defined as the frequency at which the largest absolute decrease in spike count occurred. To define how specific the suppression was to the BSF a frequency specificity ratio (FSR) was calculated. The FSR was the absolute decrease in response at the BSF divided by the mean decrease in response across frequencies. Lower values represent a more uniform spread of suppression and higher values represent a situation in which suppression was *only* observed at the BSF.

6.2.2.4. Quantitative measures of the recovery from adaptation

The responses to the conditioner tones were well characterised because they were presented on every repeat of each condition. For the 4 Hz condition, the conditioner tone responses were fit with single exponential functions, which generally gave the lowest error as calculated using the sum of squared errors (compared to a power law and a linear fit).

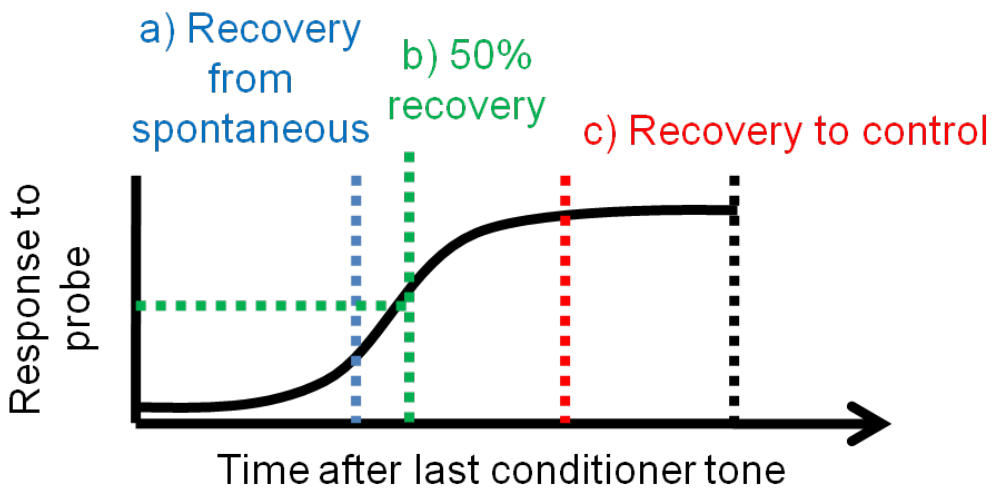


Figure 6-3. The recovery functions for each unit were characterised using three time points. a) The recovery from spontaneous (in blue) was the first time-point at which the response was significantly different to spontaneous activity, as assessed using the bootstrap method. b) The 50% recovery (in green) was the point at which the response had recovered to 50% of the difference between the mean unconditioned response and the mean spontaneous activity. If this value fell between two time-points it was calculated using linear interpolation. c) The recovery to unconditioned was the first time-point at which the response was significantly similar to the unconditioned response, as assessed using the bootstrap method.

The recovery patterns varied in shape, both across units and within units for different numbers of conditioner tones. Therefore, they were not fit with functions. Instead, the recovery from suppression was characterised using three time-points, as shown in figure 6-3. Using the bootstrap comparisons, two time-points were defined: the point at which the response distribution became significantly different from the spontaneous distribution and the point at which the adapted response distribution became significantly similar to the unconditioned response distribution. The unconditioned response distribution was taken from the first tone in the conditioner tone sequence. A 50% recovery time was also defined by finding the point at which recovery had reached 50% of the difference between the mean unconditioned response and the mean spontaneous activity. If this point fell between two time-points then linear interpolation was used to define it.

In some cases, the response time window was long enough for the response to the last conditioner tone to overlap with the response to the probe tones, at least for low ISIs. When this was possible, the analysis of the recovery from suppression began at the first time point when overlap was not likely to occur, calculated based on the response time window.

6.3. Results

6.3.1. The FR

The first question that will be addressed is the extent to which the suppression observed could be considered frequency specific. In other words,

does suppression occur mainly at the conditioner tone frequency or is it more general? A second question that will be addressed is whether, at the level of one or a small number of neurons, there is any evidence of a sharpening of tuning around the best frequency of a unit. These first two questions are both related to the form that suppression takes at the level of a small number of neurons and can thus be related back to the three models of adaptation discussed in the introduction (Grill-Spector et al., 2006). A third question is how the number of conditioners affects the suppression pattern that is observed. For example, does suppression become more or less frequency specific as the number of conditioners is increased?

6.3.1.1. Unit examples

6.3.1.1.1. Varying levels of frequency-specific adaptation

For each unit, some degree of suppression was always observed at the frequency of the conditioner, and this was especially evident for the 4 conditioner condition. Figure 6-4 shows the effect of multiple conditioner tones on the FR of two example units. In both cases, the response at the frequency of the conditioner tone (the dashed black line) decreases and this decrease is greater after 4 conditioner tones (solid light blue line) than after 1 conditioner tone (solid green line).

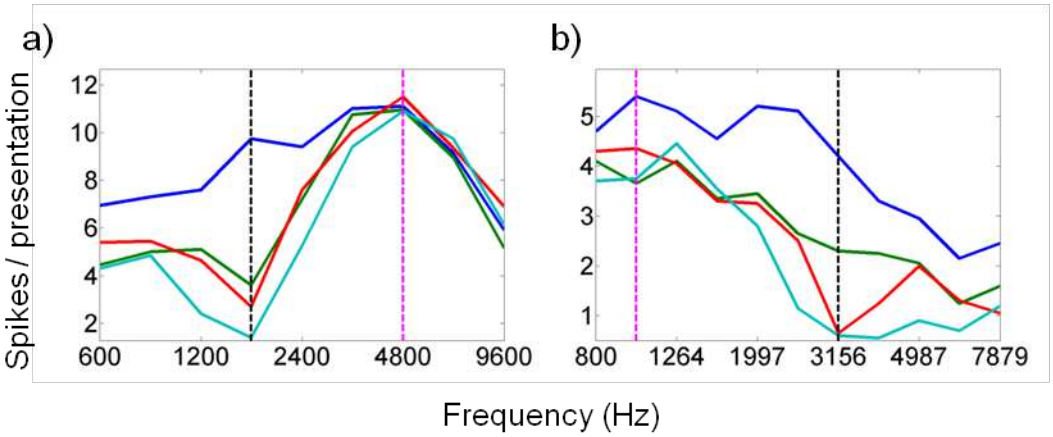


Figure 6-4. Two example units showing a decreased response at the conditioner frequency. The unconditioned FR is shown as a solid blue line. SFRs are shown after 1 (solid green), 2 (solid red) and 4 (solid light blue) conditioner tones. The control BF (dashed pink) and conditioner tone frequency (dashed black) are also shown. a) MU example. b) SU example.

The amount of suppression that occurred at frequencies other than the conditioner tone differed from unit to unit. In the example shown in figure 6-4a, suppression mainly occurred at and around the frequency of the conditioner tone, with higher frequencies around the BF (dashed magenta line) remaining relatively unaffected. A similar pattern can be observed in the example shown in figure 6-4b. The effect of one conditioner tone (solid green line) is to reduce the response at all frequencies, however further addition of conditioner tones mainly results in a further decrease in response at and around the conditioner frequency.

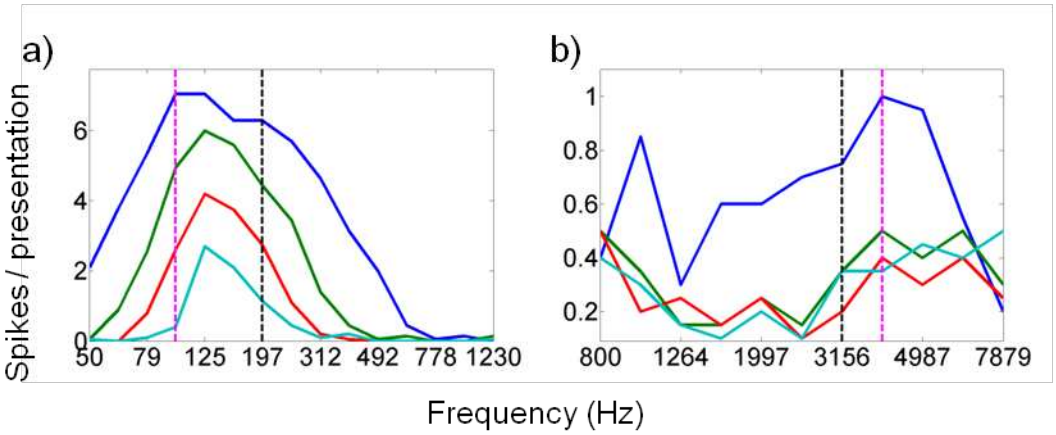


Figure 6-5. Two examples of overall suppression. Same conventions as figure 6-4. a) MU example. b) SU example.

In other units, suppression seemed to be more general, occurring at all frequencies in the FR. Two examples of this overall suppression are displayed in figure 6-5. In figure 6-5a the FR shows a general sharpening with each additional conditioner tone. As the response at BF decreases along with the rest of the FR, this would not be considered as an exemplar of the sharpening model. Instead the unit exhibits an overall suppression in which the BF response decreases but remains the dominant response. Figure 6-5b shows a unit in which the overall suppression has saturated after one conditioner tone, with further conditioner tones not leading to further decrements in response.

The examples in figure 6-4 and figure 6-5 are representative of the majority of units observed in this study, in that suppression was qualitatively either more specific or more general. These units will be assessed in more detail in the next section. Only 2 of the 54 units studied exhibited suppression that could be considered as a sharpening of the tuning and these units are

displayed in figure 6-6. In both examples the response around the BF remains relatively unaffected by either 1 or 2 conditioners, while the response at other frequencies decreases, causing a sharpening of the FR. After 4 conditioners the response at all frequencies is decreased relative to the unconditioned FR, similar to the overall suppression described above.

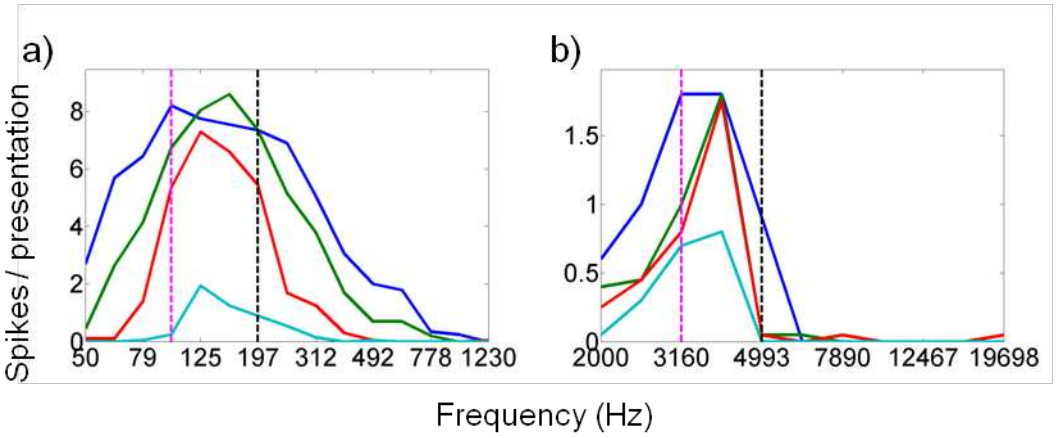


Figure 6-6. Two examples of sharpening. Same conventions as figure 6-4. a) and b) MU examples.

6.3.1.1.2. On-off units

Table 6-1 gives a summary of the number of units that displayed particular temporal response patterns. 7 of the 54 units displayed a phasic response that consisted of a distinct onset and offset component. The suppression was investigated separately for both components of the response by setting two time windows, one around each response component (as described in the methods).

	SU	MU	Total
Onset	4	19	23
Sustained	12	12	24
On-off	2	5	7
Total	18	36	54

Table 6-1. Number of units classified within each unit type.

For most of the units, the difference between the combined and onset response was subtle, however, in some units the offset response had a stronger influence on the combined response of the unit. Figure 6-7 shows three example units that displayed both onset and offset responses. In the MU example shown in figure 6-7a, the offset response changes the combined response of the unit slightly, with the BF shifting by one frequency step.

However, the suppression was specific to the frequency of the conditioner

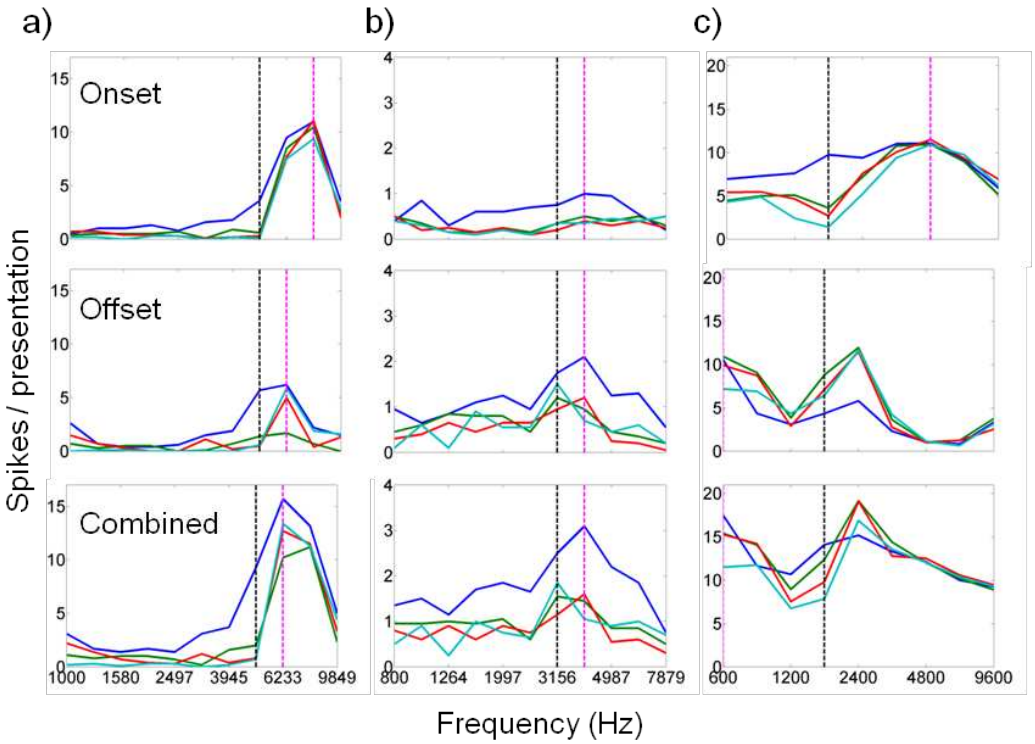


Figure 6-7. Three examples of on-off units showing FRs for the onset response, the offset response and the combined response. Same conventions as figure 6-4. a) MU example. b) MU example. c) SU example.

tone in both the onset and combined responses. The offset response of the SU example displayed in figure 6-7b is the dominant response, though overall suppression is observed in all aspects of the temporal response. Finally, a large change in the suppression pattern is observed in the MU example displayed in figure 6-7c. A combination of frequency-specific suppression for the onset and facilitation for the offset response lead to a complicated combined response pattern. Due to the small number of units with onset and offset responses, these units were not analysed further. The onset responses were used for the population analysis with the caveat that some aspects of

the more complex response patterns associated with the offset responses would be ignored.

6.3.1.2. Population analysis

6.3.1.2.1. The locus of suppression

In order to compare this data set with that of chapter 4, the frequency at which the suppression was greatest (greatest drop in absolute spike count) was defined as the best suppressed frequency (BSF). For different numbers of conditioner tones, figure 6-8 shows the distance of the BSF from the BF as a function of the distance of the conditioner frequency from the BF. The conditioner following index (CFI), which is analogous to the PFI in chapter 4, is also shown.

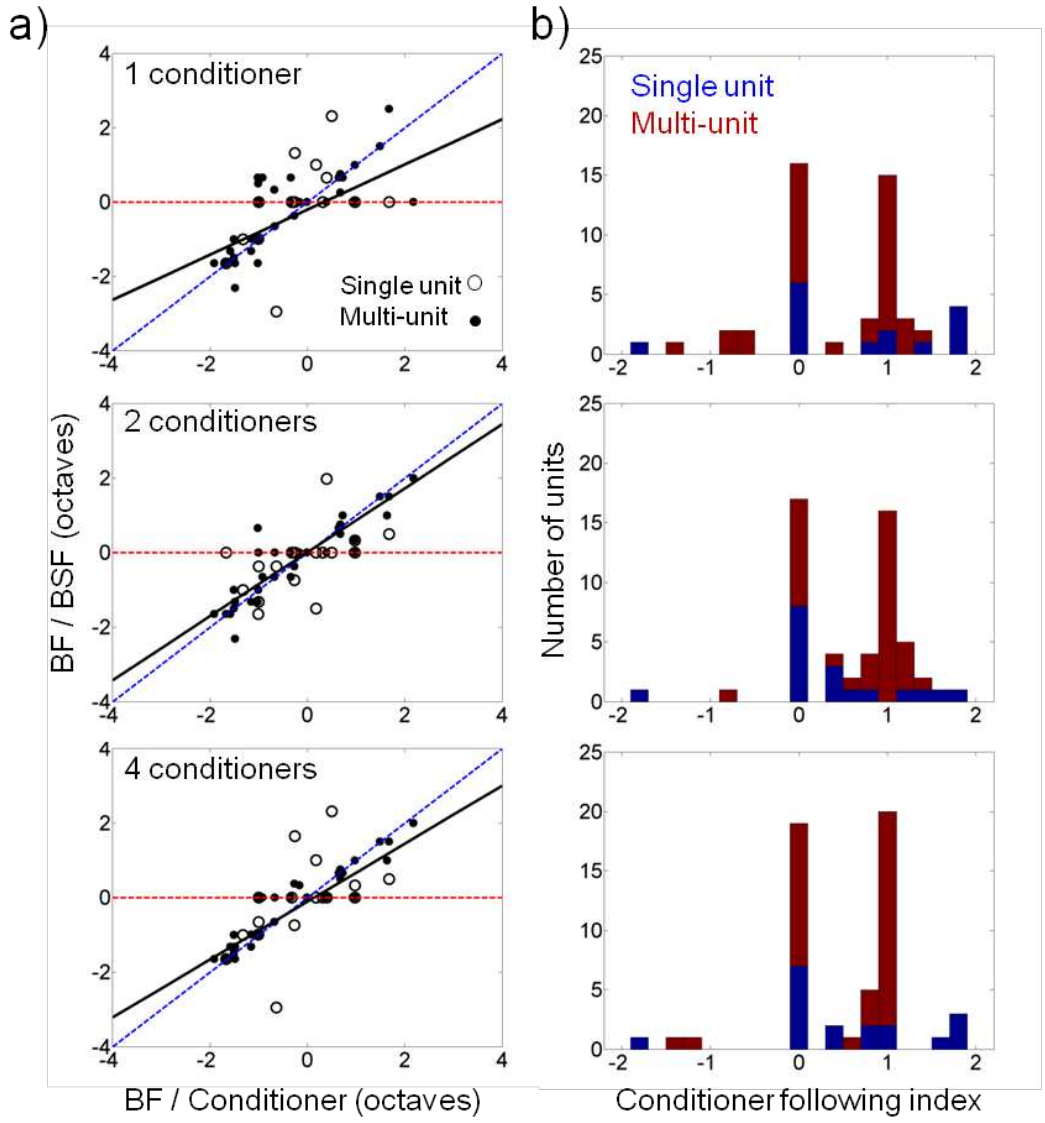


Figure 6-8. Relationship between the BSF and the conditioner tone frequency. The distance of the BSF from the BF as a function of the distance of the conditioner tone from the BF. Dashed lines indicate the trends that would be expected if the BSF was always at BF (red) or always at the conditioner tone frequency (blue). The solid black line shows a linear regression of all of the data. b) The CFI for each of the data points in a.

Essentially, a similar relationship is observed here between the BSF and the conditioner tone as was observed between the CSF/BSF and the probe tone in chapter 4. The maximum suppression is observed at the BF (CFI = 0) or at the conditioner frequency (CFI = 1) in most units. In fact, the

distribution of the CFI is more bimodal than the PFI, with less of a spread around the conditioner tone frequency. To investigate whether the peak at the BF was possibly an artefact of sampling, the simulation that was used in chapter 4 was run on the data presented in figure 6-8. Again, the BF-following peak disappeared if evenly distributed random numbers up to 1/8 octave were added to all of the BF and BSF values and the CFI was recalculated. Further the peak returned if these new values were transformed into more discrete numbers, indicating that the BF-following peak could again be an artefact of sampling. The peak at a CFI of 1 was present throughout. The gradients and R values for the regression fits for each number of conditioners is displayed in table 6-2. As the number of conditioners is increased both the gradient and the R value increase, indicating that the maximum suppression moves closer to the conditioner frequency with less variability, though, if one refers to the CFI histograms, this seems mainly due to the multi-units.

Number of conditioners	Gradient	R value
1	0.61	0.66
2	0.86	0.79
4	0.78	0.76

Table 6-2. Gradient and R values of linear regression fits between the conditioner tone frequency and the BSF, for different numbers of conditioner tones.

6.3.1.2.2. Bandwidth

Frequency specific adaptation can be defined by considering the frequency at which suppression occurs. It can also be defined by the bandwidth of the suppression: suppression can be considered more specific if the bandwidth is lower. In this section, I consider how the bandwidth of suppression changes with the number of conditioner tones.

Figure 6-9a shows that the mean suppression bandwidth (indexed by the number of significantly suppressed bands as described in the methods section) increases as the number of conditioner tones is increased. This would imply that suppression becomes less specific as the number of conditioners increases.

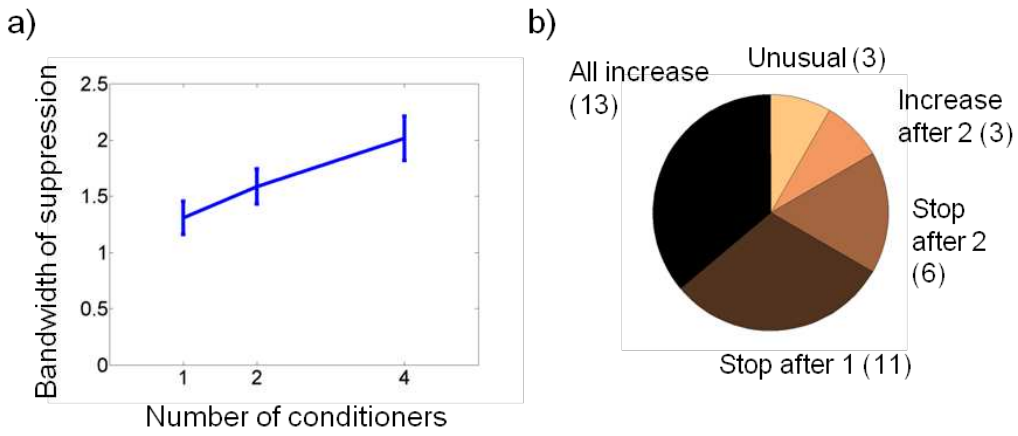


Figure 6-9. a) The bandwidth of suppression increases as the number of conditioners is increased. b) Units display varying patterns of suppression bandwidth increase (description in text).

While the mean suppression bandwidth exhibits an increase with each additional conditioner, this is not true of all units. Units were classified

depending on the number of conditioners at which bandwidth ceased to increase (see methods). As displayed in the pie chart in figure 6-9b, the majority of units increased with each additional conditioner (36%) or showed no further increase after the first conditioner (31%). Of the rest, the suppression bandwidth of 17% of units stopped increasing after 2 conditioners, in 8% of units it only increased after the second conditioner and 8% (3) of units did not fall into any of these groups and so were classed as unusual.

6.3.1.2.3. The overall form of suppression

Figure 6-10 shows three cartoon examples of how the suppression could develop as the number of conditioner tones was increased. All three examples have a similar frequency of maximum suppression irrespective of the number of conditioner tones. They differ in the relative increase in suppression bandwidth and magnitude with additional conditioner tones. The pattern shown in figure 6-10a can be discounted as I have already shown that the bandwidth of suppression increases with the number of conditioner tones. However, given the data presented so far, it is difficult to disambiguate between the examples shown in figure 6-10b and c.

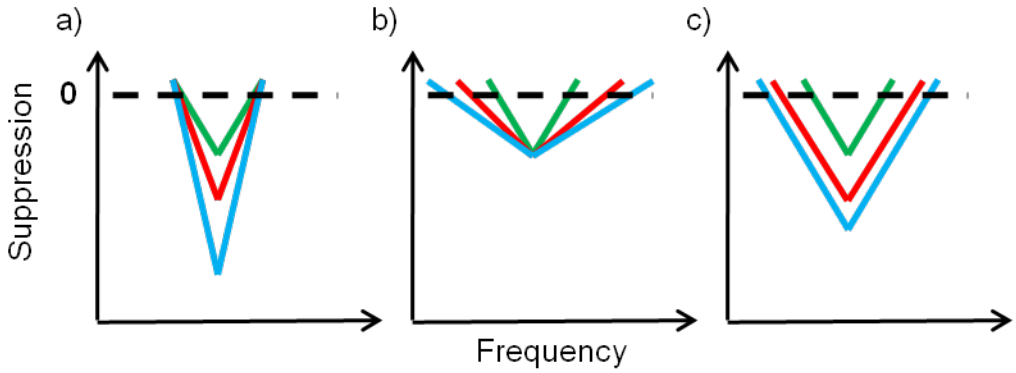


Figure 6-10. 3 examples of the form that suppression could take as the number of conditioner tones is increased from 1 (solid green) to 2 (solid red) or 4 (solid light blue). a) The bandwidth of the suppression does not change but the magnitude increases. b) The magnitude of suppression does not change but the bandwidth increases. c) Both the bandwidth and magnitude of suppression increase.

To choose between these two possibilities it is necessary to consider how the magnitude of suppression builds up with the number of conditioner tones. Figure 6-11 shows two analyses that help to decide between the two options. Firstly, the mean magnitude of the maximum suppression increases with the number of conditioner tones (shown in figure 6-11a). Secondly, the mean suppression patterns for the population, displayed in figure 6-11b, are commensurate with the pattern displayed in figure 6-10c, where the bandwidth and magnitude of suppression increase with the number of conditioner tones. The suppression that is observed is akin to an iceberg that rises, thus exhibiting more of the tip, with each additional conditioner tone.

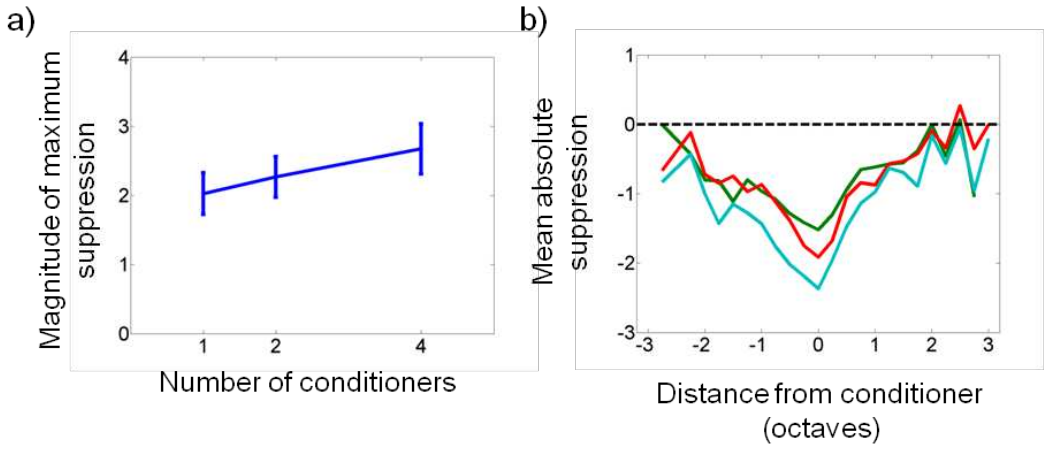


Figure 6-11. How the maximum suppression magnitude and the mean suppression change with the number of conditioner tones. a) The maximum suppression magnitude increases as the number of conditioner tones increases. b) The mean absolute suppression for the population of units as a function of the distance from the conditioner tone frequency, after 1 (solid green), 2 (solid red) and 4 (solid light blue) conditioner tones.

In order to compare the tuning of suppression with that reported at longer timescales the percentage tuning was calculated in exactly the same way as previously (Condon and Weinberger, 1991). Specifically, individual FR difference functions were normalised by dividing the difference FR by the absolute maximum value in that difference FR and then multiplying by 100:

$$normdiffFR = \frac{SFR - CFR}{\left|\max(SFR - CFR)\right|} * 100$$

Equation 6-1

where *SFR* is the suppressed FR, *CFR* is the control FR and *normdiffFR* is the normalised difference FR. Figure 6-12a shows the mean normalised difference FR which can be compared directly to the difference FR in figure 6-12b (from Condon and Weinberger, 1991).

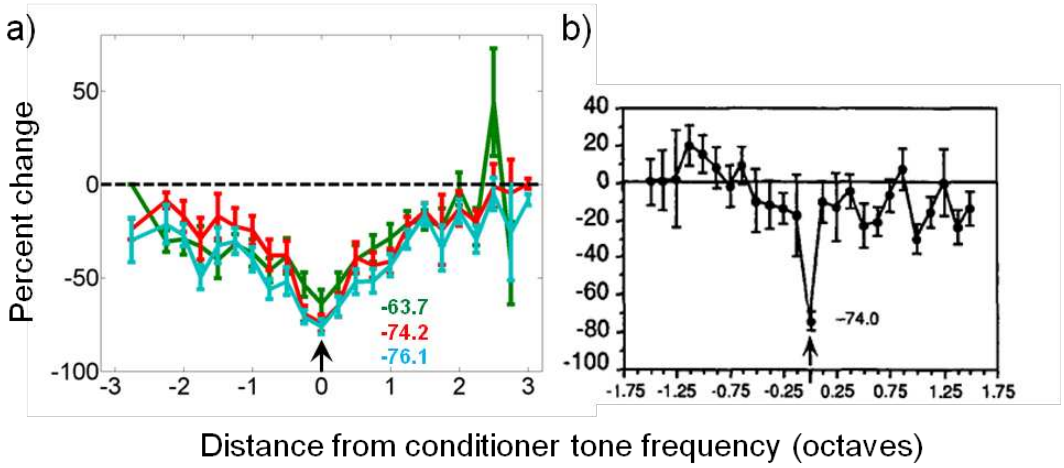


Figure 6-12. Comparison of short-term and long-term tuning of suppression. The mean percentage change in suppression, calculated using equation 6-1, as a function of the distance from the conditioner tone frequency. The arrow indicates the frequency of the conditioner tone and the numbers close to the minimum points indicate the magnitude of the percentage suppression at that point. Bars show standard error. a) Data from this study. b) Data from a study on long-term suppression in guinea pig (Condon and Weinberger, 1991).

Suppression over the long-term presentation of isofrequency tones is much narrower than that observed in this study, though the peak percentage decrease is similar after 2 or 4 conditioner tones (indicated as red and cyan numbers in figure 6-12a and a black number in figure 6-12b). Suppression in figure 6-12b is characterised by a single narrow peak at the conditioner tone frequency with little suppression at other frequencies. In figure 6-12a, the suppression is still centred on the conditioner tone frequency but there is a gradual decrease in suppression as the frequency difference is increased in either direction leading to a broader suppression.

6.3.1.2.4. Is the amount of SSA linked to characteristics of the unit?

A salient question is whether the degree to which suppression is frequency specific, as opposed to general, is related to characteristics of the unit. The ratio of the maximum suppression to the mean suppression (the FSR, see methods) was used to grade the specificity of the suppression. The distance of the conditioner from the BF proved to be a poor predictor of any facet of suppression apart from the locus of maximal suppression, displaying no significant relationship with the suppression bandwidth or the FSR. This

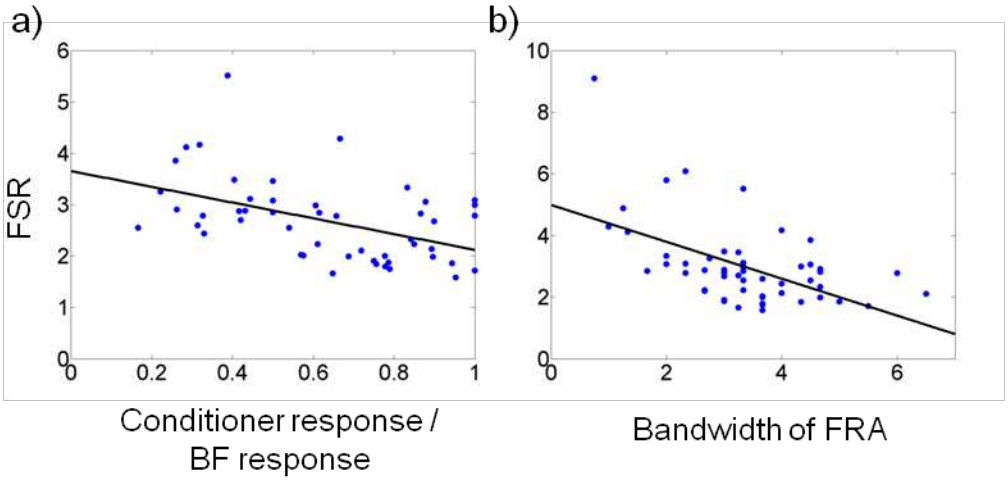


Figure 6-13. 2 measures that significantly predict the FSR. The FSR as a function of a) the ratio of the conditioner tone response and the response at BF and b) the bandwidth of the control FR. The 4 conditioner tone case is shown here, but significant relationships were observed irrespective of the number of conditioner tones. The solid black lines indicate linear regressions of the data.

may not be surprising given that the analogue of this measure (the distance from the probe to the CF/BF) was also a poor predictor for suppression characteristics in chapter 4.

Two measures that significantly predict the FSR are displayed in Figure 6-13. There is a small (gradient = -1.5) but significant ($p<<0.01$) negative correlation between the FSR and the response to the conditioner normalised to the response at BF ($R = 0.47$). This implies that the greater the response to the conditioner the more general the suppression is. The story is complicated by the fact that the bandwidth of the FR is also a significant predictor of the FSR (gradient = -0.6, $R = 0.54$). This would imply that frequency specific adaptation is more likely to be observed for FRs with narrower tuning. A confounding factor, however, is that there is a slight positive correlation between the two predictors (gradient = 0.07, $R = 0.3$, $p=0.04$).

6.3.1.2.5. How does the BF change after suppression?

I have shown that the maximum suppression is most often observed at and around the conditioner tone frequency. A further question is whether the suppressed FR exhibits the same preference for certain stimuli as the control FR, for example does a neuron have the same BF? To investigate this question directly, for each suppressed FR the BF was calculated as the frequency at which the response was greatest. Figure 6-14a shows a histogram of the distance between the suppressed BF and the control BF for each number of conditioner tones. Around 60% of units have suppressed BFs within 0.5 octaves of the control BF, irrespective of the number of conditioner tones.

Even though the suppressed BF could change markedly from the control BF, the response at the control BF within each suppressed FR was similar to the response at the suppressed BF, as shown in figure 6-14b. The relationship between the response at the control and suppressed BFs was confirmed statistically for each number of conditioner tones (Spearman's Rho: r = 0.90 – 0.95, p << 0.01). Thus, units continued to respond well, if not maximally, to a tone presented at the BF, irrespective of the frequency or number of conditioner tones that preceded it.

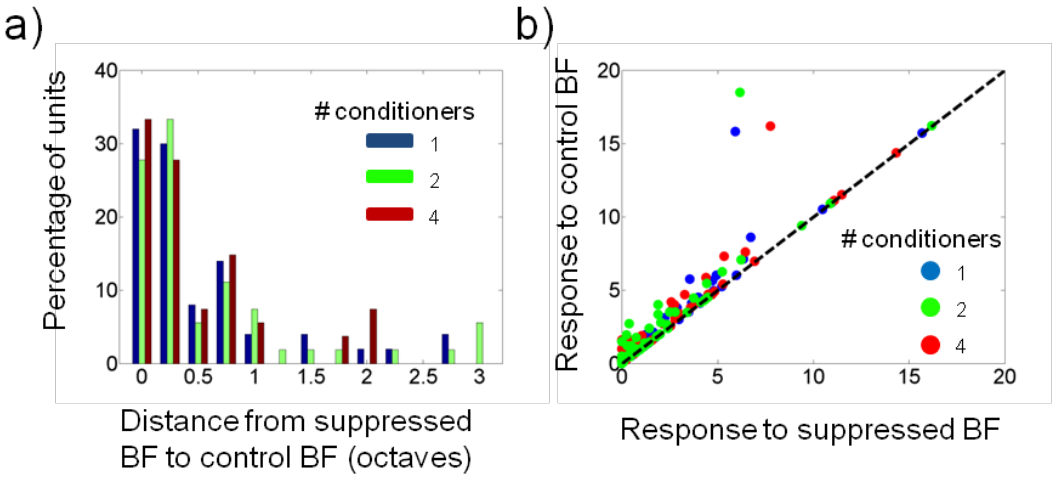


Figure 6-14. Measures of the suppressed BF relative to the control BF. a) Histogram of the distance from the suppressed BF to the control BF. b) The response at the BF of the control FR as a function of the response at the BF of the SFR.

6.3.2. The recovery from suppression

The aim of this study was to investigate the effect of varying the number of conditioner tones on the recovery from suppression. Due to the large number of repetitions of the conditioner tones (one for each repetition of each probe tone gap) it was generally possible to fit the spike count

responses to successive conditioner tones. The objective was then to compare the properties of these fits with characteristics of the recovery functions.

6.3.2.1. Unit examples

The conditioner tones were presented at 4, 8 or 16 Hz. The majority of units were presented with conditioner tones at 4 Hz, as indicated in table 6-3, and so the population analysis will be limited to these units. Unit examples for all three PR groups will be presented.

PR	4 Hz	8 Hz	16 Hz	Total
SU	7	4	2	13
MU	31	2	1	34
Total	38	6	3	47

Table 6-3. Proportion of units tested at each PR.

6.3.2.1.1. Conditioner tone presentation rate of 16 Hz

Figure 6-15 shows two example units when the conditioner tones were presented at 16 Hz. The right panel in figure 6-15a shows the recovery functions for a multi-unit as the number of conditioner tones is varied. The general trend that can be observed is that the recovery functions move vertically as the number of conditioner tones is increased. With one

conditioner tone the probe tone response recovers to the control response after around 60 ms, however with 4 or above conditioner tones the unit fails to recover to control within the 200 ms period tested. One question is the extent to which the recovery can be predicted from the adaptation in the response to the conditioner tones (shown in the left panels in figure 6-15). In general, the adaptation was well fitted with an exponential function, however in this example, and in the example shown in figure 6-15b, the adaptation profile is unusual and so was not characterised using a fitted curve. Nevertheless, it is possible that there is a relationship between the recovery and the spike count response to the last conditioner tone that occurred before the probe tone (indicated in colours corresponding to the number of conditioner tones in the left panels of figure 6-15).

Unfortunately, a relationship between recovery and the response to the last conditioner tone is not obvious in the example single unit displayed in figure 6-15b. The adaptation profile is similar to the unit in figure 6-15a, however the change in recovery with different numbers of conditioner tones is less obvious. Indeed, the unit takes longer to recover after one conditioner tone than it does after more than one conditioner tone.

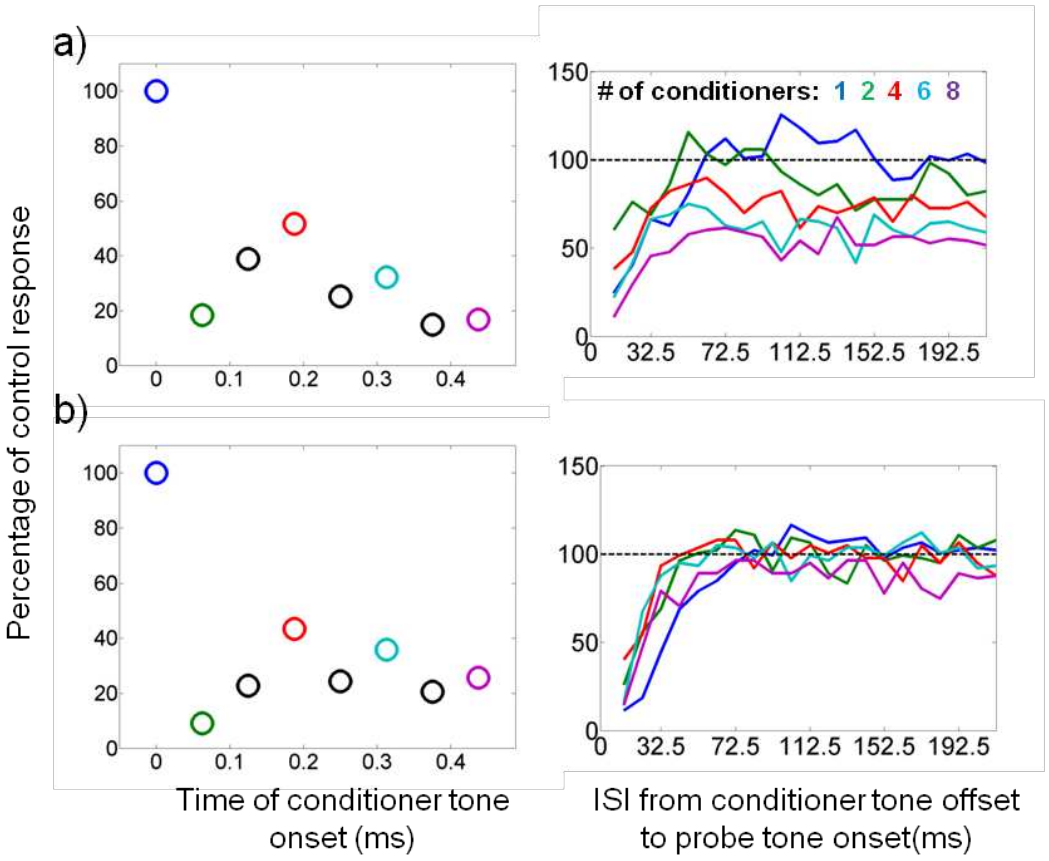


Figure 6-15. Examples of adaptation and recovery functions for units at a PR of 16 Hz. The number of conditioner tones for each recovery function is colour-coded and displayed in the right panel of a). The last conditioner tone before each recovery function is also colour-coded to correspond to the number of conditioner tones.

6.3.2.1.2. Conditioner tone presentation rate of 8 Hz

Three units that were tested with conditioner tones presented at 8 Hz are displayed in figure 6-16 in the same format that the units were presented in figure 6-15. The effect of increasing the number of conditioner tones is similar with a PR of 8 Hz as it was with a PR of 16 Hz when the first 200 ms of recovery is considered (refer to figure 6-16a and b). An increase in the number of conditioner tones causes a slower recovery such that the unit takes longer to recover to the control response or does not recover at all

within the 200 ms tested. The adaptation profiles shown in the panels to the left of figure 6-16a and b are different to those for the 16 Hz PR, exhibiting a slower decrease in response. There is a qualitative relationship between the response to the conditioner tones and the difference in the recovery functions as more conditioner tones are added. To highlight this, dotted lines have been drawn on to the adaptation profiles to show the difference in response to the last conditioner tone that is presented before the probe tones. Both the recovery curves and the highlighted tones in the adaptation profile are spaced wider for the unit in figure 6-16a than they are for the unit in figure 6-16b. I will return to this qualitative relationship in section 6.3.2.2 when I assess the population of units tested at a 4 Hz PR. The unit in figure 6-16c has an adaptation profile that is fitted well with an exponential function and it displays a large drop in response after the first conditioner tone. It is therefore not surprising that the recovery functions for this unit are generally similar. This unit also displays an overlap of the last conditioner tone response with the probe tone responses at low ISIs (the vertical dashed line indicates the point at which overlap should cease to occur, estimated from the manually-set response time window). For units in which overlap occurred, only those points on or after the vertical dashed line were used to derive statistics that characterised the recovery function.

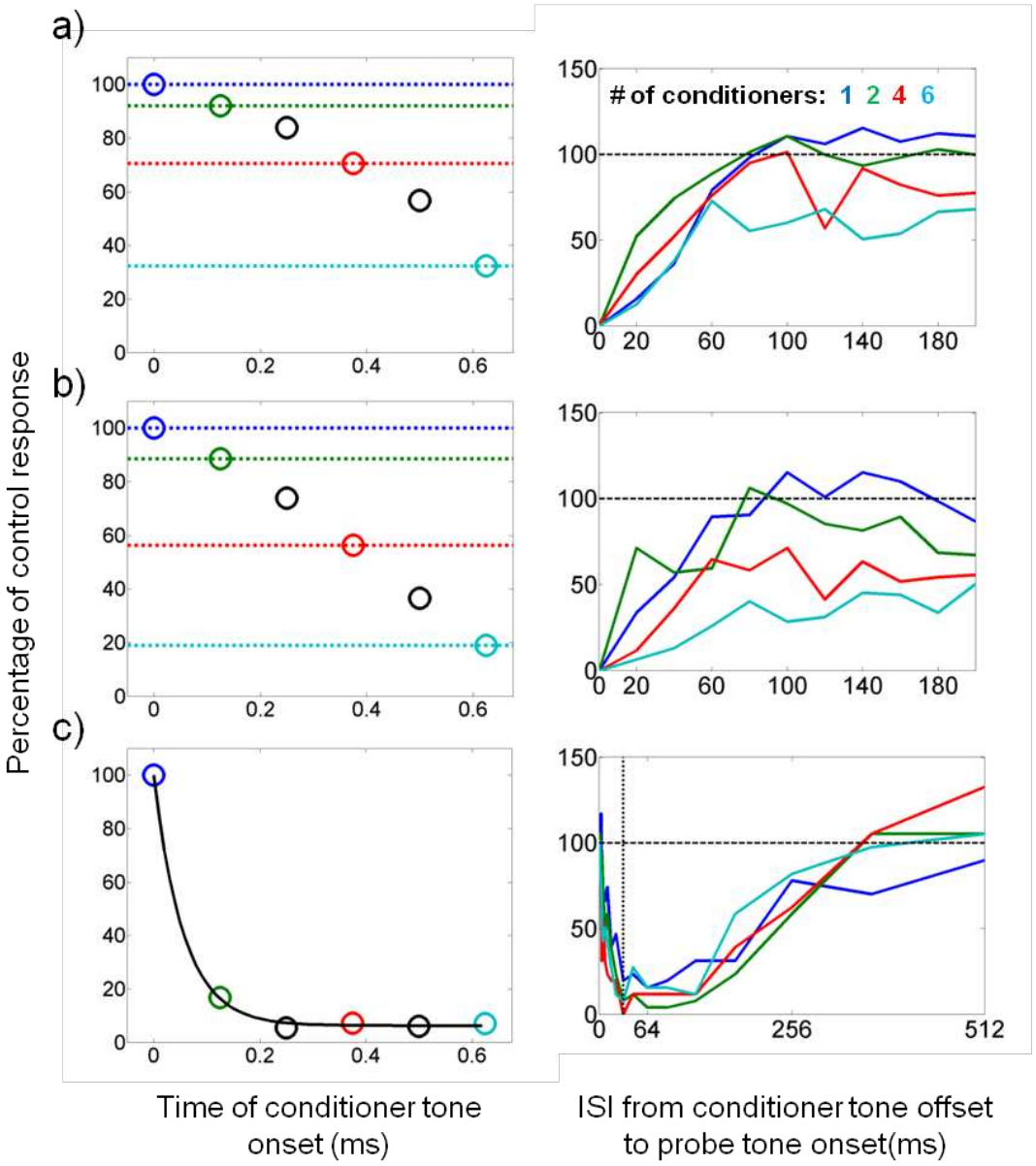


Figure 6-16. Example units when the conditioner tones were presented at 8 Hz. The adaptation of the unit in c) has been fitted with an exponential function. Same colour-code conventions as figure 6-15.

6.3.2.1.3. Conditioner tone presentation rate of 4 Hz

The adaptation spike counts for those units in which the conditioner tones had been presented at 4 Hz were well fit with an exponential function

(see methods for a description of the fitting). The units were visually

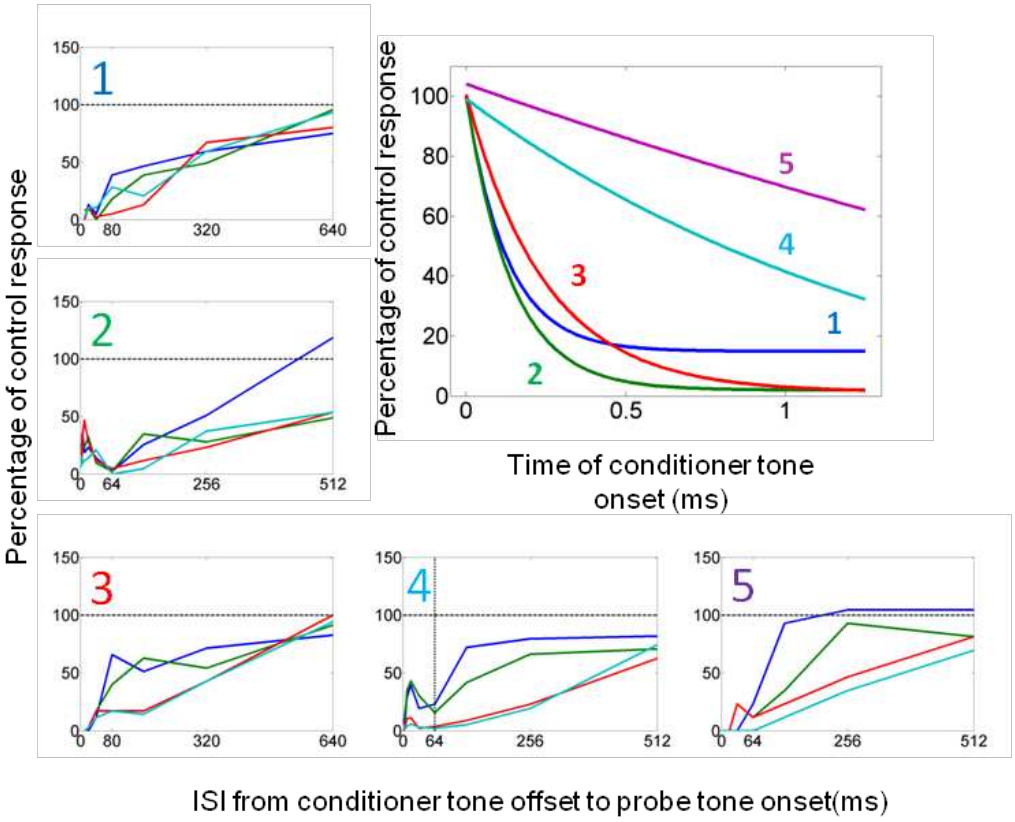


Figure 6-17. How the adaptation profile relates to the recovery functions for the 4 Hz PR conditions. The main panel shows the fitted exponential adaptation functions for 5 units while each of the smaller panels shows the recovery from suppression for each of the units. The units are ranked in terms of the adaptation time constant with unit 1 having the lowest (the fastest to adapt) and unit 5 having the highest (the slowest to adapt). Same colour-code conventions as figure 6-15.

inspected in ascending order of time constant to assess whether any qualitative relationships between the adaptation and the recovery could be observed. Figure 6-17 shows 5 example units ordered in terms of the time constant of adaptation, with the fitted adaptation function also shown. There is a gradual increase in the disparity between the recovery functions for different numbers of conditioner tones as the time constant increases. So for

unit 1, the recovery functions are similar irrespective of the number of conditioner tones, while for unit 5 there is a slowing of the rate of recovery for each addition of conditioner tones. This implies that there may be a relationship over the population of units between the rate at which a unit adapts and the difference in recovery functions with different numbers of conditioner tones.

6.3.2.2. Population analysis

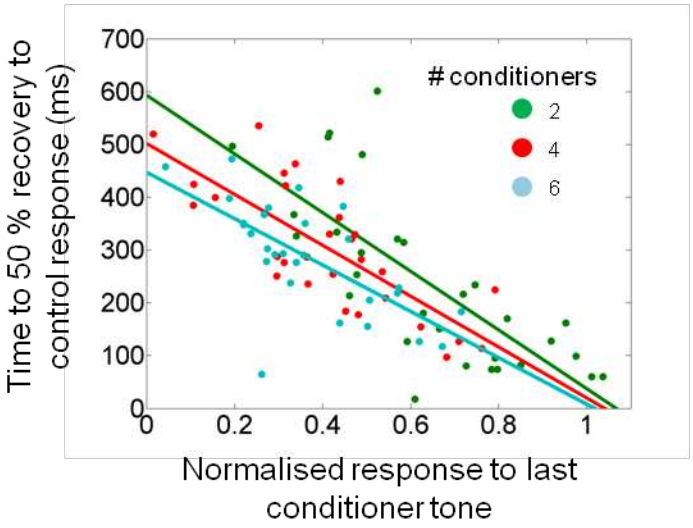


Figure 6-18. Relationship between the normalised response to the last conditioner tone and the time to 50% recovery.

To investigate whether the qualitative relationship between the adaptation and the recovery could be analysed quantitatively, various measures of each were compared statistically using linear models. The adaptation was characterised with the coefficient of the fitted exponential function and the response to the last conditioner tone in each condition, normalised to the control response. The recovery was characterised as

discussed in section 6.2.2.4, with 3 time points: the recovery from spontaneous activity, the recovery to control and the recovery to 50% of control (t50). Of these 3 measures, the t50 and the recovery from spontaneous activity showed significant relationships with the response to the last conditioner tone. The time constant of adaptation was not correlated with any of the measures of recovery.

Figure 6-18 shows that there is a negative correlation between the response to the last conditioner tone (normalised to the control response) and the t50. So, as the response to the last conditioner tone increases the unit tends to recover quicker, possibly because the unit is in a less adapted state. There is also a correlation between the response to the last conditioner tone and the time at which the response recovers from spontaneous activity. The gradients and correlation coefficients for each time measure are shown in table 6-4.

Number of conditioner tones	Time to 50% recovery (R value)	Time to recovery from spontaneous (R value)
2	-560 (0.75)**	-390(0.56) **
4	-480 (0.77)**	-490 (0.56) **
6	-440 (0.67)**	-300 (0.37)*

Table 6-4. Gradients of linear regressions of t50 and the time to recovery from spontaneous activity with the response to the last conditioner tone. * *p<0.05*. ** *p<0.01*.

6.4. Discussion

6.4.1. Summary of findings

6.4.1.1. The FR

- Qualitatively, the pattern of suppression for a given unit could vary along a continuum from a general suppression that affected all frequencies in the FR to a more specific suppression around the conditioner tone frequency

- Suppression that could be considered as a sharpening of the FR was only observed in two units

- The maximum suppression after one conditioner tone was generally at or around the frequency of the conditioner

- As the number of conditioner tones was increased the gradient of the linear regression of the maximum suppression and the frequency of the conditioner tone increased while the variability decreased: suppression became more frequency specific

- The bandwidth of suppression increased as the number of conditioner tones increased and so suppression can also be considered to have become less frequency specific with additional conditioner tones

- The bandwidth of suppression for individual units displayed varying degrees of dependence on the number of conditioner tones

- The magnitude of suppression also increased with the number of conditioner tones

- The tuning of the frequency-specific suppression was considerably wider than that observed at longer timescales

- Suppression tended to be more specific (as indexed by the FSR) for conditioner tones that elicited lower spike count responses and in units that had narrower tuning

- However, a small positive relationship between the response to the conditioner and the FR bandwidth make it difficult to rule out the possibility that there is an interaction between the two predictors

- The distance of the conditioner from the BF was not found to be a good predictor of suppression characteristics, apart from the locus of maximal suppression

- In general, neurons continued to preferentially respond to probe tones at the best frequency irrespective of the frequency or the number of conditioner tones

6.4.1.2. The recovery from adaptation

- Units displayed varying degrees of difference between recovery functions as the number of conditioner tones was increased: in some cases recovery changed as the number of conditioner tones was increased, in others there was no difference

- Qualitatively, the amount by which recovery functions differed seemed to be related to the rate at which adaptation occurred: larger differences were observed when the units adapted slowly

- Quantitatively, there was a negative correlation between the response to the last conditioner tone (normalised to the control response) and both the time to 50% recovery and the time to recovery from spontaneous activity

6.4.2. Comparison with previous studies

6.4.2.1. The FRA

6.4.2.1.1. The frequency specificity of suppression

An informative way to think about the effects that have been described in this chapter is to consider the type of suppression that would be observed in AN fibres using this stimulus paradigm. Due to the simple relationship between the response to the conditioner tone and the amount of forward suppression, one would expect that the decrement in response at each frequency in the FRA would be almost equal. In support of this notion it has been shown that, at a single tone level, the amount of suppression does not vary with the frequency of the probe tone (Abbas, 1979). This general suppression is thus similar to that seen in some neurons in the cortex that have been reported in this chapter. It is possible that such neurons represent the output of a processing stream that has had little excitatory or inhibitory modification from the input at the AN, at least as far as this particular aspect

of processing is concerned. To investigate whether general suppression was a characteristic of the neuron, as opposed to a stimulus-related effect, it would be necessary to vary the conditioner frequency within the same neuron.

If general suppression represents a relatively unchanged feature that is inherited from earlier auditory processing stations, then the degree of frequency specific suppression indicates to what extent the AN response has been modified. These data indicate that the degree of specificity varies along a continuum, as does the relationship between the suppression bandwidth and the number of conditioners. Such variability of response features is a facet of the RFs of cortical neurons in A1 (Moshitch et al., 2006).

6.4.2.1.2. The relation to models of adaptation

The models of adaptation that were mentioned in the introduction (section 6.1) were developed to explain how *f*MRI adaptation occurred and so are not directly relevant to this study. However, it is interesting to consider how the data in this chapter would be viewed at the haemodynamic level. The data presented in this chapter are consistent with a mixture of the fatigue and sharpening models of adaptation. However, as effects of latency were not tested it is not possible to rule out the facilitative model. The general suppression that has been noted would result in a reduction in the response of neurons that were initially responsive, which is the definition of the fatigue model. Frequency-specific suppression, however, would lead to a reduction of response at a narrow band of frequencies, causing a sharpening of neural tuning to those frequencies that the neuron preferentially responds to. The

distinction here is that neurons that used to respond to the tone will no longer respond to it after repetition. If there is a relationship between the degree of frequency specificity of suppression and the distance of the conditioner tone from the preferred frequency then overall this would lead to a sharpening of tuning. At the population level, those neurons with a BF at the conditioner tone frequency would exhibit a general suppression, while those with a BF away from it would exhibit a decrease in response at the conditioner tone frequency that would lead to less neurons responding to the tone and thus a sharpening of the representation.

6.4.2.1.3. The relation to responses to alternating tone sequences

The correlation between the spike count in response to the conditioner tone and the degree of frequency specific suppression that is observed has implications for the way in which neurons respond to alternating tone sequences. No significant correlation was found between the distance of the conditioner tone from BF and the FSR. A direct comparison of the frequency specificity described in this study and the results of the streaming experiment presented in chapter 5 is not possible, because those data were analysed in terms of the difference in frequency between the 'A' and 'B' tones. However, frequency differences are only one means to measure the relation of the 'B' tone to the 'A' tone, and are susceptible to the various different shapes and bandwidths of RFs that have been described in auditory cortical neurons (Sutter, 2000; Moshitch et al., 2006). For example, the actual spike count response to a 'B' tone that is 0.5 octaves away from the

BF will vary depending on the bandwidth of the RF: at the two extremes it could still be well within the RF and elicit a large response relative to BF or it could be outside of it and thus elicit no response. It is therefore not surprising that the distance of the conditioner (or probe) from the BF is a poor predictor of many aspects of suppression over a population of neurons.

A different means by which one can measure the response of a tone relative to the BF is in the ratio of the spike counts. Such a measure does not suffer from the high variability of RF shapes described above, but it is used with the assumption that effects are to some extent due to response history rather than stimulus history. Here, there was a negative correlation between the conditioner tone response and the degree of specificity of the suppression. This implies that tones that elicit stronger responses (nearer to the BF, for example the 'A' tone) will suppress a wider range of frequencies of subsequent tones than will tones that elicit weaker responses (away from the BF, for example the 'B' tones). So, while the 'A' tone would suppress the 'B' tone due to the larger bandwidth of suppression, it may not be true that the 'B' tone suppresses the 'A' tone to the same extent. This could help to explain the differential adaptation of the 'A' and 'B' tones that is observed as the sequence continues and the 'dominance' of the 'A' tone response as the PR is increased. In support of this is the fact that, even when suppressed, units tended to respond maximally or near-maximally to the control best frequency. These hypotheses will be discussed in further detail in chapter 7.

6.4.2.1.4. The relation to responses to long-term FSA

The data presented in this chapter indicate that the locus of maximal suppression is most likely to be at the frequency of the conditioner tone and this is in agreement with a study that addressed suppression over a longer time course (Condon and Weinberger, 1991). However, the exquisite narrow tuning of suppression that has been reported at longer timescales is not apparent in the short-term suppression presented within this chapter. Instead, suppression is broadly tuned, spanning all of the frequencies that were tested (around 6 octaves). The suppression displayed in the mean population data at longer timescales generally reflects the pattern of suppression seen in all of the multi-units tested (Condon and Weinberger, 1991). In contrast, the mean population data in figure 6-12a is calculated based on units that displayed various degrees of frequency-specific suppression and thus the mean suppression is more broadly tuned. A general suppression such as that reported in this chapter was not reported at a longer timescale; all the units displayed suppression that was specific to a narrow band of frequencies around the conditioner tone frequency. Thus, it is likely that such narrowly-tuned and highly frequency-specific suppression develops over a much longer timescale than was used in this chapter.

6.4.2.2. The recovery from adaptation

The shapes of the recovery from suppression were variable, even within units as the number of conditioner tones was varied, making them difficult to fit universally with functions. In general, adaptation spike counts

were fit with exponential functions; however there was no correlation between the time constants of adaptation and the characteristics of recovery. A strong predictor of the recovery characteristics was the normalised response to the last conditioner tone. It is perhaps not surprising that this was the case: the time constant gives an indication of the rate at which a unit adapts without specifying where along the adaptation function the response is occurring. Conversely, the normalised response to the last conditioner tone indicates precisely how adapted the unit is. The correlation between adaptation and recovery characteristics is of interest, but further investigation is necessary to fully characterise the way in which the two interact, as discussed in section 6.4.2.3.

6.4.2.3. Improvements and further work

I have shown that there is a continuum of suppression from frequency-specific to general. One major question is whether this gradient is due to unit or conditioner tone differences. For example, is it true that the degree of frequency specificity depends on the distance of the conditioner tone from the BF? To investigate this it would be necessary to vary the frequency of the conditioner tone within the same unit. One aspect of the physiological responses to tone sequences that has still to be explained completely is the differential suppression of the response to the 'B' tone as the presentation rate is increased. So, even though the responses to both the 'A' and the 'B' tone decrease, the response to the 'B' tone decreases more, such that the ratio between the responses (B/A) also decreases. Although it is

likely that forward suppression has a role in this differential effect it is not known exactly how the suppression builds up. There is some evidence from this chapter that there are differential suppression effects depending on the frequency of the conditioner tone. By varying the presentation rate of the conditioner tones in the paradigm used here it would be possible to investigate whether the distinction between general and frequency-specific suppression became more apparent at faster rates.

There are two improvements to the paradigm used to investigate the effect of multiple conditioner tones on the recovery from adaptation, depending on the question that is being asked. If the focus of the study were to simply investigate the effect of increasing the number of conditioner tones on the recovery, then it would be necessary to characterise the recovery with fitted functions. To maximise the fit, a large number (more than were used in this study) of time samples and stimulus repetitions would be necessary. Another way of considering the recovery is within the framework of the tone sequences used to investigate streaming. Instead of using a large number of time samples, one could sample at times that are relevant to the rate at which the conditioner tones were presented. For example, for a 4 Hz PR, the time samples could be at 0, 250, 500, 750 and 1000 ms, points at which tones would occur within the streaming stimulus. Also, information about the position of the conditioner tone within the RF (for example the distance between the tone frequency and the BF and the level of the tone relative to

the RF threshold) would help to characterise differences between the ‘A’ and ‘B’ tone adaptation and recovery.

7. Discussion and conclusions

In this thesis I studied various aspects of the contextual responses to successive tones in the primary auditory cortex of the anaesthetised guinea pig. In chapter 4 I investigated whether probe tones with different frequencies (and levels) were suppressed by the same or different ranges of preceding conditioner tones. In chapter 5 I considered whether the responses to streaming stimuli in single units in the anaesthetised guinea pig were similar to those reported in both single units and ensemble neural activity in awake animals. Finally, in chapter 6 I examined both the spectral and temporal characteristics of the build up of suppression in a sequence of isofrequency tones.

7.1. A general model of suppression in the primary auditory cortex

The frequency-specificity of forward suppression and SSA suggest that at some level in the auditory system there is a convergence of neural inputs each of which can be selectively suppressed. The results presented in chapters 4 and 6 suggest some constraints by which a model of this convergence could be developed. Firstly, as mentioned, it should be possible to independently suppress specific inputs to the model neuron. Secondly, the frequencies to which the neuron responds most, including the BF, should be similar when the neuron is suppressed to when it is not. Thirdly, inputs should overlap to a large extent in order to explain the large bandwidth and threshold of suppression when the probe tone directly follows the

conditioner tone (when ISI = 0). The inputs should then be able to independently recover such that inputs that respond to frequencies that are closer to the probe tone recover at a slower rate than the other inputs.

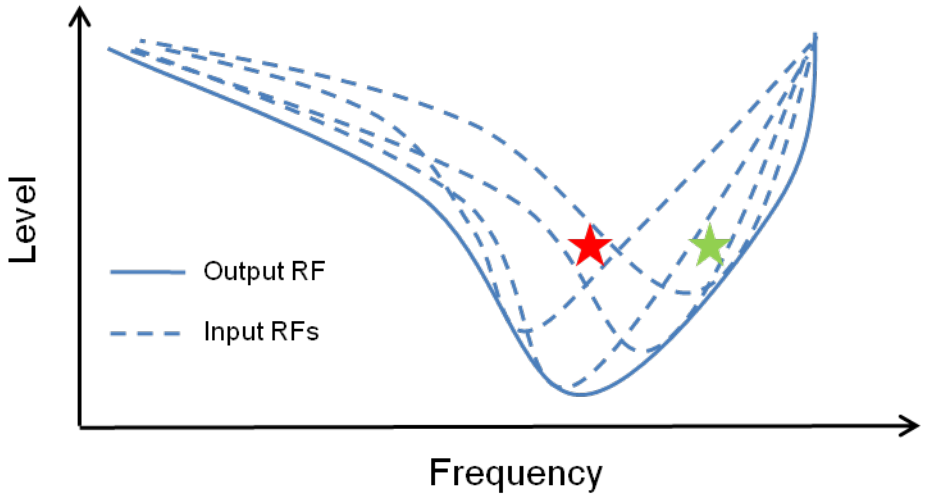


Figure 7-1. A schematic of the way in which a neural RF may be constructed from a convergence of inputs. The stars indicate the hypothetical position of two conditioner tones (explanation in text).

Figure 7-1 shows how the inputs to a neuron may be arranged given the constraints that I have outlined above. The figure is a schematic that is meant to illustrate the points that I have discussed, not a testable model of neural convergence. Also, the schematic only considers the RF that is constructed from spiking responses measured extracellularly. An intracellular study in IC demonstrated that inhibitory and subthreshold excitatory responses occurred over a wide (at least 1.5 – 2 octaves) frequency range (Xie et al., 2007). These responses, which are undetectable using extracellular recording techniques, may contribute to the frequency-specific responses

that have been observed in this thesis and in previous studies (Malone and Semple, 2001; Ulanovsky et al., 2003).

The model neuron displayed in figure 7-1 could display some of the features of suppression that have been reported in this thesis. For example, a conditioner tone presented around the BF of the neuron (indicated by a red star) would activate a large number of inputs to the neuron which would lead to stimulus-driven suppression within these inputs and thus a general suppression across the RF. In contrast, a conditioner tone presented away from BF (indicated by a green star) would predominantly cause activation and thus suppression of a narrower range of inputs leading to frequency-specific suppression of the RF.

7.2. Does cortical sequential two-tone forward suppression contribute to SSA?

A feature that characterises frequency-specific adaptation over the course of hundreds of milliseconds to seconds (Ulanovsky et al., 2003) to minutes (Condon and Weinberger, 1991) is the highly specific nature of the adaptation. The method used to compute the hyperacuity described by Ulanovsky et al. (2003) was based on signal detection theory and only considered discrimination of the responses to two tones. It is thus difficult to compare their results with those from chapters 4 and 6 where a larger number of tone frequencies were tested. However, Condon and Weinberger (1991) investigated the effect of repeating a single tone frequency on the frequency response areas of multi-units and so their results can be more

readily compared to those presented in this thesis. After around 5 - 7 minutes of the 1.25 Hz repetition of a single tone frequency the adaptation was highly specific (on the order of 0.125 octaves or less: Condon and Weinberger, 1991). In chapter 6, I used a similar paradigm to that of Condon and Weinberger (1991) but on a much shorter time scale with the same tone repeating at a rate of 4 Hz. At this shorter time scale the suppression was wider than that observed at longer time scales and became even wider as the number of conditioner tones was increased. So, as the number of conditioner tones increased suppression became less specific to the frequency of the conditioner tone. The wider bandwidth of suppression with tone repetition is the opposite of that observed by Condon and Weinberger (1991).

However, if one considers the larger temporal gap between the tones that were considered by Condon and Weinberger (1991) then the data from chapter 4 may suggest why suppression became highly frequency-specific. When the temporal gap between the conditioner and probe tones was increased in chapter 4, the bandwidth of tone frequencies that caused suppression decreased around the frequency of the probe tone. This implies that suppression becomes more frequency specific as the temporal gap is increased because the tuning of suppression becomes narrower. This finding may go some way to explaining the acutely narrow tuning of frequency specific suppression over long time courses (Condon and Weinberger, 1991; Ulanovsky et al., 2003). I showed in chapter 6 that the tuning of suppression over hundreds of milliseconds was considerably wider than that observed

over minutes (Condon and Weinberger, 1991). It is possible that such acute tuning is an extrapolation of the increase in frequency specificity with increasing temporal gap that was noted in chapter 4. Frequency specific effects may also depend on the duration of the conditioner tones. For example, the tones used in Condon and Weinberger (1991) were 300 ms in duration compared to 50 ms in chapter 4. In the IC, Malone and Semple (2001) noted that the duration of the conditioner tone could effect the time that a neuron took to recover from suppression. In addition they noted a significant negative relationship ($p<0.05$) between the response to the conditioner tone and the decrement in the response to the probe tone for the *quickstep* stimulus (200 ms tones with a 100 ms gap) but not for the *step* stimulus (2 second tones with no gap). This implies that non-specific (response history-driven) suppression dominates when short duration stimuli are used.

Further investigation would be necessary to support the notion that suppression becomes more frequency specific as the temporal gap and/or the duration of the conditioner tone increase. For example, the paradigm that was used in chapter 6 to investigate the effect of multiple conditioner tones on the FRA of a neuron could be adapted to directly investigate how the tuning of suppression changes at different time scales. The presentation rate of the conditioner tones could be varied to investigate the extent of frequency specific adaptation on the FRA at different rates. Additionally the specificity of suppression could be indexed at different points along a

sequence of isofrequency tones by varying the number of conditioner tones that occur before a FR is computed. It is possible that the narrowly tuned suppression observed after minutes of adaptation could be observed after a much shorter time. For example, the adaptation of responses to a sequence of tones at a single frequency with a 2 Hz PR is complete after around 2 seconds (see figure 5-14). If similar mechanisms lead to SSA at short and long time scales then one would expect to observe a decrease in the bandwidth of suppression as the time scale was gradually increased. The results of such an experiment would allow a direct comparison between the SSA observed over minutes with that observed over seconds and would thus allow more insight into whether the same mechanisms underlie SSA at all time scales.

7.3. Are the responses of neurons in the anaesthetised guinea pig consistent with those in awake preparations?

The steady-state responses of neurons in the anaesthetised guinea pig were generally in accordance with those observed in awake animals (Fishman et al., 2001; Fishman et al., 2004) apart from at the fastest PR tested (16 Hz). In contrast, the adaptation of responses displayed a recovery that was not consistent with the psychophysical build-up of streaming when a signal detection model was used to compare the two. If the adaptation was fitted with either exponential or power-law functions the time constant of adaptation was generally too fast to account for the psychophysical build-up. It is not clear if the recovery from adaptation was a consequence of the

anaesthetic, though it has not previously been reported in awake preparations (Micheyl et al., 2008). As further investigation into this phenomenon is necessary I will concentrate on the effects of urethane anaesthesia on the adaptation of responses as characterised using exponential or power-law functions.

The effects of anaesthetics on neural responses are complex, depending, for example, on the anaesthetic agent (Angel and Gratton, 1982; Angel, 1991; Astl et al., 1996), the species (Dodd and Capranica, 1992) and the auditory area under investigation (Capsius and Leppelsack, 1996). The anaesthetic used in this thesis was urethane (ethyl carbamate) and so this discussion will focus on literature pertaining to the effects of this drug. There are two main effects of urethane. Firstly there is a general decrease in neuronal firing manifested as a decrease in the spontaneous activity (Capsius and Leppelsack, 1996), a decrease in the stimulus-related response amplitude (Albrecht and Davidowa, 1989; Capsius and Leppelsack, 1996), a decrease in the probability of a response (Albrecht and Davidowa, 1989) and a decrease in the number of active recording sites (Capsius and Leppelsack, 1996). Secondly, there is an increase in the latency of responses (Angel and Gratton, 1982; Albrecht and Davidowa, 1989; Angel, 1991; Capsius and Leppelsack, 1996). Of these effects of urethane, there are two that may contribute to the differences observed between the results presented in this thesis and those from awake animals: the decrease in response amplitude and the decrease in the probability of a response. It has been also been shown that adaptation

leads to an increase in the proportion of failures in the responses of A1 neurons at longer time scales (Ulanovsky et al., 2004). The spike count responses from awake starling field L with 40 ms tones were around twice those observed in chapter 5, and for 125 ms tones were around 6 times as high (Micheyl et al., 2008). These effects taken together may explain the more rapid and complete adaptation observed in the anaesthetised guinea pig though this is an inference rather than a direct result. To support this hypothesis it would be necessary to directly compare the awake and anaesthetised responses to tone sequences in the same animal, which is a possible avenue for further research into the effects of anaesthesia on the responses to streaming stimuli.

The build-up of streaming has previously been accounted for using a signal detection model applied to the responses of VCN neurons in the urethane-anaesthetised guinea pig (Pressnitzer et al., 2008). It is therefore strange that the build-up could not be accounted for using cortical responses in the anaesthetised guinea pig. However, there are 2 possible reasons for this. Firstly, the effects of urethane have been shown to differ between populations from different fields of the avian analogue of the auditory cortex, with some showing less effect than others (Capsius and Leppelsack, 1996). Although not studied directly, it is possible that urethane anaesthesia affects the VCN in a different manner to the auditory cortex. Secondly, the stimuli used in the two experiments differ. Pressnitzer et al. (2008) used the ABA_ABA_ stimulus with 125 ms tones and gaps, while in this thesis, for

comparison with the studies of Fishman et al. (2001, 2004), the ABAB stimulus was used with 50 ms tones and (at 8 Hz) 75 ms gaps. It is difficult to compare the two stimuli directly. Another avenue for future investigations is to use the ABA_ABA_ in an anaesthetised guinea pig.

7.4. How does forward suppression develop into the neural responses observed using streaming stimuli?

The methods used in chapters 4 and 6 represent two different approaches to investigating forward suppression. In chapter 4 I investigated the influence of a range of conditioner tones of different frequencies and levels on a single probe tone (though more than one probe tone was presented for each unit). In chapter 6 I investigated the effect of repeatedly presenting a conditioner tone at a single frequency on a range of probe tones that constituted the frequency response area of a unit. One aspect of suppression that is strongly evidenced by both of the approaches is that the suppression is strongest when the conditioner and probe tones have similar characteristics (in this case frequency). This aspect of suppression has been reported extensively both for frequency (Brosch and Schreiner, 1997; Bartlett and Wang, 2005; Brosch and Scheich, 2008) and for binaural levels or spatial location (Reale and Brugge, 2000; Zhang et al., 2005; Nakamoto et al., 2006). Suppression thus acts to decrease the response to successive stimuli that have the same characteristics. However, while this may be true, it seems that the stimuli that a neuron preferentially responds to in quiet (with no preceding sounds) are the same stimuli that the neuron will preferentially

respond to in the presence of context (after preceding sounds). So, even though the greatest suppression was observed at the frequency of the conditioner tone in chapter 6, the response at BF was either at or close to the maximum response irrespective of the frequency or the number of conditioner tones. This suggests that suppression acts to enhance the selectivity of neurons that were previously sensitive to a larger range of stimuli and were thus optimized for detection. Indeed, the type of suppression that was observed (general or frequency-specific) was correlated to the normalised response to the conditioner tone. A more general suppression was observed (the frequency-specificity ratio was lower) when the response to the conditioner tone was greater and a more frequency-specific suppression was observed when the conditioner tone response was smaller. This pattern of suppression results in a maintenance of the response at the frequency to which the neuron preferentially responds, as is shown diagrammatically in figure 7-2.

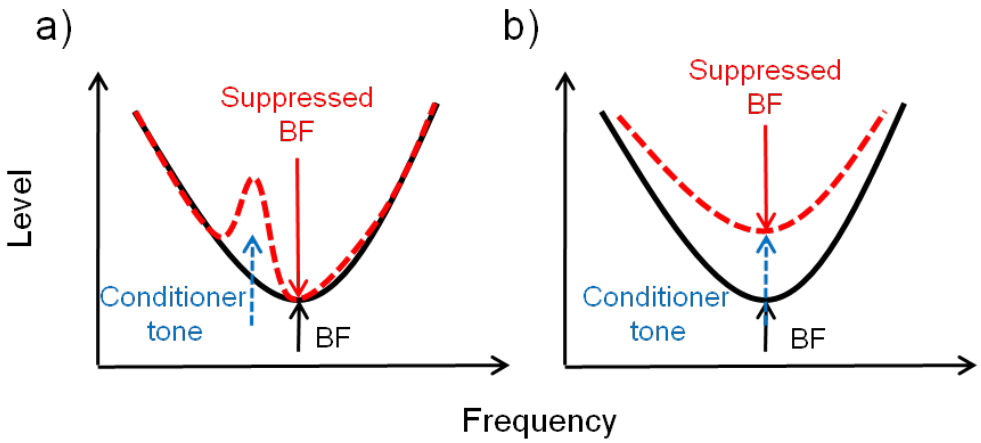


Figure 7-2. The suppressed BF remains similar to the BF for both a) frequency-specific and b) general suppression patterns.

The notion that the BF response is more robust to suppression than weaker responses is consistent with the responses to tone sequences and helps to explain why the response to the ‘A’ tone dominates the response to the ‘B’ tone. The general suppression observed with tones that are closer to the BF (the ‘A’ tone) will be manifested as a decrease in the subsequent response to both the ‘A’ and ‘B’ tones. Conversely, the more frequency-specific suppression observed with the non-BF ‘B’ tone will only be manifested as a decrease in the subsequent response to ‘B’ tones. One would predict that this effect would be magnified if the presentation rate was increased and indeed the results from chapter 5 show an increase in the differential suppression with an increase in PR (the B/A ratio decreases).

8. Bibliography

Abbas PJ (1979) Effects of stimulus frequency on adaptation in auditory-nerve fibers. J Acoust Soc Am 65:162-165.

Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. Science 275:220-224.

Abeles M, Goldstein MH, Jr. (1972) Responses of single units in the primary auditory cortex of the cat to tones and to tone pairs. Brain Res 42:337-352.

Ainsworth A, Dostrovsky JO, Merrill EG, Millar J (1977) An improved method for insulating tungsten micro-electrodes with glass [proceedings]. J Physiol 269:4P-5P.

Aitkin LM, Webster WR (1972) Medial geniculate body of the cat: organization and responses to tonal stimuli of neurons in ventral division. J Neurophysiol 35:365-380.

Aitkin LM, Prain SM (1974) Medial geniculate body: unit responses in the awake cat. J Neurophysiol 37:512-521.

Aitkin LM, Kenyon CE, Philpott P (1981) The representation of the auditory and somatosensory systems in the external nucleus of the cat inferior colliculus. J Comp Neurol 196:25-40.

Aitkin LM, Webster WR, Veale JL, Crosby DC (1975) Inferior colliculus. I. Comparison of response properties of neurons in central, pericentral, and external nuclei of adult cat. J Neurophysiol 38:1196-1207.

Albrecht D, Davidowa H (1989) Action of urethane on dorsal lateral geniculate neurons. Brain Res Bull 22:923-927.

Alho K (1995) Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hear 16:38-51.

Angel A (1991) The G. L. Brown lecture. Adventures in anaesthesia. Exp Physiol 76:1-38.

Angel A, Gratton DA (1982) The effect of anaesthetic agents on cerebral cortical responses in the rat. Br J Pharmacol 76:541-549.

Anstis S, Saida S (1985) Adaptation to auditory streaming of frequency-modulated tones. J Exp Psychol Hum Percept Perform 11:257-271.

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

Atienza M, Cantero JL (2001) Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity (MMN). Brain Res 901:151-160.

Bartlett EL, Wang X (2005) Long-lasting modulation by stimulus context in primate auditory cortex. J Neurophysiol 94:83-104.

Beauvois MW, Meddis R (1996) Computer simulation of auditory stream segregation in alternating-tone sequences. J Acoust Soc Am 99:2270-2280.

Bee MA, Klump GM (2004) Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. J Neurophysiol 92:1088-1104.

Bee MA, Klump GM (2005) Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences. Brain Behav Evol 66:197-214.

Bergerbest D, Ghahremani DG, Gabrieli JD (2004) Neural correlates of auditory repetition priming: reduced fMRI activation in the auditory cortex. J Cogn Neurosci 16:966-977.

Best AR, Wilson DA (2004) Coordinate synaptic mechanisms contributing to olfactory cortical adaptation. J Neurosci 24:652-660.

Bilger RC (1959) Additivity of Different Types of Masking. J Acoust Soc Am 31:1107-1109.

Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, Hyde JS (1994) Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. Brain Res Cogn Brain Res 2:31-38.

Blackburn CC, Sachs MB (1989) Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. J Neurophysiol 62:1303-1329.

Bleeck S, Sayles M, Ingham NJ, Winter IM (2006) The time course of recovery from suppression and facilitation from single units in the mammalian cochlear nucleus. Hear Res 212:176-184.

Bock GR, Webster WR, Aitkin LM (1972) Discharge patterns of single units in inferior colliculus of the alert cat. J Neurophysiol 35:265-277.

Boettcher FA, Salvi RJ, Saunders SS (1990) Recovery from short-term adaptation in single neurons in the cochlear nucleus. Hear Res 48:125-144.

Bregman AS (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Bregman AS, Pinker S (1978) Auditory streaming and the building of timbre. Can J Psychol 32:19-31.

Bregman AS, Ahad PA, Crum PA, O'Reilly J (2000) Effects of time intervals and tone durations on auditory stream segregation. Percept Psychophys 62:626-636.

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

Brosch M, Schreiner CE (2000) Sequence sensitivity of neurons in cat primary auditory cortex. Cereb Cortex 10:1155-1167.

Brosch M, Scheich H (2008) Tone-sequence analysis in the auditory cortex of awake macaque monkeys. Exp Brain Res 184:349-361.

Brosch M, Schulz A, Scheich H (1998) Neuronal mechanisms of auditory backward recognition masking in macaque auditory cortex. Neuroreport 9:2551-2555.

Brosch M, Schulz A, Scheich H (1999) Processing of sound sequences in macaque auditory cortex: response enhancement. J Neurophysiol 82:1542-1559.

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.

Buunen TJ, Rhode WS (1978) Responses of fibers in the cat's auditory nerve to the cubic difference tone. J Acoust Soc Am 64:772-781.

Caird D, Klinke R (1987) Processing of interaural time and intensity differences in the cat inferior colliculus. Exp Brain Res 68:379-392.

Calford MB (1983) The parcellation of the medial geniculate body of the cat defined by the auditory response properties of single units. J Neurosci 3:2350-2364.

Calford MB, Aitkin LM (1983) Ascending projections to the medial geniculate body of the cat: evidence for multiple, parallel auditory pathways through thalamus. J Neurosci 3:2365-2380.

Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. J Neurophysiol 73:1876-1891.

Capsius B, Leppelsack HJ (1996) Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. Hear Res 96:59-70.

Carandini M, Heeger DJ, Senn W (2002) A synaptic explanation of suppression in visual cortex. J Neurosci 22:10053-10065.

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

Carlyon RP (2004) How the brain separates sounds. Trends Cogn Sci 8:465-471.

Carlyon RP, Cusack R, Foxton JM, Robertson IH (2001) Effects of attention and unilateral neglect on auditory stream segregation. J Exp Psychol Hum Percept Perform 27:115-127.

Carlyon RP, Plack CJ, Fantini DA, Cusack R (2003) Cross-modal and non-sensory influences on auditory streaming. Perception 32:1393-1402.

Carver FW, Fuchs A, Jantzen KJ, Kelso JA (2002) Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation: rate dependence and transient to steady-state transition. Clin Neurophysiol 113:1921-1931.

Chimento TC, Schreiner CE (1991) Adaptation and recovery from adaptation in single fiber responses of the cat auditory nerve. J Acoust Soc Am 90:263-273.

Chung S, Li X, Nelson SB (2002) Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. Neuron 34:437-446.

Cody AR, Russell IJ (1987) The response of hair cells in the basal turn of the guinea-pig cochlea to tones. J Physiol 383:551-569.

Condon CD, Weinberger NM (1991) Habituation produces frequency-specific plasticity of receptive fields in the auditory cortex. Behav Neurosci 105:416-430.

Creutzfeldt O, Hellweg FC, Schreiner C (1980) Thalamocortical transformation of responses to complex auditory stimuli. Exp Brain Res 39:87-104.

Cusack R (2005) The intraparietal sulcus and perceptual organization. J Cogn Neurosci 17:641-651.

Cusack R, Deeks J, Aikman G, Carlyon RP (2004) Effects of location, frequency region, and time course of selective attention on auditory scene analysis. J Exp Psychol Hum Percept Perform 30:643-656.

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

Dean I, Robinson BL, Harper NS, McAlpine D (2008) Rapid neural adaptation to sound level statistics. J Neurosci 28:6430-6438.

Denham S (2001) Cortical synaptic depression and auditory perception. In: Computational Models of Auditory Function (Greenberg S, Slaney M, eds), pp 281-296. Amsterdam: IOS Press.

Deutsch D (1974) An auditory illusion. Nature 251:307-309.

Diamond IT, Neff WD (1957) Ablation of temporal cortex and discrimination of auditory patterns. J Neurophysiol 20:300-315.

Diamond IT, Goldberg JM, Neff WD (1962) Tonal discrimination after ablation of auditory cortex. J Neurophysiol 25:223-235.

Dodd F, Capranica RR (1992) A comparison of anesthetic agents and their effects on the response properties of the peripheral auditory system. Hear Res 62:173-180.

Duifhuis H (1973) Consequences of peripheral frequency selectivity for nonsimultaneous masking. J Acoust Soc Am 54:1471-1488.

Eggermont JJ (1991) Rate and synchronization measures of periodicity coding in cat primary auditory cortex. Hear Res 56:153-167.

Eggermont JJ (1992) Stimulus induced and spontaneous rhythmic firing of single units in cat primary auditory cortex. Hear Res 61:1-11.

Eggermont JJ (1999) The magnitude and phase of temporal modulation transfer functions in cat auditory cortex. J Neurosci 19:2780-2788.

Ehret GN, Romand R (1997) The central auditory system. Oxford: Oxford University Press.

Evans EF (1975) Cochlear nerve and cochlear nucleus. In: Handbook of sensory physiology (Keidel WD, Neff WD, eds), pp 1-108. Berlin: Springer.

Evans EF, Whitfield IC (1964) Classification of Unit Responses in the Auditory Cortex of the Unanaesthetized and Unrestrained Cat. J Physiol 171:476-493.

Evans EF, Pratt SR, Spenner H, Cooper NP, Versnel H, Schoonhoven R, Kollmeier B, Harrison RV, Pickles JO, Scharf B (1992) Comparisons of Physiological and Behavioral Properties - Auditory Frequency-Selectivity. Auditory Physiology and Perception 83:159-169.

Eytan D, Brenner N, Marom S (2003) Selective adaptation in networks of cortical neurons. J Neurosci 23:9349-9356.

Fay RR (1998) Auditory stream segregation in goldfish (Carassius auratus). Hear Res 120:69-76.

Finlayson PG (1999) Post-stimulatory suppression, facilitation and tuning for delays shape responses of inferior colliculus neurons to sequential pure tones. Hear Res 131:177-194.

Finlayson PG, Adam TJ (1997) Excitatory and inhibitory response adaptation in the superior olive complex affects binaural acoustic processing. Hear Res 103:1-18.

Fischer C, Morlet D, Giard M (2000) Mismatch negativity and N100 in comatose patients. Audiol Neurootol 5:192-197.

Fishman YI, Arezzo JC, Steinschneider M (2004) Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. J Acoust Soc Am 116:1656-1670.

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2001) Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hear Res 151:167-187.

Fitzpatrick DC, Kuwada S, Kim DO, Parham K, Batra R (1999) Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. J Acoust Soc Am 106:3460-3472.

Freeman TC, Durand S, Kiper DC, Carandini M (2002) Suppression without inhibition in visual cortex. Neuron 35:759-771.

Gaese BH, Ostwald J (2001) Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. J Neurophysiol 86:1062-1066.

Gatehouse S, Noble W (2004) The Speech, Spatial and Qualities of Hearing Scale (SSQ). Int J Audiol 43:85-99.

Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol 32:613-636.

Gollisch T, Herz AV (2004) Input-driven components of spike-frequency adaptation can be unmasked in vivo. J Neurosci 24:7435-7444.

Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10:14-23.

Grimault N, Bacon SP, Micheyl C (2002) Auditory stream segregation on the basis of amplitude-modulation rate. J Acoust Soc Am 111:1340-1348.

Grimault N, Micheyl C, Carlyon RP, Arthaud P, Collet L (2000) Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency. J Acoust Soc Am 108:263-271.

Gutschalk A, Oxenham AJ, Micheyl C, Wilson EC, Melcher JR (2007) Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. J Neurosci 27:13074-13081.

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25:5382-5388.

Hari R, Kaila K, Katila T, Tuomisto T, Varpula T (1982) Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implications for their neural generation. Electroencephalogr Clin Neurophysiol 54:561-569.

Harms MP, Melcher JR (2002) Sound repetition rate in the human auditory pathway: representations in the waveshape and amplitude of fMRI activation. J Neurophysiol 88:1433-1450.

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

Heil P, Rajan R, Irvine DR (1994) Topographic representation of tone intensity along the isofrequency axis of cat primary auditory cortex. Hear Res 76:188-202.

Herbert H, Aschoff A, Ostwald J (1991) Topography of projections from the auditory cortex to the inferior colliculus in the rat. J Comp Neurol 304:103-122.

Hocherman S, Gilat E (1981) Dependence of auditory cortex evoked unit activity on interstimulus interval in the cat. J Neurophysiol 45:987-997.

Hulse SH, MacDougall-Shackleton SA, Wisniewski AB (1997) Auditory scene analysis by songbirds: stream segregation of birdsong by European starlings (Sturnus vulgaris). J Comp Psychol 111:3-13.

Izumi A (2002) Auditory stream segregation in Japanese monkeys. Cognition 82:B113-122.

Jaaskelainen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levanen S, Lin FH, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW (2004) Human posterior auditory cortex gates novel sounds to consciousness. Proc Natl Acad Sci U S A 101:6809-6814.

Javel E (1996) Long-term adaptation in cat auditory-nerve fiber responses. J Acoust Soc Am 99:1040-1052.

Javitt DC, Steinschneider M, Schroeder CE, Vaughan HG, Jr., Arezzo JC (1994) Detection of stimulus deviance within primate primary auditory cortex: intracortical mechanisms of mismatch negativity (MMN) generation. Brain Res 667:192-200.

Jesteadt W, Bacon SP, Lehman JR (1982) Forward masking as a function of frequency, masker level, and signal delay. J Acoust Soc Am 71:950-962.

Joris PX, Yin TC (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am 91:215-232.

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541-577.

Kaas J, Axelrod S, Diamond IT (1967) An ablation study of the auditory cortex in the cat using binaural tonal patterns. J Neurophysiol 30:710-724.

Kaltenbach JA, Meleca RJ, Falzarano PR, Myers SF, Simpson TH (1993) Forward masking properties of neurons in the dorsal cochlear nucleus: possible role in the process of echo suppression. Hear Res 67:35-44.

Kanwal JS, Medvedev AV, Micheyl C (2003) Neurodynamics for auditory stream segregation: tracking sounds in the mustached bat's natural environment. Network 14:413-435.

Kessel RG, Kardon RH (1979) Tissues and organs: a text-atlas of scanning electron microscopy. San Fransisco: Freeman.

Kiang N-S (1965) Discharge patterns of single fibres in the cat's auditory nerve. Cambridge: MIT Press.

Kidd G, Jr., Feth LL (1982) Effects of masker duration in pure-tone forward masking. J Acoust Soc Am 72:1384-1386.

King C, McGee T, Rubel EW, Nicol T, Kraus N (1995) Acoustic features and acoustic changes are represented by different central pathways. Hear Res 85:45-52.

Kraus N, McGee T, Carrell T, King C, Littman T, Nicol T (1994) Discrimination of speech-like contrasts in the auditory thalamus and cortex. J Acoust Soc Am 96:2758-2768.

Kvale MN, Schreiner CE (2004) Short-term adaptation of auditory receptive fields to dynamic stimuli. J Neurophysiol 91:604-612.

Langner G (1992) Periodicity Coding in the Auditory-System. Hear Res 60:115-142.

Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 63:442-455.

Lu ZL, Williamson SJ, Kaufman L (1992) Behavioral lifetime of human auditory sensory memory predicted by physiological measures. Science 258:1668-1670.

MacDougall-Shackleton SA, Hulse SH, Gentner TQ, White W (1998) Auditory scene analysis by European starlings (Sturnus vulgaris): perceptual segregation of tone sequences. J Acoust Soc Am 103:3581-3587.

Macken WJ, Tremblay S, Houghton RJ, Nicholls AP, Jones DM (2003) Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. J Exp Psychol Hum Percept Perform 29:43-51.

Malmierca MS, Rees A, Le Beau FE, Bjaalie JG (1995) Laminar organization of frequency-defined local axons within and between the inferior colliculi of the guinea pig. J Comp Neurol 357:124-144.

Malmierca MS, Blackstad TW, Osen KK, Karagulle T, Molowny RL (1993) The central nucleus of the inferior colliculus in rat: a Golgi and computer

reconstruction study of neuronal and laminar structure. J Comp Neurol 333:1-27.

Malone BJ, Semple MN (2001) Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. J Neurophysiol 86:1113-1130.

Mardia KV, Jupp PE (2000) Directional Statistics, 2nd Edition. Chichester: J. Wiley.

Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. Proc Natl Acad Sci U S A 95:5323-5328.

Massaro DW, Cohen MM, Idson WL (1976) Recognition masking of auditory lateralization and pitch judgments. J Acoust Soc Am 59:434-441.

McAdams S, Bregman AS (1979) Hearing musical streams. Comput Music J 3:26-63.

McCabe S, Denham M (1997) A model of auditory streaming. J Acoust Soc Am 101:1611-1621.

Merrill EG, Ainsworth A (1972) Glass-coated platinum-plated tungsten microelectrodes. Med Biol Eng 10:662-672.

Merzenich MM, Roth LL, Andersen RA, Knight PL, Colwell SA (1977) Some basic features of organization of the central auditory nervous system. In: Psychophysics and Physiology of Hearing (Evans EF, Wilson JP, eds). London: Academic Press.

Micheyl C, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 48:139-148.

Micheyl C, Bee MA, Oxenham AJ, Klump GM (2008) Multi-second Adapatation of Neural Responses to Tone Sequences in the Avian Forebrain and its Relationship with the Build-up of Auditory Streaming. In: ARO Poster #895. Phoenix, Arizona, USA.

Micheyl C, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Courtenay Wilson E (2007) The role of auditory cortex in the formation of auditory streams. Hear Res 229:116-131.

Miller GA, Taylor WG (1948) The Perception of Repeated Bursts of Noise. J Acoust Soc Am 20:171-182.

Miller GA, Heise GA (1950) The Trill Threshold. J Acoust Soc Am 22:637-638.

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

Moore BC (1993) Frequency analysis and pitch perception. In: Human Psychophysics (Yost WA, Popper AN, Fay RR, eds). New York: Springer-Verlag.

Moore BC, Glasberg BR, Roberts B (1984) Refining the measurement of psychophysical tuning curves. J Acoust Soc Am 76:1057-1066.

Moore BCJ (2003) An introduction to the psychology of hearing, Fifth Edition. San Diego: Academic Press.

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. J Neurophysiol 95:3756-3769.

Naatanen R, Winkler I (1999) The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 125:826-859.

Naatanen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence" in the auditory cortex. Trends Neurosci 24:283-288.

Nakamoto KT, Zhang J, Kitzes LM (2006) Temporal nonlinearity during recovery from sequential inhibition by neurons in the cat primary auditory cortex. J Neurophysiol 95:1897-1907.

Oesterreich RE, Strominger NL, Neff WD (1971) Neural structures mediating differential sound intensity discrimination in the cat. Brain Res 27:251-270.

Oshurkova E, Scheich H, Brosch M (2008) Click train encoding in primary and non-primary auditory cortex of anesthetized macaque monkeys. Neuroscience 153:1289-1299.

Palombi PS, Backoff PM, Caspary DM (1994) Paired tone facilitation in dorsal cochlear nucleus neurons: a short-term potentiation model testable in vivo. Hear Res 75:175-183.

Perez-Gonzalez D, Malmierca MS, Covey E (2005) Novelty detector neurons in the mammalian auditory midbrain. Eur J Neurosci 22:2879-2885.

Pfeiffer RR (1966) Classification of response patterns of spike discharges for units in the cochlear nucleus: tone-burst stimulation. Exp Brain Res 1:220-235.

Pfingst BE, O'Connor TA (1981) Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. J Neurophysiol 45:16-34.

Phillips DP, Orman SS (1984) Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. J Neurophysiol 51:147-163.

Phillips DP, Sark SA (1991) Separate mechanisms control spike numbers and inter-spike intervals in transient responses of cat auditory cortex neurons. Hear Res 53:17-27.

Phillips DP, Orman SS, Musicant AD, Wilson GF (1985) Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. Hear Res 18:73-86.

Phillips DP, Semple MN, Calford MB, Kitzes LM (1994) Level-dependent representation of stimulus frequency in cat primary auditory cortex. Exp Brain Res 102:210-226.

Phillips DP, Calford MB, Pettigrew JD, Aitkin LM, Semple MN (1982) Directionality of sound pressure transformation at the cat's pinna. Hear Res 8:13-28.

Pickles JO (1988) An introduction to the physiology of hearing, Second Edition. London: Academic Press.

Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I. Evaluation of components. Electroencephalogr Clin Neurophysiol 36:179-190.

Plomp R, Mimpen AM (1979) Speech-reception threshold for sentences as a function of age and noise level. J Acoust Soc Am 66:1333-1342.

Pressnitzer D, Hupe J (2006) Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. In: Equipe Audition (FRE 2929).

Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual organization of sound begins in the auditory periphery. Curr Biol 18:1124-1128.

Qin L, Kitama T, Chimoto S, Sakayori S, Sato Y (2003) Time course of tonal frequency-response-area of primary auditory cortex neurons in alert cats. Neurosci Res 46:145-152.

Read HL, Winer JA, Schreiner CE (2002) Functional architecture of auditory cortex. Curr Opin Neurobiol 12:433-440.

Reale RA, Brugge JF (2000) Directional sensitivity of neurons in the primary auditory (AI) cortex of the cat to successive sounds ordered in time and space. J Neurophysiol 84:435-450.

Rees A, Palmer AR (1989) Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. J Acoust Soc Am 85:1978-1994.

Relkin EM, Pelli DG (1987) Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures. J Acoust Soc Am 82:1679-1691.

Relkin EM, Turner CW (1988) A reexamination of forward masking in the auditory nerve. J Acoust Soc Am 84:584-591.

Relkin EM, Doucet JR (1991) Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons. Hear Res 55:215-222.

Rennaker RL, Carey HL, Anderson SE, Sloan AM, Kilgard MP (2007) Anesthesia suppresses nonsynchronous responses to repetitive broadband stimuli. Neuroscience 145:357-369.

Roberts B, Glasberg BR, Moore BC (2002) Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. J Acoust Soc Am 112:2074-2085.

Rogers WL, Bregman AS (1993) An experimental evaluation of three theories of auditory stream segregation. Percept Psychophys 53:179-189.

Rose JE, Greenwood DD, Goldberg JM, Hind JE (1963) Some discharge characteristics of single neurons in the inferior colliculus of the cat. I. Tonotopical organisation, relation of spike-counts to tone intensity, and firing patterns of single elements. J Neurophysiol 26:294-320.

Russell IJ, Sellick PM (1978) Intracellular studies of hair cells in the mammalian cochlea. J Physiol 284:261-290.

Ryan A, Miller J (1978) Single unit responses in the inferior colliculus of the awake and performing rhesus monkey. Exp Brain Res 32:389-407.

Sachs MB, Abbas PJ (1974) Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. J Acoust Soc Am 56:1835-1847.

Sadagopan S, Wang X (2008) Level invariant representation of sounds by populations of neurons in primary auditory cortex. J Neurosci 28:3415-3426.

Saldana E, Feliciano M, Mugnaini E (1996) Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. J Comp Neurol 371:15-40.

Sally SL, Kelly JB (1988) Organization of auditory cortex in the albino rat: sound frequency. J Neurophysiol 59:1627-1638.

Sanchez-Vives MV, Nowak LG, McCormick DA (2000a) Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. J Neurosci 20:4286-4299.

Sanchez-Vives MV, Nowak LG, McCormick DA (2000b) Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. J Neurosci 20:4267-4285.

Sanes DH, Malone BJ, Semple MN (1998) Role of synaptic inhibition in processing of dynamic binaural level stimuli. J Neurosci 18:794-803.

Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the FMRI adaptation paradigm. Neuron 49:307-318.

Scharlock DP, Neff WD, Strominger NL (1965) Discrimination of Tone Duration after Bilateral Ablation of Cortical Auditory Areas. J Neurophysiol 28:673-681.

Scholl B, Gao X, Wehr M (2008) Level dependence of contextual modulation in auditory cortex. J Neurophysiol 99:1616-1627.

Schreiner C (1981) Poststimulatory effects in the medial geniculate body of guinea pigs. In: Neuronal mechanisms of hearing (Syka J, Aitkin L, eds), pp 191-196. New York: Plenum.

Schreiner CE, Langner G (1988) Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. J Neurophysiol 60:1823-1840.

Schreiner CE, Sutter ML (1992) Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. J Neurophysiol 68:1487-1502.

Schreiner CE, Langner G (1997) Laminar fine structure of frequency organization in auditory midbrain. Nature 388:383-386.

Schreiner CE, Mendelson JR, Sutter ML (1992) Functional topography of cat primary auditory cortex: representation of tone intensity. Exp Brain Res 92:105-122.

Sellick PM, Patuzzi R, Johnstone BM (1982) Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique. J Acoust Soc Am 72:131-141.

Shamma SA, Symmes D (1985) Patterns of inhibition in auditory cortical cells in awake squirrel monkeys. Hear Res 19:1-13.

Shannon RV (1990) Forward masking in patients with cochlear implants. J Acoust Soc Am 88:741-744.

Shore SE (1995) Recovery of forward-masked responses in ventral cochlear nucleus neurons. Hear Res 82:31-43.

Smith RL (1977) Short-term adaptation in single auditory nerve fibers: some poststimulatory effects. J Neurophysiol 40:1098-1111.

Smith RL, Zwislocki JJ (1975) Short-term adaptation and incremental responses of single auditory-nerve fibers. Biol Cybern 17:169-182.

Snyder JS, Large EW (2004) Tempo dependence of middle- and long-latency auditory responses: power and phase modulation of the EEG at multiple time-scales. Clin Neurophysiol 115:1885-1895.

Snyder JS, Alain C (2007) Toward a neurophysiological theory of auditory stream segregation. Psychol Bull 133:780-799.

Strominger NL, Oesterreich RE, Neff WD (1980) Sequential auditory and visual discriminations after temporal lobe ablation in monkeys. Physiol Behav 24:1149-1156.

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

Sutter ML, Schreiner CE (1991) Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. J Neurophysiol 65:1207-1226.

Sutter ML, Schreiner CE (1995) Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. J Neurophysiol 73:190-204.

Symmes D, Chapman LF, Halstead WC (1955) The Fusion of Intermittent White Noise. J Acoust Soc Am 27:470-473.

Tan AYY, Zhang LI, Merzenich MM, Schreiner CE (2004) Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. Journal of Neurophysiology 92:630-643.

Tanaka H, Fujita N, Watanabe Y, Hirabuki N, Takanashi M, Oshiro Y, Nakamura H (2000) Effects of stimulus rate on the auditory cortex using fMRI with 'sparse' temporal sampling. Neuroreport 11:2045-2049.

Theunissen FE, Doupe AJ (1998) Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. J Neurosci 18:3786-3802.

Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci U S A 94:719-723.

Turner CW, Relkin EM, Doucet J (1994) Psychophysical and physiological forward masking studies: probe duration and rise-time effects. J Acoust Soc Am 96:795-800.

Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. Nat Neurosci 6:391-398.

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

van Noorden LPAS (1975) Temporal Coherence in the Perception of Tone Sequences. In. Eindhoven, the Netherlands: University of Technology.

Varela JA, Sen K, Gibson J, Fost J, Abbott LF, Nelson SB (1997) A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. J Neurosci 17:7926-7940.

Viemeister NF, Plack CJ (1993) Time analysis. In: Human Psychophysics (Yost WA, Popper AN, Fay RR, eds). New York: Springer-Verlag.

Vliegen J, Oxenham AJ (1999) Sequential stream segregation in the absence of spectral cues. J Acoust Soc Am 105:339-346.

Vliegen J, Moore BC, Oxenham AJ (1999) The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. J Acoust Soc Am 106:938-945.

Volkov IO, Galazyuk AV (1992) Peculiarities of Inhibition in Cat Auditory-Cortex Neurons Evoked by Tonal Stimuli of Various Durations. Experimental Brain Research 91:115-120.

Wallace MN, Rutkowski RG, Palmer AR (2000) Identification and localisation of auditory areas in guinea pig cortex. Exp Brain Res 132:445-456.

Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol 74:2685-2706.

Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. Nature 435:341-346.

Watkins PV, Barbour DL (2008) Specialized neuronal adaptation for preserving input sensitivity. Nat Neurosci 11:1259-1261.

Weber DL, Green DM (1978) Temporal factors and suppression effects in backward and forward masking. J Acoust Soc Am 64:1392-1399.

Wehr M, Zador AM (2003) Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature 426:442-446.

Wehr M, Zador AM (2005) Synaptic mechanisms of forward suppression in rat auditory cortex. Neuron 47:437-445.

Werner-Reiss U, Porter KK, Underhill AM, Groh JM (2006) Long lasting attenuation by prior sounds in auditory cortex of awake primates. Exp Brain Res 168:272-276.

Westerman LA, Smith RL (1984) Rapid and short-term adaptation in auditory nerve responses. Hear Res 15:249-260.

Widin GP, Viemeister NF (1979) Intensive and temporal effects in pure-tone forward masking. J Acoust Soc Am 66:388-395.

Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical FMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. J Neurophysiol 97:2230-2238.

Winkler I, Naatanen R (1992) Event-related potentials in auditory backward recognition masking: a new way to study the neurophysiological basis of sensory memory in humans. Neurosci Lett 140:239-242.

Winter IM, Robertson D, Yates GK (1990) Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres. Hear Res 45:191-202.

Xie R, Gittelman JX, Pollak GD (2007) Rethinking tuning: in vivo whole-cell recordings of the inferior colliculus in awake bats. J Neurosci 27:9469-9481.

Yin TC, Chan JC, Irvine DR (1986) Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise. J Neurophysiol 55:280-300.

Zhang J, Nakamoto KT, Kitzes LM (2005) Modulation of level response areas and stimulus selectivity of neurons in cat primary auditory cortex. J Neurophysiol 94:2263-2274.

Zhang LI, Tan AY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. Nature 424:201-205.