Sensitivity to Interaural Timing Differences within the Envelopes of Acoustic Waveforms

David Louis Greenberg

Submitted for the Degree of

Doctor of Philosophy

University College London

I, David Louis Greenberg, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Interaural-timing-differences (ITDs) are a cue for sound-source localisation and can be conveyed in the temporal-fine-structure (TFS) of low-frequency tones or in the envelope of high-frequency, amplitude-modulated sounds such as sinusoidally amplitude-modulated (SAM) and transposed-tones. Sensitivity to these cues has been measured in human psychophysical experiments and has revealed that the tranposed-tone elicits just-noticeable-differences (JNDs) in ITDs that are equivalent to those of low-frequency pure-tones when the modulation frequency is below 512-Hz. At modulation frequencies above 512-Hz performance rapidly declines for the transposed-tone while sensitivity to ITDs in pure-tones is robust until around 1200-Hz. Furthermore, transposed-tones elicit JNDs smaller than SAM tones. In the present study, ITD JNDs are assessed psychophysically for pure-tones and transposed-tones using off-midline reference locations. The results demonstrate that frequency, whether the ITD is conveyed in the TFS or the envelope, and location, all have a significant effect on human ITD JNDs and suggest that a difference exists in how ITDs are coded neuronally when conveyed by either high- or low-frequency sounds.

ITD-sensitive neurons located within several brainstem nuclei display a high degree of phase-locking to both the TFS of low-frequency pure-tones and the envelopes of SAM and transposed-tones. Echoing the psychophysical findings, phase-locking to the waveform envelope at low modulation frequencies is equivalent to that of low-frequency pure-tones, while declining at high rates of modulation to a lesser degree for tranposed-tones than SAM tones. In order to assess factors critical to the localisation of high-frequency sounds a series of electrophysiology experiments were conducted. Recordings were made from single neurons within the inferior colliculus of the guinea pig in response to ITDs conveyed by 18 unique envelope shapes to evaluate how the envelope

segments; Pause, Attack, Sustain and Decay each effect ITD JNDs. Amplitude-modulations with envelope shapes comprising relatively long Pause but short Attack durations have been found to elicit the greatest ITD discrimination of high-frequency sounds.

Acknowledgements

First and foremost, I would like to thank my supervisor, David McAlpine, for all of his help and guidance throughout my thesis project. The past years under his supervision have been incredibly enjoyable and he has facilitated a great deal of my personal growth and development. I am grateful for having had the opportunity to work so closely with him for the duration of my PhD.

This thesis would not have been possible without the tremendous help from two individuals: Torsten Marquardt and Mathias Dietz. I am grateful to Torsten for helping me to set up the psychophysical experiments and to Mathias for helping to put together the electrophysiological experiments.

When I first started my PhD I was fortunate enough to work with Andy Forge, Jonathan Gale and Dan Jagger who I would like to thank for their mentoring at different stages over the years and for providing me with a fantastic introduction to microscopy, dissection, histochemistry and molecular biology.

I have been fortunate enough to be involved with a great number of different programmes at UCL that have allowed me to develop in ways that have shaped the person that I am. Through these I have had the pleasure of working and interacting with a great number of remarkable individuals. I would like to thank Priya Singh, Stuart Rosen, Debi Vickers, Cherilee Rutherford, Roland Schaette, Vit Drga, Bridgitte Harley, Sally Dawson, Ruth Taylor, Ghada Al-Malky, Denise Goldman, Jennifer Linden, Ifat Yasin and Ben Robinson. To my viva examiners, Jennifer Bizley and Kevin Munro, a special thank you for your insights and your significant contribution to such a memorable experience.

Many other people provided me with technical and administrative support over the years without whom the process would have been impossible. In particular I

would like to thank Cathy, June, Cheryl, Ricky, Sarah, Kate, Neil, AC, Nick, Robert and Adam,

I would also like to thank my fellow EI students, who have contributed hugely to my experience over the years, Stefano, Nicolas, Eleanor, Jason, John, Lara, Alex, Jacob and Rachael.

To Jessica, I am now and will be forever grateful to you for everything that you have done for me over the past 3 years; you deserve a PhD for that alone.

Finally, I am forever grateful to my parents for their unconditional support over the years and to Anna and Jonathan for simply being who they are. I know that wherever I am and whatever I am doing I will always have your love and advice to guide me and a place to call home.

Contents

List of Abbreviations

µs	Microsecond
ANF	Auditory Nerve Fibre
B/MLD	Binaural/Masking Level Difference
BF	Best Frequency
CD	Characteristic Delay
CF	Characteristic Frequency
CI	Cochlear Implant
dB	Decibel
FRA	Frequency Response Area
Hz	Hertz
IC	Inferior Colliculus
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
ITD	Interaural Time Difference
JND	Just Noticeable Difference
kHz	Kilohertz
LSO	Lateral Superior Olive
MAA	Minimum Audible Angle
MNTB	Medial Nucleus of the Trapezoid Body
ms	Millisecond
MSO	Medial Superior Olive
PSTH	Post Stimulus Tine Histogram
PSW	Pseudo Square Wave
SAM	Sinusoidally Amplitude Modulated
SOC	Superior Olivary Complex
TFS	Temporal Fine Structure
TT	Transposed Tone

1 General Introduction

The ability to locate the source of a sound is a fundamental aspect of hearing, vital at all levels within the animal kingdom for purposes ranging from communication to the survival of both predator and prey. There are multiple cues available that can be used to establish where a sound is coming from in three dimensional space and many species, including humans, make specific use of binaural ('two-eared') cues to do so. For a sound that originates at a location other than the midline there are differences in the time of arrival at each ear, referred to as the interaural time difference (ITD). For an ongoing sound an interaural phase difference (IPD) follows the initial ITD. The head also casts an acoustic shadow on the ear furthest from a sound source thus generating an interaural level difference (ILD). Frequency dependent limitations exist for the ability of the auditory system to encode these sound localisation cues as well as the existence of locations that would prove ambiguous without further cues. Across-frequency processing of auditory features arises from the filtering properties of the outer ear, contributing to the spectral cues relevant to sound localisation in both the horizontal and vertical planes. Furthermore, the difficult task of establishing the distance of a sound source is in part made possible as a change in distance results in a change of both stimulus intensity and frequency spectrum.

Despite the existence of these binaural cues having been recognised for at least a century, the manner in which they are processed by the mammalian brain remains a subject of considerable debate. This thesis presents a dual approach, incorporating both psychophysical and electrophysiological experimental findings, to provide complimentary perspectives on the binaural mechanisms involved in processing localisation cues that are conveyed by both low- and high-frequency sounds. Initially, human sensitivity to IPDs within the envelope of high frequency sounds was compared to sensitivity to IPDs within the

temporal fine structure of low-frequency sounds, extending the known limitations of human sensitivity to reference IPDs greater than zero. Following this, a study of the neural sensitivity to ITDs carried within different envelope shapes will be presented to provide a complimentary insight in to the mechanisms that underlie localisation performance.

1.1 Sound Localisation Cues

The duplex theory (Rayleigh 1907) posits that humans use ITDs predominantly for localising low-frequency sounds while ILDs are used in the localisation of high-frequency sounds. This can be understood based on the physical properties of low and high-frequency sounds. As the frequency of a sound is increased its wavelength decreases resulting in a greater proportion of the sound wave energy being reflected if it is directed toward a solid surface such as a human head. This reflection of acoustic energy results in an acoustic shadow at the ear furthest from the source. The attenuation of the signal at the far ear relative to the near ear results in the ILD. ILDs of low fequency sounds are relatively minor regardless of sound-source azimuth due to the manner in which a long wavelength, relative to the size of the head, is easily diffracted. ITDs are present at all frequencies but at high frequencies (when the waveform period is short compared to the maximum ITD that can be experienced, as dictated by head width) the accompanying IPD, in isolation, may be an ambiguous cue for sound source localisation due to multiple locations producing the same IPD.

In order to assess the physical properties of ITDs, Woodworth (1938) used a solid sphere to model the shape of the head and measured ITDs as a function of azimuth. The model had a distance between the two ears of 22-23 cm, approximating that of an adult human. The measurements found a maximum ITD of approximately 660 μs when the sound source was at 90° azimuth to one

ear. Following this, Fedderson et al. (1957) measured ITDs using an electronic counter that was switched on and off by signals from microphones placed at the ear canal entrance. Measurements were made between 0 and 180° azimuth at 10° intervals with the resulting values matching closely those recorded by Woodworth (1938) (Figure 1.1). Kuhn (1977) measured ITDs and ILDs using a mannequin in an anechoic chamber with microphones at the head surface. A loudspeaker was placed at a distance of approximately 3 meters, level with the mannequin's ear canals. Measurements were made in 20-Hz increments between 260 and 500 Hz and in 100-Hz increments up to about 7000 Hz for angles of incidence between 0 and 90° azimuth. The results showed that ITD is frequency independent below approximately 500 Hz and above 3000 Hz with a minimum occurring between 1400 Hz and 1600 Hz for angles of incidence between 15 and 60° azimuth. Feddersen et al. (1957) also measured ILDs across frequency and azimuth (Figure 1.2) using a probe microphone within the subject's ear canal and a sound source that was moved around the subject while a level-recorder provided a continuous measurement. The ILD was calculated by subtracting the data from symmetrical locations about the subject. Where the peak ILD occurs in terms of azimuth is frequency dependent as is the peak value itself. At 1000 Hz, peak values are around 10 dB, but can exceed 20 dB at frequencies above 5000 Hz.

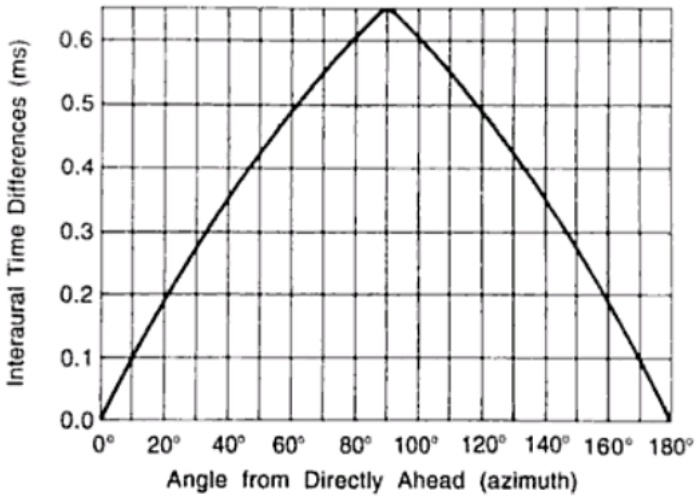


Figure 1.1 ITDs as a function of azimuth as recorded by Feddersen et al. (1957). For the average human head this creates ITDs from 0 to 690μs for pure tone signals. The greatest ITD occurs at 90 degrees azimuth.

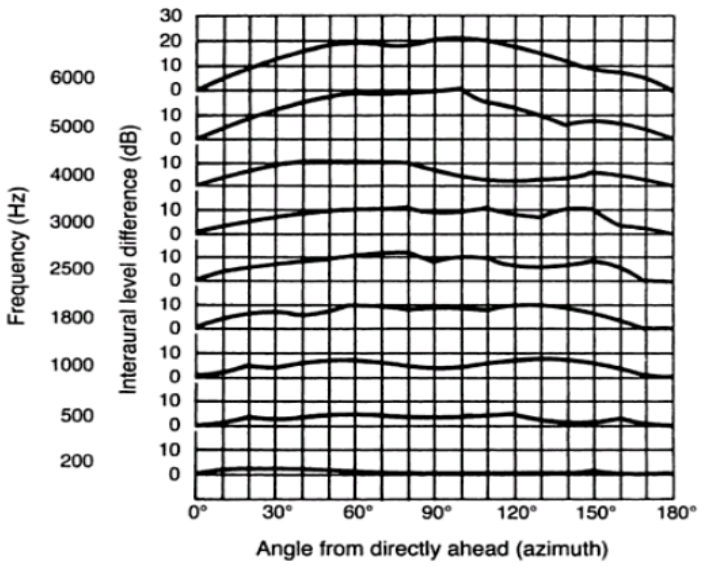


Figure 1.2 ILDs across frequency and azimuth as recorded by Feddersen et al. (1957). At low frequencies ILDs are minor due to the long wavelength relative to head size. For high frequencies ILDs increase in value towards 90 degrees azimuth.

In addition to ITDs and ILDs, spectral cues that result from the natural resonance of the ear canal and pinna provide location dependent gain for some frequencies over others (Gardner and Gardner 1973). These spectral cues allow for the disambiguation of sound source locations that have equal ITDs and ILDs such as those located on the median plane or originating on the circumference of a circular conical slice extending from either ear, termed the cone of confusion. Spectral cues can arise due to changes in the frequency spectrum of a sound at the eardrum due to its angle of origin and were studied in humans using probe microphone measurements by Carlile and Pralong (1994). Results showed that for anterior sound locations a gain of 15 dB for frequencies of 3000 Hz to 5000 Hz was possible and the same frequency range was attenuated by 10 dB to 15 dB when the sound originated at posterior locations. Spectral cues can also account for the ability to detect the elevation of a sound source. Butler and Belendiuk (1977) performed a series of experiments that recorded frequency response curves via microphones in the ear canals of eight subjects in response to broadband noise presented at elevations of ±30°, ±15°, and 0° in the median sagittal plane. The spectra of these recorded sounds revealed a notch in the frequency response curves that systematically lowered in frequency as the sound source was moved from a positive to a negative elevation offering an insight in to the nature of the cues important for establishing the elevation of a sound source.

1.2 Sensitivity to Sound Localisation Cues

Human listeners display an exquisite sensitivity to the source location of pure tones in the free-field as demonstrated by Stevens and Newman (1936). Subjects were presented with tones between 60 Hz and 10000 Hz at a sensation level of 30 dB for the lowest and highest test frequencies, and at 50 to 60 dB in the 400 Hz to 4000 Hz range. The sound source was one of 13 locations between 0 and

180° azimuth and the subject's task was to name the position of the sound source while the error in judgement was recorded. The average error of the two subjects tested (Stevens and Newman) was plotted as a function of tone frequency (Figure 1.3). Localisation errors were 10-14° azimuth for frequencies below 1000 Hz increasing to a peak at 20° azimuth for a 3000-Hz tone. Above 4000 Hz localisation accuracy improved, with a 10000-Hz tone producing a 13° azimuth localisation error.

Following Stevens and Newman's (1936) measurements of error in auditory localisation, Mills (1958) measured the relative precision of auditory localisation. This was achieved by assessing the smallest detectable difference between the azimuths of two identical sounds in the free-field, referred to as the minimum audible angle (MAA). The stimuli were pure tones between 250 Hz and 10000 Hz presented at a sensation level of 50 dB. The smallest difference in azimuth at which a subject could discriminate a change to the right from a change to left was recorded for sound source locations between 0 and 90° azimuth (Figure 1.4). The MAA was smallest for a sound source at 0° azimuth and for frequncies between 250 Hz and 1000 Hz at approximately 1°. Above 1000 Hz the MAA increases to a maximum of about 3° around 1500 Hz with 2° MAA between 3000 and 6000 Hz. A sound source at 75° azimuth would produce a MAA between 6° and 10° for frequencies up to 1000 Hz with a rapid increase to above 30° for frequencies between 1000 Hz and 2000 Hz. These studies demonstrate that human sound localisation ability on the horizontal plane is greatest when the target sound is either below 1500 Hz or greater than 4000 Hz and is located close to the midline of the auditory field.

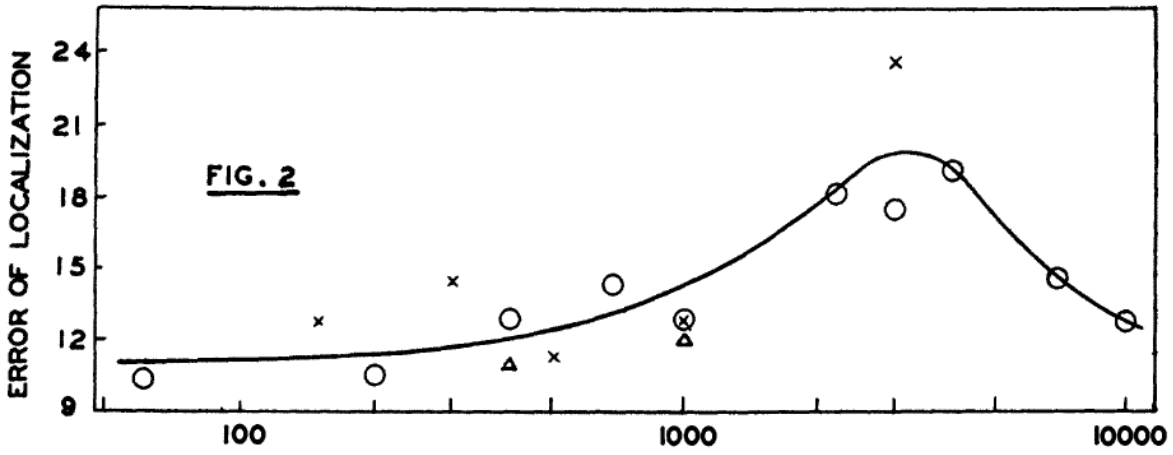


Figure 1.3 The average error of localisation plotted as a function of tone frequency from Stevens and Newman (1936). Localisation errors are as low as 10-14° azimuth for frequencies below 1000 Hz increasing to a peak at 20° azimuth for a 3000 Hz tone.

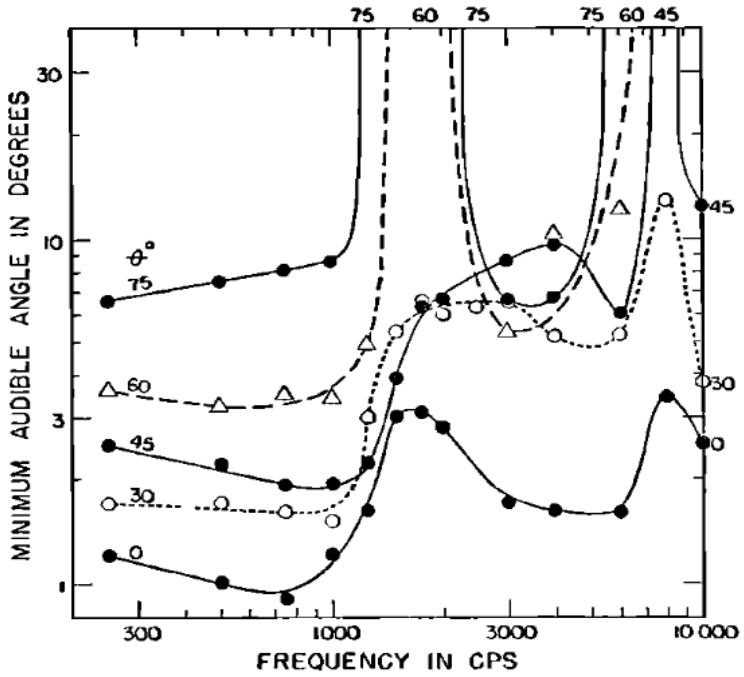


Figure 1.4 MAA for pure tones of 250-10000Hz for reference azimuths 0-90 degrees as recorded by Mills (1958). Between 250 and 1000Hz, MAA is smallest. Above 1000Hz, the MAA greatly increases as a function of reference azimuth.

Using headphones it is possible to isolate specific acoustic cues used for sound localisation and present them to a listener in the form of a lateralisation task. This enables the assessment of the sensitivity and weighting of the binaural system for a specific cue by recording the smallest interaural difference that elicits a noticeable change in a sounds lateral position. Zwislocki and Feldman (1956) measured ITD discrimination thresholds as a function of both pure tone frequency and sensation level. The results showed that human listeners are able to discriminate ITDs as small as 13-30 μs at frequencies below 1200 Hz (Figure 1.5). The study also demonstrated that increasing the sensation level of a signal from 10 dB to 110 dB sensation level affects IPD thresholds in a frequency specific manner (Figure 1.6). Greatest sensitivity to IPDs for pure tones below 1250 Hz appears to occur when the signal is presented at a sensation level between 50 and 90 dB. Yost (1974) assessed discrimination of IPDs for pure tones from 250 Hz to 4000 Hz and found for frequencies above 2000 Hz, changes in IPD are not perceived by a listener as a shift in the lateralisation of a sound image. In agreement with earlier studies, for frequencies below 1000 Hz, IPDs of around 2° to 5° are discriminable (Klumpp and Eady 1956; Zwislocki and Feldman 1956; Yost 1974). A common observation is that detection of interaural differences is greatest when the reference and target signals are on or close to the midline and a shift towards either ear results in a decrease in sensitivity (Yost 1974; Hafter and Maio 1975; Domnitz and Colburn 1977).

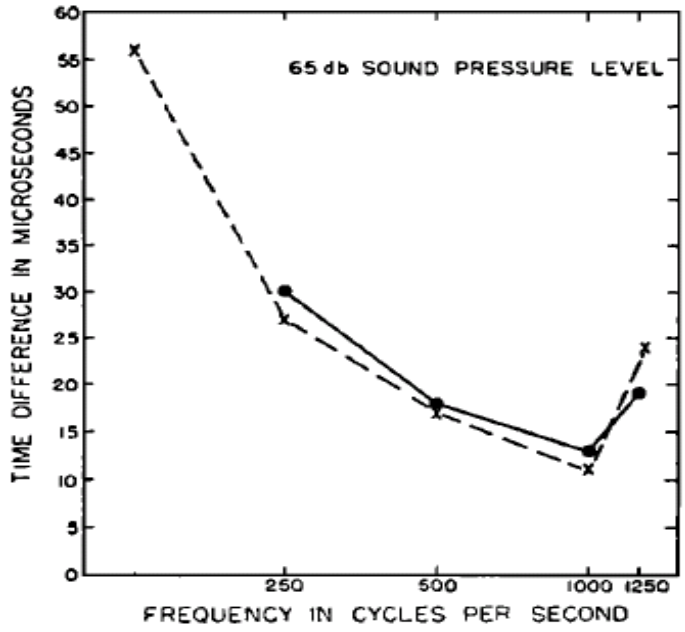


Figure 1.5 Threshold ITD as a function of pure tone frequency from Zwislocki and Feldman (1956). Human listeners are able to discriminate ITDs as small as 13-30 μs at frequencies below 1000 Hz. Solid line = Zwislocki and Feldman (1956) Dashed line = data from Klumpp from a personal communication to Zwislocki and Feldman.

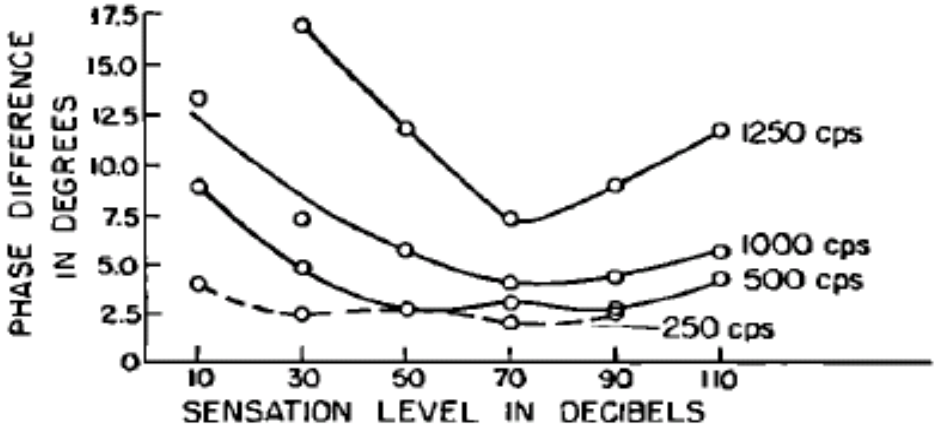


Figure 1.6 Threshold IPD as a function of pure tone frequency and sensation level from Zwislocki and Feldman (1956). With increasing frequency, sensation level has a varying impact on IPD detection thresholds.

ILDs as a cue for sound localisation have also been extensively studied. Mills (1958; 1960) conducted experiments that investigated thresholds ILDs for sound frequencies of 250 Hz to 10000 Hz presented via headphones at a sensation level of 50 dB. The ILD just-noticeable-difference (JND) was found to be 1 dB at 1000 Hz decreasing to around 0.5 dB for frequencies between 2000 Hz and 5000 Hz. Although there is a reduction in sensitivity to ILDs around 1000 Hz (Grantham 1984), thresholds remain approximately constant across a frequency range of 200 Hz to 5000 Hz (Yost and Dye 1988). Time-intensity trading tasks further enable the assessment of the use of ILDs in sound localisation. A stimulus with a fixed ITD is presented and the ILD required to shift the perceived sound source location to one that is defined by the task, often the midline, is measured. The trading ratio for pure tones has been found to be 192.3 μs/dB for 200 Hz, 81.2 μs/dB for 400 Hz, 66.2 μs/dB for 600 Hz, and 62.5 μs/dB for 1000 Hz (Young 1976) although amongst studies there is a large variability in the reported trading ratios. This has been attributed to the type of stimulus used and its presentation level, whereby high-intensity low-frequency tones produce the smallest ratios and low intensity clicks produce the largest. A 500 Hz tone at 70 dB, for example, resulted in values from 13 to 25 μs/dB (Hafter and Carrier 1972) while low-pass clicks at 15 dB sensation level had values of around 120 μs/dB dropping to 20 μs/dB when the stimulus was raised 40 dB (Deatherage and Hirsh 1959). Click trains, low-pass filtered at 1000 Hz produce a time-intensity trade of approximately 25 μs/dB which increases to 60 μs/dB when the click train is high pass filtered at 4000 Hz (Harris 1960). The steep slope of the time-intensity trade for the high-frequency stimulus confirms that ILDs are an important cue for sound localisation of high-frequency stimuli.

Despite the inability of human listeners to discriminate ITDs conveyed in high-frequency pure tones it is possible for listeners to discriminate ITDs in complex waveforms, even when the signal is comprised only of frequencies above 1500 Hz (Leakey, Sayers et al. 1958). Rather than the ITDs being conveyed in the fine

structure of a signal, they can instead be carried in the envelope of high-frequency complex waveforms. Although human sensitivity to changes in ITDs of transient stimuli that have energy only above 1500 Hz, is poorer than for transients with lower–frequency energy (Yost, Wightman et al. 1971), listeners are able to detect interaural delays in the envelope of a 3900-Hz carrier tone modulated 100% at 300 Hz just as well as they could detect interaural delays of a 300-Hz tone Henning (1974). This particular study also reported that with modulation frequencies around 600 Hz, ITD discrimination declined rapidly compared to the 600-Hz pure tone, a factor that was further explored in the present experiments. Sensitivity to ITDs in the envelope of high frequencies stimuli compared to ITDs in low-frequency temporal fine structure is generally poor with the suggestion being that this is in part due to the limitations of the neural pathway involved (Joris 2003). This is explored further in Section 1.6 - *Neural representation of ITDs* and Section 1.7 – *Processing amplitude modulated tones*.

1.3 The Binaural Advantage

Binaural hearing enables the mechanism that allows a target-signal and a noise-interferer to be spatially separated, improving a listener's speech intelligibility in the presence of background noise. This effect is termed 'spatial release from masking' (Bronkhorst and Plomp 1992). Spatial release from masking arises due to the difference in the ITD of competing sound sources (binaural unmasking) (Bronkhorst and Plomp 1988) combined with the improved signal-to-noise ratio at one ear due to the head-shadow effect, known as the better-ear effect (Arsenault and Punch 1999). These mechanisms allow a person with normal hearing to follow speech in the presence of background noise remarkably well, an ability first described as the 'cocktail party effect' (Cherry 1953). The advantage of binaural over monaural hearing has been measured at 30% in a

monosyllabic-word discrimination task using increasing numbers of extraneous speakers (Pollack and Pickett 1958). Although this binaural advantage has been well documented (Arsenault and Punch 1999; Hawley, Litovsky et al. 1999; Hawley, Litovsky et al. 2004) the precise neural mechanisms involved remains in question.

It is possible to improve speech understanding in the presence of masking noise by presenting the two signals 180 degrees out of phase as demonstrated by Licklider (1948). The study measured the intelligibility of speech-in-noise using a set of antiphasic (speech in phase and noise out of phase or vice versa) and homophasic (both the speech and noise either in or out of phase) conditions. Intelligibility was higher for the antiphasic compared to the homophasic conditions and for binaural antiphasic conditions compared to monaural conditions. Hirsh (1948) additionally measured the threshold of tones between 100 Hz and 5000 Hz in the presence of a 59.1 dB, 7000-Hz band of noise utilising a similar set of antiphasic and homophasic parameters. The binaural threshold was again lower for the antiphasic conditions compared to homophasic conditions by as much as 12–15 dB. These findings are explained by the fact that a tone is heard more easily when it is lateralised in the middle of the head while the noise is located out at the ears or vice versa than when both the tone and noise are localised in the same place, made possible by manipulating the signal IPD. The difference in binaural thresholds for antiphasic and homophasic conditions decreases as frequency increases above 200 Hz due to the higher frequency tones becoming less localisable on the basis of phase (Hirsh 1948).

That certain binaural stimulus configurations can lead to improvements in the detection of a signal in the presence of noise is a concept termed the ‘masking level difference’ (MLD) by Webster (1951) and commonly the ‘binaural masking level difference’ (BMLD). The amount of noise required to mask a signal often depends upon the interaural phase relations between the signal and

masker such that more masking noise is required if the signal is 180 degrees out of phase at the two ears (N_0S_π) than if it is in phase (N_0S_0) (Jeffress, Blodgett et al. 1952; Jeffress, Blodgett et al. 1956). The advantage is greatest for a signal near 250 Hz and falls off progressively toward 1400 Hz while for a signal that is in phase, for frequencies between 250 and 6000 Hz, the level of masking remains relatively constant (Webster 1951).

1.4 Models of Binaural Processing

The place theory of sound localisation (Jeffress 1948) is recognised as one of the earliest descriptions of a mechanism used to account for the remarkable sensitivity to ITDs shown by humans as well as mammals, birds and reptiles. The model suggests that in the presence of a binaural input, neural signals (action potentials) generated from each ear travel along delay lines of varying length towards central coincidence-detecting units (Figure 1.7). Each coincidence detector responds maximally when the externally applied ITD offsets the difference in propagation time of two excitatory components along the delay lines from each ear. In this model the internal delay to each coincidence detector must be known by the auditory system such that the ITD can be extracted from the place of maximum response. The Jeffress model formed the basis of many further psychophysical models of ITD sensitivity that utilise a conceptually similar mechanism known as cross correlation due to its ability to successfully explain data from a number of binaural psychoacoustic studies (Hirsh 1948; Blodgett, Wilbanks et al. 1956; Sayers 1964).

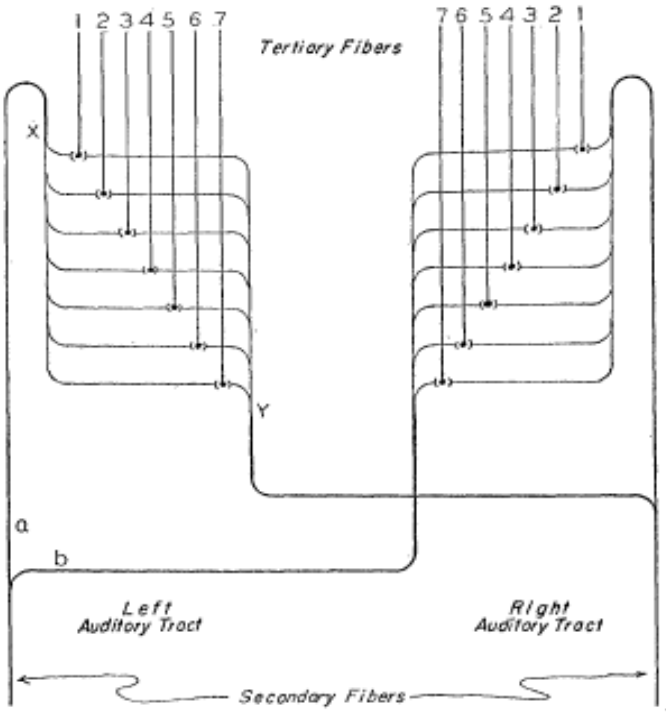


Figure 1.7 The Jeffress model of ITD coding (1948). Jeffress proposed that an anatomical difference in path-lengths could be used by the auditory nervous system to encode ITDs. From Jeffress (1948)

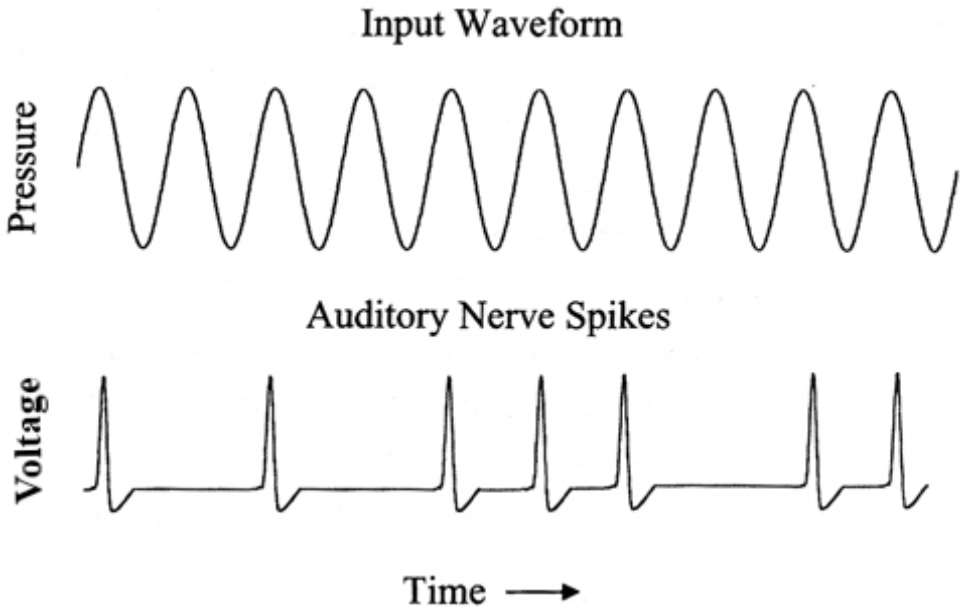


Figure 1.8 ITD sensitive neurons respond cyclically to pure tones with maximum responses occurring at period multiples of the stimulus waveform.

Amongst the first to develop a model based on cross correlation were Sayers and Cherry (1957). Their model provided emphasis for the contributions of small internal delays as well as a mechanism to incorporate ILDs to account for the existing psychophysical data. A further extension to the Jeffress model by Colburn (1973) considers that information utilised by central neural structures must be found within the auditory nerve given that this is the only route for a signal to take following the process of peripheral transduction from sound-wave to nerve impulse. The auditory nerve activity model (Colburn 1973) addresses this fact and considers the possibility that some processes of interaural discrimination may be significantly limited by internal noise arising at the periphery due to the pseudo-random nature of signal transduction. Durlach and Colburn (1978) proposed that, up to that point, the majority of existing models of binaural processing could be described by a generic model of binaural interaction that included peripheral bandpass-filtering, signal rectification, stochastic neural representation of the signal, a correlation or coincidence mechanism, consideration of peripheral ILDs and a central decision making process.

Later models such as that of Lindemann (1986) extended the Jeffress model by incorporating both an inhibition mechanism for the outputs of adjacent coincident detectors and a monaural mechanism that becomes active in the presence of small signal intensities. These features together produce a sharpening of the peaks of the coincidence detector outputs and a shift in the peaks location along the internal delay axis for changes in ILD thus providing a possible mechanism for the time-intensity trading seen in the psychophysical data (Deatherage and Hirsh 1959; Harris 1960; Hafter and Carrier 1972; Young 1976).

In addition to the models describing mechanisms for processing sound localisation cues is the equalisation-cancellation (EC) model of Durlach (1963).

The EC model was developed to account for the finding that detection of a target signal in masking noise is dependent on the interaural relationship between the target and the masker (Jeffress, Blodgett et al. 1952), including that a 180° phase shift in the target signal improves detection thresholds by up to 15 dB (Hirsh 1948) (the BMLD as discussed in Section 1.3 - *The Binaural Advantage)*. The model proposes that the auditory system attempts to transform the total signal in each ear via a series of interaural level adjustments, internal time delays and internal phase shifts ('E' process) followed by a subtraction of the signal in one ear from that in the other ('C' process). When the target interaural signals contain some difference that the masker does not, the remaining signal components after the C process are utilised for detection.

1.5 Physiology of Auditory Perception

As a general introduction to the neurophysiology component of this thesis it is necessary to first review the process of auditory transduction: When a sound wave impinges on the tympanic membrane the airborne pressure variations are converted into mechanical vibrations across the ossicles. The stapes is the last of the three bones in the ossicular chain and is in direct contact with the oval window of the cochlea – the hearing organ of the inner ear. The stapes transmits the vibrations in response to the sound wave via the oval window, reproducing the pressure variations within the perilymphatic fluid of the scala vestibuli, one of the three, coiled chambers of the cochlea. The movement of the perilymphatic fluid acts upon the basilar membrane, producing a travelling wave. The structure of the basilar membrane is such that a maximum deflection occurs where the resonance of the membrane matches the frequency of the stimulating tone. The hearing organ itself, the organ of Corti, lies along the length of the basilar membrane and has protruding from its surface the inner and outer hair cells. As the basilar membrane moves, a shearing motion is created between the

stereocilia bundles on the apical surface of the hair cells and the tectorial membrane. This creates a tension on the tip-links connecting individual stereocilia resulting in the opening of cation channels allowing potassium (K+), present in the surrounding endolymph, a greater probability of entering the hair cell, causing it to depolarise. At the same time calcium (Ca2+) also enters the cell, resulting in hyperpolarisation as it binds to the cation channel. In turn this mediates the release of neurotransmitter at the basal end of the inner hair cells triggering action potentials in afferent auditory nerve fibres (ANFs). The outer hair cells enable cochlea amplification in the presence of low intensity sounds by converting electrical signals back to mechanical, and generating additional movement of the basilar membrane.

The ANF action potentials are *phase-locked* to the individual cycles of low-frequency sounds (Rose, Brugge et al. 1967) thus providing the binaural-temporal information that converges onto ITD sensitive neurons in the auditory brainstem. ITD-sensitive neurons respond cyclically to pure tones over a range of ITDs with maximum responses occurring at period multiples of the stimulus waveform (Figure 1.8). Early studies exploring ANF responses showed phase-locking to be constant up to approximately 2000 Hz in the guinea pig (Tasaki 1954), cat (Rupert, Moushegian et al. 1963) and squirrel monkey (Rose, Brugge et al. 1967) after which it declines until approximately 6000 Hz (Johnson 1980). It is interesting to note that the barn owl (*Tyto alba*), a specialist in exploiting auditory spatial cues for locating prey, is something of an exception, having ANFs that are able to phase-lock to frequencies up to 9000 Hz (Sullivan and Konishi 1984).

1.6 Neural representation of ITDs

ITDs are temporally encoded by the generation of phase-locked action potentials originating at the synapses of inner hair cells in response to the fine structure and envelope of acoustic stimuli. The action potentials propagate along the tonotopically organised auditory nerve maintaining the temporal information, and the frequency specific location along the organ of Corti from which they arise, and provide input to the neural mechanisms responsible for sound localisation (Rose, Greenwood et al. 1963; Guinan, Norris et al. 1972; Tsuchitani 1977; Stiebler and Ehret 1985; Oxenham, Bernstein et al. 2004). Prior to entering the cortical regions of the brain, the propagating signal traverses a network of brainstem nuclei beginning where the auditory nerve synapses with neurons in the ipsilateral cochlear nucleus (Ryugo 1992) which in turn projects to nuclei in the ipsilateral and contralateral superior olivary complex (SOC). The SOC has two major nuclei, the medial superior olive (MSO), considered the nucleus that processes ITDs, and the lateral superior olive (LSO), which processes ILDs (Tsuchitani and Boudreau 1967; Tollin 2003; Grothe, Pecka et al. 2010). Alongside the binaural excitatory input from both cochlear nuclei (Goldberg and Brown 1968), the MSO neurons also receive inhibitory inputs from the medial nucleus of the trapezoid body (MNTB) (Cant and Hyson 1992; Grothe and Sanes 1993; Smith 1995). The MNTB has axonal projections that innervate both the LSO and MSO (Smith, Joris et al. 1998) where the variety of nerve ending morphologies suggest roles in both excitatory and inhibitory mechanisms (Clark 1969). While the MSO and LSO both process binaural cues for sound localisation, the manner in which they respond is different. MSO neurons are binaurally excited (EE) whilst LSO neurons are excited by sound in the ipsilateral ear but inhibited by sound in the contralateral ear (EI) (Goldberg and Brown 1968; Yin and Chan 1990; Batra, Kuwada et al. 1997).

The MSO and LSO send excitatory projections to the inferior colliculus (IC) (Loftus, Bishop et al. 2004), the major auditory nucleus in the midbrain. The IC is tonotopically organised, comprising neurons that fire maximally in the presence of a particular interaural delay (Roth, Aitkin et al. 1978; Stiebler and Ehret 1985; Skottun, Shackleton et al. 2001; Grothe, Pecka et al. 2010). Both the unique firing patterns of binaurally stimulated IC neurons and the degree to which ascending auditory pathways project to this region of the midbrain qualify the IC as an important target in auditory research.

ITDs in the temporal fine structure of low-frequency sounds

Rose et al. (1963) conducted pioneering experiments exploring the discharge characteristics of single neurons within the IC in response to pure tones. The study demonstrated both the frequency tuning of individual neurons as well as the tonotopic organisation of the external and central nuclei of the IC. Hind et al. (1963) conducted one of the primary studies on the neural representation of auditory space, observing discharge patterns of single neurons in response to binaural stimulation. The study documented that the discharge pattern of an IC neuron can vary substantially depending on whether the binaural stimuli are in or out of phase. Rose et al (1966) explored in further detail the discharge patterns of ITD sensitive neurons found in the IC of the cat. The paper describes how single neurons can have a peak discharge at a specific ITD regardless of frequency, termed the characteristic delay (CD) (Figure 1.9). The discharge curves plotted firing rate as a function of ITD revealing each neurons CD. In addition to this, by recording the firing rate of a neuron for acoustic stimuli that were alternately binaural and monaural it was possible to observe how firing rate could be decreased in the presence of the binaural stimuli. In this way the study was able to demonstrate the possible existence of an excitatory-inhibitory mechanism underlying sound localisation abilities.

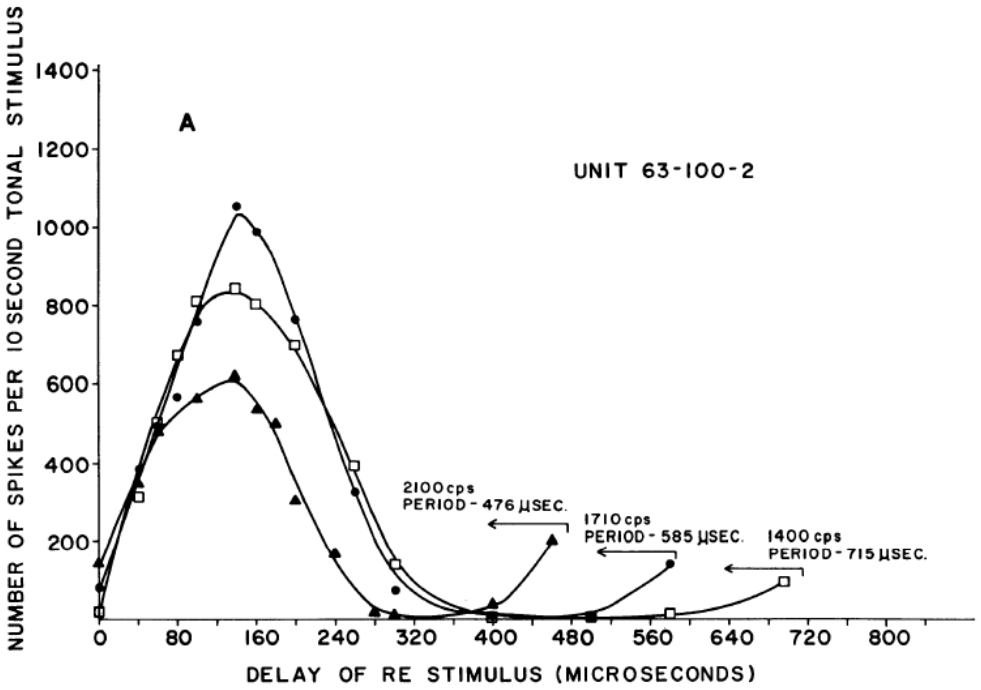


Figure 1.9 Periodic discharge curves to binaural stimulation at three frequencies presented with an ITD. The period of each curve equals that of the stimulating frequency but a maximum response for each occurs at the same delay of 140 µsec, referred to as the characteristic delay. From Rose et al. (1966)

ITD sensitive neurons that respond to the fine structure of low-frequency signals have since been recorded extensively in animals including the rabbit (Kuwada, Stanford et al. 1987), cat (Yin and Kuwada 1983) and gerbil (Spitzer and Semple 1998) and yet the majority of evidence in support of the coincidence detector model is based on recordings made in the barn owl (Carr and Konishi 1990) with only limited evidence for the same neural arrangement in mammals (Smith, Joris et al. 1993; Beckius, Batra et al. 1999). Electrophysiological experiments in the mammalian auditory pathway have demonstrated that the distribution of ITDs to which the majority of delay sensitive neurons are tuned, lies in the range -300 µs to +300 µs (for rabbits, cats and gerbils) (Yin and Kuwada 1983; Kuwada, Stanford et al. 1987; Spitzer and Semple 1998) but there are also neurons tuned

to far larger ITDs (Yin, Chan et al. 1986; McAlpine, Jiang et al. 2001; Hancock and Delgutte 2004; Pecka, Brand et al. 2008). Since some neurons fire maximally for ITDs that are outside of the physically producible range for most small mammals it follows that the ITD to which a given neuron is maximally sensitive is independent of the head width.

Recordings from ITD sensitive neurons reveal that the peak firing rates are not limited to the physiological range but that the steepest part of rate-ITD-function does fall within the physiological range (Figure 1.10) (McAlpine, Jiang et al. 2001; Hancock and Delgutte 2004). A change in ITD within the physiological range would therefore cause the greatest change in firing rate, even though the firing rate itself may not be the neurons maximum. We can hypothesise that this response pattern contributes to the information required to code auditory space and is used by the higher auditory centres for sound localisation and could account for the enhanced sound localisation ability at the midline.

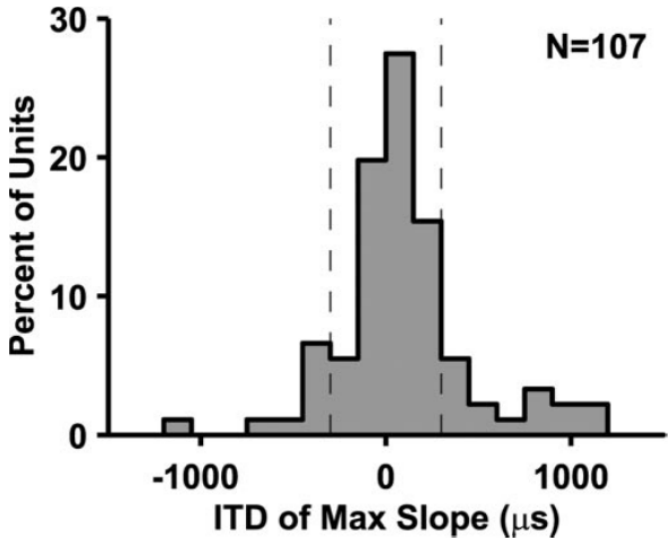


Figure 1.10. Recordings from ITD sensitive neurons reveal that the peak firing rates are not limited to the physiological range but the peak of the distribution that describes the relationship between the steepest part of rate-ITD-function does fall within the physiological range. From Hancock et al. (2004).

ITDs in the envelope of high-frequency sounds

Sensitivity to ITDs in the envelope of high-frequency sounds is present in the firing rates of neurons in the IC (Yin, Kuwada et al. 1984). Utilising interaural delays in the envelope of a signal that was amplitude modulated (AM) at 250 Hz it was demonstrated that IC neurons, with best frequencies greater than 2500 Hz, are able to phase-lock to the signal envelope thus providing sensitivity to IPDs of the modulated waveform. Two examples of AM tones are the sinusoidally amplitude modulated (SAM) tone and the transposed-tone, both are discussed further in Section 1.7 – *processing amplitude modulated tones*, but it is important to note here that assessing ITD sensitivity using SAM tones and transposed-tones has demonstrated that ITD thresholds are smaller for transposed-tones than for SAM tones at all modulation frequencies (Griffin, Bernstein et al. 2005). The ability of neurons located elsewhere in the auditory pathway to phase-lock to the envelope of high-frequency sounds has also been recorded, in the MSO (Yin and Chan 1990), LSO (Joris and Yin 1995) and ANFs (Javel 1980; Palmer 1982; Joris and Yin 1992). Despite the bias of the MSO towards low frequencies and the LSO towards high frequencies (Guinan, Norris et al. 1972) and their respective roles in the processing of ITDs and ILDs, it has been proposed that these roles are not as constrained as was once thought (Joris and Yin 1995).

Dreyer and Delgutte (2006) conducted physiological experiments observing how ANFs respond to SAM tones, transposed-tones and pure tones. The frequency of the pure tone stimuli was set according to the nerve fibre characteristic frequency, as were the carrier frequencies of the SAM and transposed-tones. A synchronisation index was formed from period histograms in order to quantify the strength of phase-locking to the cycles of the stimulus waveform. Phase-locking is better for transposed-tones than SAM tones at all modulation frequencies and levels (Figure 1.11).

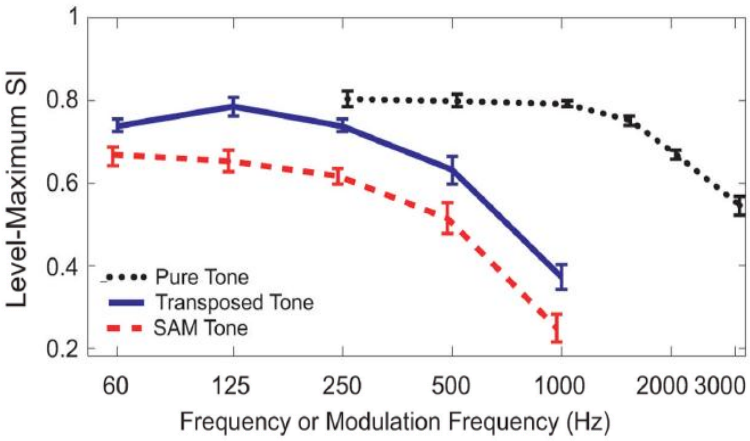


Figure 1.11 A comparison of auditory nerve fibre phase locking capabilities in response to transposed tones, SAM tones and pure tones. Phase locking is better for transposed tones than SAM tones at all modulation frequencies and levels. From Dreyer and Delgutte (2006).

Figure 1.12 presents further data from Dreyer et al. (2006) for a low and high characteristic frequency nerve fibre (344 Hz and 7000 Hz respectively). The recordings reveal similar precision of phase-locking for pure tones and transposed-tones when stimulus levels are close to threshold. Phase-locking to transposed-tones and SAM tones worsens more rapidly with increasing stimuli level than does phase-locking to pure tones but phase-locking for transposed-tones is again better than for SAM tones.

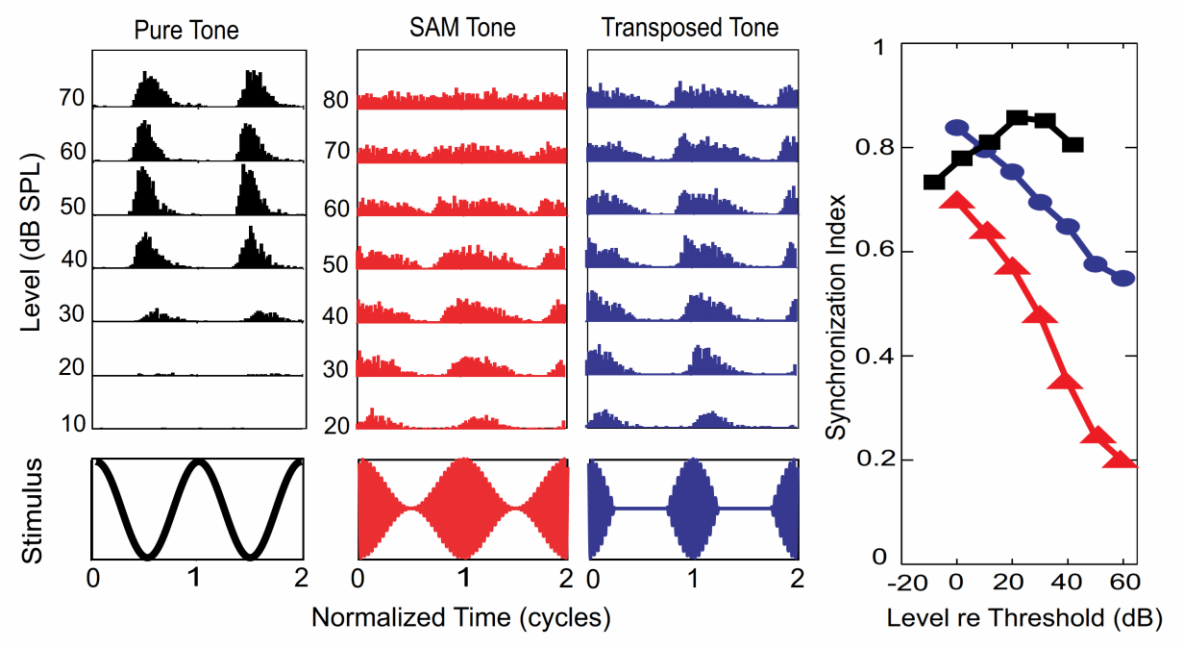


Figure 1.12. From Dreyer et al. (2006). A comparison of auditory nerve fibre phase locking for changing level in pure tones, SAM tones and transposed tones. Synchronisation decreases rapidly at higher levels for SAM tones and transposed tones. Black squares = pure tone, Blue circles = transposed tone, red triangle = SAM tone.

1.7 Processing of Amplitude Modulated Tones

Peripheral processing of amplitude modulated tones

A simple model of peripheral processing of acoustic signals includes three stages; band-pass filtering, half-wave rectification, and low-pass filtering (Figure 1.13) (Bernstein 2001). Band-pass filtering represents the mechanical properties of the inner ear, whereby hair cells respond maximally over a restricted range of sound frequencies. This frequency-specific response of the inner hair cells is itself driven by the frequency tuning of the basilar membrane. The basal end of the basilar membrane is narrow and rigid and vibrates maximally in the presence

of high-frequency sounds. The apical end of the basilar membrane is relatively wide and less rigid and vibrates maximally in the presence of low-frequency sounds. Half-wave rectification mirrors the probability of a hair cell firing being linked to the hair cells response preference being to one direction of stereocilia displacement. The low-pass filtering represents the fact that at 1000 to 2000 Hz the ability of auditory nerve fibres to phase-lock to the waveform temporal fine-structure begins to deteriorate (Palmer and Russell 1986).

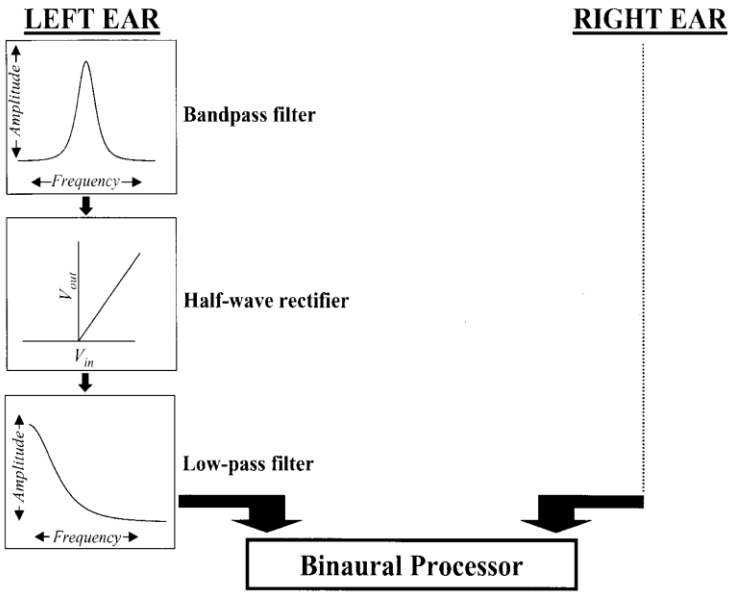


Figure 1.13 A simple model of monaural processing that includes three stages; band-pass filtering, half-wave rectification, and low-pass filtering. From (Bernstein, 2001)

The output from this peripheral processing model for both a pure tone and a complex tone is represented in Figure 1.13. Distinct portions of the processed pure tone waveform have an amplitude level of zero. The envelope that is extracted from the complex waveform has no distinct portions with an amplitude level of zero. This is particularly relevant as this thesis aims to explore the effect of the inter-cycle Pause time on ITD sensitivity of binaurally sensitive neurons.

Considering the outputs from the peripheral processing model (Figure 1.14) it is therefore of interest to construct a high-frequency complex waveform for which the peripheral processing output is the same as for a low-frequency tone of the same modulation frequency. The result would be that the high-frequency channels of the binaural processor could use the envelope-based information that mirrors the information normally available only in low-frequency channels (i.e. temporal information conveyed in the low-frequency temporal fine structure) and enables a useful experimental comparison. Van de Par et al. (1997) described a technique for achieving such a waveform, whereby a half-wave-rectified, low-pass-filtered, low-frequency tone is multiplied by a high-frequency sinusoidal carrier. The result is a complex waveform known as a transposed-tone (Figure 1.15) and it has been utilised in both electrophysiology and psychophysical experiments due to its unique properties.

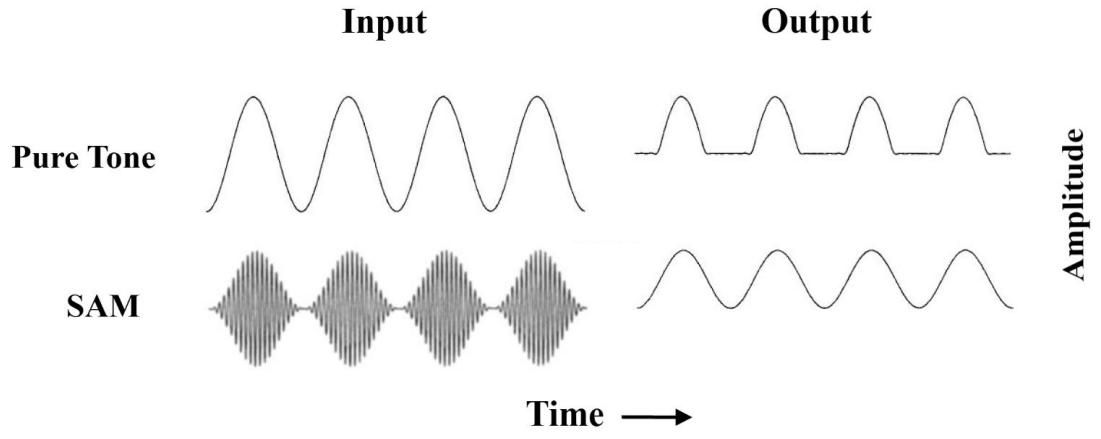


Figure 1.14 The comparative differences in output from the model of monaural processing given a pure tone (top) and a complex tone (bottom) as inputs.

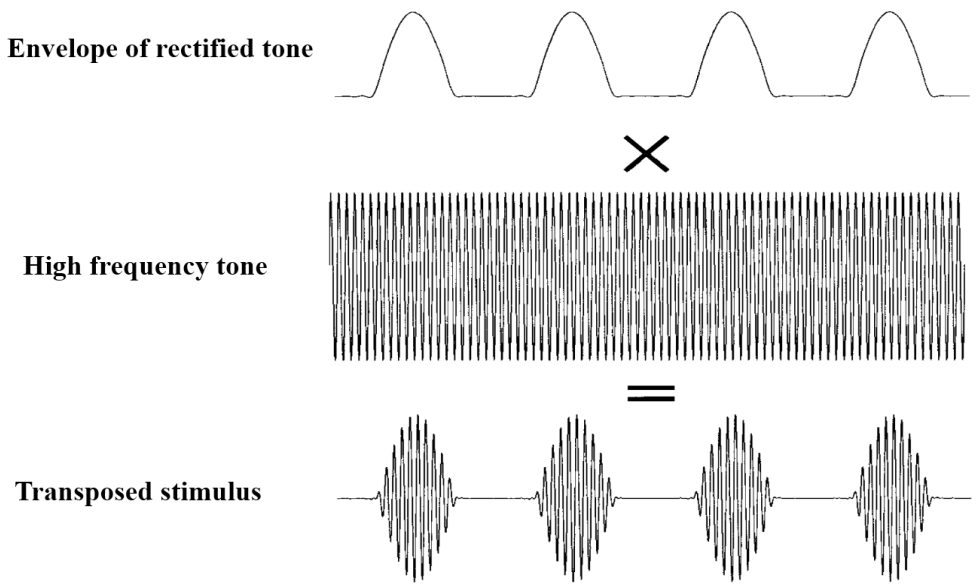


Figure 1.15. A half-wave-rectified, low-pass-filtered, low-frequency tone is multiplied by a high-frequency sinusoidal carrier. The result is a complex waveform known as a transposed tone.

Human Sensitivity to Envelope ITDs

As has been described, action potentials propagating along the auditory nerve are phase-locked to the half-wave rectified waveform. For high-frequency sounds, this is equivalent to the envelope of the waveform. Blauert (1982) showed an effect of this psychophysically by recording listeners' sidedness ratings (from 1-5) for a pure tone and a SAM tone with ITDs ranging from 0 to 500 μs (Figure 1.16). The low-frequency pure tone was always perceived as lateralised to a greater extent (i.e. heard to originate further off to the side) than the SAM tone when the actual ITDs were the same. This suggests that although envelope ITDs can be used in sound lateralisation tasks, they are not as potent a cue as are ITDs conveyed by low-frequency pure tones.

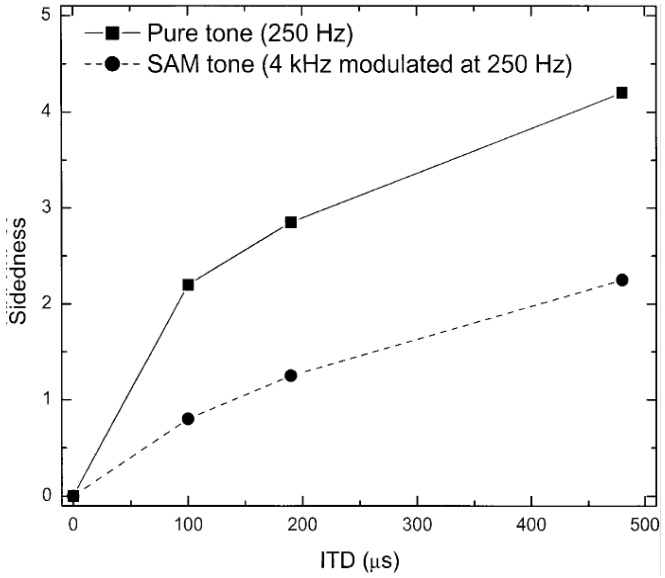


Figure 1.16 A comparison of sidedness rating for ITDs in pure tones and SAM tones. The low-frequency pure tone was always perceived as lateralised to a greater extent than the SAM tone when ITDs were the same. (Blauert, 1982)

Bernstein (2001) measured the sensitivity to changes in ongoing ITD using SAM tones and transposed-tones with a 4000-Hz carrier modulated at frequencies equal to the low-frequency pure tones also tested (Figure 1.17). The threshold ITDs obtained with the transposed-tone stimuli were consistently smaller than those obtained with the SAM tones. The ability to discriminate changes in ITD using transposed-tones also matched closely those measured with low-frequency pure tones below 512 Hz modulation. Carrier tones modulated at 512 Hz however, result in threshold ITDs far larger than for pure tones of equivalent frequency (Figure 1.18) (Bernstein and Trahiotis 2002). It is possible that the central auditory system is unable to process envelope fluctuations as high as 512 Hz, even if this information is maintained in the auditory nerve (Bernstein and Trahiotis (1994)). This may be due to frequency components close to the centre frequency of the carrier tone passing through the peripheral filter but with increasing rates of modulation, frequency-components

that occur further from the peripheral filter centre-frequency fall increasingly within the 'skirt' of the filter resulting in their attenuation in that particular channel. Despite this, off-frequency listening could still take place although the effect of attenuated frequency side-bands may, in part, underlie the reduction in discriminability of ITDs in transposed-tones and SAM tones with modulation frequencies above 256 Hz.

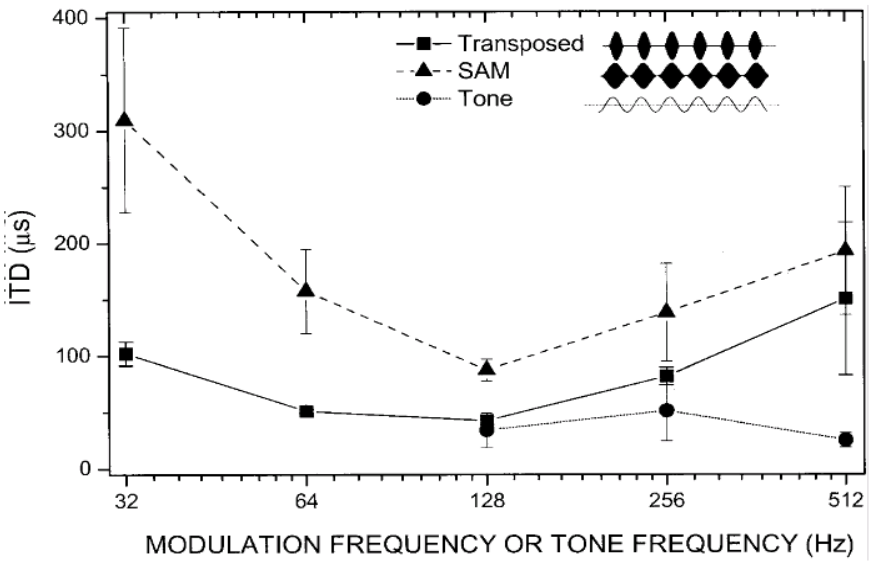


Figure 1.17 Threshold ITDs for transposed tones are more similar to pure tone ITD thresholds than are SAM tone ITD thresholds. From (Bernstein, 2001)

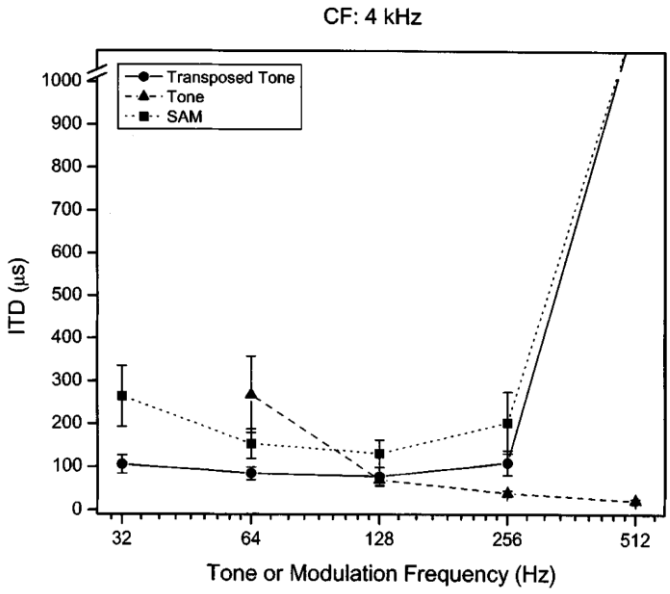


Figure 1.18 At 512Hz modulation, ITD JNDs are far larger for transposed tones and SAM tones than for pure tones. From (Bernstein et al, 2002)

The effect of the peripheral auditory filter width can be assessed by increasing the carrier frequency, resulting in a broadening of the corresponding filter. One might then expect less attenuation of the sidebands and improved ITD thresholds, yet, in fact, the opposite is observed (Henning 1974; Bernstein and Trahiotis 1994) (Bernstein and Trahiotis 2002) (Figure 1.19). It would seem that it is not simply a result of the peripheral processing that produces worse ITD thresholds for transposed-tones with increasing carrier frequencies. An additional factor is the possible existence of a limiting process for the ability of the auditory system to follow envelope fluctuations above 150 Hz. Bernstein and Trahiotis (2002) have suggested that this limiting process operates separately to peripheral band pass filtering.

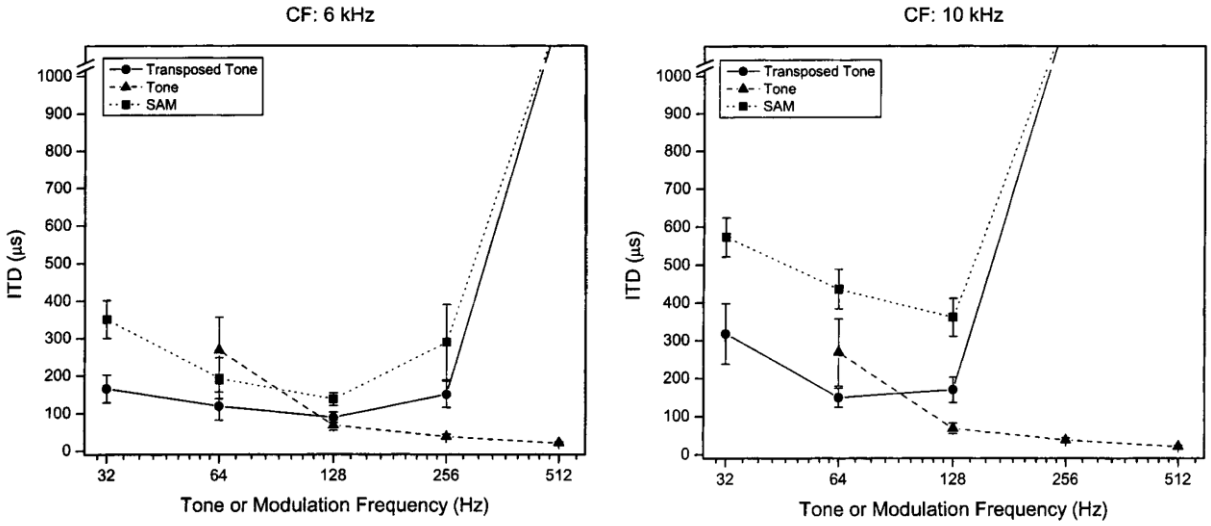


Figure 1.19 ITD JNDs are larger at low modulation frequencies when the carrier wave frequency is increased. From (Bernstein et al, 2002)

2 Sensitivity to IPDs of pure tones and transposed-tones as a function of reference IPD

2.1 Introduction

In this chapter data are presented that compares the sensitivity of human listeners to lateralisation cues within the two domains that form a complex acoustic signal, the temporal-fine-structure and the envelope-structurewill. Such assessments of binaural hearing are set in an historical context spanning 150 years from Helmholtz's Sensation of Tone (1863) and Lord Rayleigh's Theory of Sound (1877), across the studies outlined in Chapter 1, to the modern day assessments of Cochlear Implant (CI) users' ability to discriminate speech in background noise (Friesen, Shannon et al. 2001; Müller, Schon et al. 2002).

Sensitivity to ITDs conveyed in the temporal-fine-structure of low-frequency sounds underpins sound-source localisation in the horizontal plane and contributes to speech reception thresholds in background noise (Strelcyk and Dau 2009), an ability commonly referred to as the 'cocktail party effect' (Cherry 1953). The ITD JND conveyed in the temporal-fine-structure of a sound is smallest (ITD sensitivity is best) at zero degrees azimuth and increases (ITD sensitivity worsens) with increasing reference ITD (e.g. (Mills 1958; Hershkowitz and Durlach 1969; Domnitz 1973; Yost 1974; Domnitz and Colburn 1977) and can be determined for frequencies below about 1400 Hz (Zwislocki and Feldman 1956). Nevertheless, it has long been known that ITDs conveyed in the envelopes of modulated high-frequency sounds are detectable, even when the spectral content of the acoustic signal comprises only frequencies above 1500 Hz (Leakey, Sayers et al. 1958; Henning 1974; Nuetzel and Hafter 1976; Bernstein and Trahiotis 1994). Sensitivity to such envelope ITDs extends the Duplex theory (Rayleigh 1907) beyond the proposal that separate mechanisms underlie localisation of low- and high-frequency sounds. The

temporal precision required to exploit ITD cues in spatial-listening tasks is maintained by specialisations along the auditory pathway, culminating in a process of binaural coincidence detection whereby individual neurons exhibit neural firing patterns modulated by sub-millisecond changes in the ITD of the stimulus (Rose, Gross et al. 1966; Yin and Kuwada 1983).

Most models of binaural processing implement some form of central weighting function in which the distribution of ITD detectors favours ITDs within the ecological range (Colburn 1977; Stern, Zeiberg et al. 1988; Shackleton, Meddis et al. 1992). Even those models in which an explicit representation of ITDs is not assumed (i.e. there is no assumption that each ITD is encoded by a unique detector) posit a preference for discrimination performance biased towards the natural range of ITDs determined by the width of the head (Harper and McAlpine 2004).

The human psychophysical experiments presented here assessed the ability to discriminate IPDs conveyed in the temporal-fine-structure of low-frequency tones and in the envelope of amplitude-modulated high-frequency tones as a function of reference IPDs. The primary aim was to investigate whether changes in ITD JND were related to the physiological range of ITDs determined by the human head size or to the periodicity of the timing cues in the stimulus. Additionally, the experimental data presented extend the parameter space for assessing ITD JNDs beyond the midline, incorporating lower frequencies and finer reference IPD resolution than has been previously reported. The amplitude modulated signal in the present experiments takes the form of the transposed-tone as described by Van de Par and Kohlrausch (1997), and previously applied by Bernstein and Trahiotis (2002; 2003) in their comparison of envelope ITD sensitivity with modulated waveforms. Specifically, these latter authors examined listeners' abilities to discriminate ITDs in low-frequency tones or high-frequency (4-kHz carrier) SAM tones and transposed-tones. Transposed-

tones, which contain a sharper envelope shape compared to SAM tones, are designed to mimic the temporal information conveyed by the low-frequency peripheral auditory system.

The psychophysical component of this thesis took the form of three experiments and are presented here as 1) an assessment of sensitivity to IPDs conveyed in low-frequency pure tones as a function a reference IPD 2) a comparison of sensitivity to ITDs conveyed in low-frequency pure tones and in the modulated envelope of high-frequency transposed-tones and 3) the effect of low-pass masking noise on ITD sensitivity to transposed-tones.

Part of the data collected served as a comparison to existing literature, primarily the work of Bernstein and Trahiotis (2002), in which they compared the ITD JNDs of pure tones and transposed at 0 degrees azimuth for transposed tone modulation frequencies between 32 and 512 Hz and pure tone frequencies between 64 and 1024 Hz. We expected to find similar data trends in our measurements whereby increasing the pure tone frequency from 64 to 1024 Hz results in a decrease of ITD JND and that ITD JNDs conveyed in the envelope of transposed tones is equivalent for modulation frequencies up to 256 Hz to that of pure tones of equivalent frequency. We developed the experimental parameters of the existing literature further by assessing the impact of increasing the reference IPD on IPD JNDs for both pure tones and transposed tones. Our hypothesis for this experiment was that ITD sensitivity will decrease as the perceived location of the signal moves away from the mid-line and that sensitivity to IPDs conveyed in the fine structure of pure tones will be most commonly, if not always, greater than sensitivity to IPDs conveyed in the envelope structure of transposed tones. Regarding the third experiment that assessed the effect of low-pass masking noise on ITD sensitivity, we hypothesise that all of the acoustic information required to lateralise the high frequency tranposed tone is present in the high frequency auditory channel and that ITD

sensitivity is not simply a result of low frequency cochlear distortion. ITD sensitivity should therefore remain largely constant with the addition and subsequent increase in the level of masking noise. Any effect of the masking noise is hypothesised to be caused by perceptual distraction or by contributing additional distortion to the auditory periphery. In either case, performance would not be expected to be affected by the addition of masking noise to the degree that it be required as a prerequisite to the ITD discrimination tasks utilising high frequency transposed tones.

2.2 Methods

2.2.1 Subjects

Nine normal-hearing listeners (three female) aged 22-26 years participated in the reported experiments. Seven participated in the pure-tone experiment, although one was unable to demonstrate consistent ITD detection for a 1024 Hz pure tone and so was excluded from the analysis. Three participated in the transposed-tone experiment and three in a third experiment (one of whom took part in experiment 1 and 2 and two who took part only in experiment 3) assessing the effect of low-pass masking noise on ITD sensitivity. Each of the participants was given a consent form to voluntarily sign to indicate that they were willing and able to take part in the hearing tasks and that detailed what was required of them during the listening task. Although the sample size may appear small, by comparing the data from the three most reliable subjects it was possible to observe that additional subjects did not provide more robust trends or reliability of thresholds but rather increased the mean thresholds due to some participants who struggled with certain tasks and increased the standard deviation of the population means. Furthermore, the first experiment that had seven participants consisted of approximately 12.5 hours of total testing time for each participant.

As the nature of the task was highly repetitive the threshold for boredom and tiredness was generally low which meant that only one hour of testing could be done by a subject on any given day in order to maintain threshold reliability. The result of this was many separate visits to the sound-proof booth which meant that the inclusion of any additional subjects was highly undesirable as well as unnecessary. Prior to data collection, all subjects took part in two hours of training with the experimental stimuli in order to avoid any learning effects in the data. A subject would be excluded from taking part in the full test protocol if they failed to understand the task or could not perform the task during the training. Eight of the subjects received financial compensation for the time spent taking part in the experiment whilst the ninth subject was the author DG.

2.2.2 Apparatus and Stimuli

Experiments took place in a triple-walled, sound attenuating booth. A computer monitor displayed visually the interval options and gave feedback on each subject response by way of a red (incorrect) and green (correct) marker. A response was made by either selecting an on-screen button with a computer mouse or by selecting the associated key on a number pad. Threshold IPD was measured for two types of stimuli: low-frequency pure tones and low-frequency tones transposed to a 4000 Hz carrier. The stimulus duration was 300 ms (including 20 ms raised-cosine ramps). The frequencies of the pure tones were 64, 128, 256, 512 or 1024 Hz (Experiment 1) and the rates of modulation of the transposed stimuli were 32, 64, 128 and 256 Hz (Experiment 2). Stimuli were generated digitally utilising a MATLAB script used for a wide range of psychophysical experiments that allowed for the creation and presentation of acoustic signal with a wide range of parameter variables. The stimuli were presented via an SB Audigy 2 ZS sound card via Sennheiser HDA 200 headphones at 75 dB SPL as established by a calibrated system consisting of a

Brüel & Kjær type 2636 measuring amplifier and a type 4152 6cc artificial ear simulator. The reference IPD was varied for each experimental condition between 0, 30, 60, 90 or 120°. The reference IPDs were presented as either left-leading or right-leading and the deviant IPD had always a higher magnitude (i.e. deviating away from zero IPD). Experiment 1 and 2 therefore consisted of assessing IPD JNDs for 10 different reference IPD conditions. In Experiment 3, measuring the effect of a low-pass masking noise on ITD sensitivity to transposed-tones, modulation frequencies of 128 and 256 Hz, and a reference IPD of 0° were employed. The overall level of the tones was slightly higher than in experiments 1 and 2 (80 dB SPL compared to 75 dB SPL). Masking noise was a 1300 Hz low-pass filtered diotic noise of 400 ms duration (including 20 ms raised-cosine ramps) The masking noise was presented at 3 different overall levels of 60, 70 and 80 dB SPL (corresponding to spectrum levels of 29, 39 and 49 dB).

2.2.3 Procedure

Threshold IPDs were measured using a three-interval, two-alternative-forced choice adaptive task. The inter-stimulus interval was 400 ms. One of the second or third intervals held the same IPD as the leading interval, and is referred to as the reference interval, whilst the other held a different IPD (target interval). The subject was required to choose which of these two intervals the target interval was. A one-down one-up procedure was used for the first two reversals followed by a three-down one-up staircase for the following 13 reversals. The threshold IPD was derived from the last eight reversals resulting in the 79.4% correct point (Levitt 1971). The start value of the target IPD magnitude was 45° larger than the reference IPD and was varied by a factor of $\sqrt{2}$ (1.414) for the first five reversals and $\sqrt[4]{2}$ (1.189) for the last ten reversals. Each test subject completed each stimulus frequency condition in random order three times. The 10 reference

IPDs within each frequency condition were also presented in random order using a random-order function built in to the MATLAB script. Statistical analyses, in the form of repeated-measure ANOVAs, were performed on the log-transformed geometric mean data.

Pilot studies (described in Appendix V) revealed the range of reference IPDs and test frequencies that would allow for detailed analysis of the test subject's IPD sensitivity without assessing every possible reference IPD and test frequency. The pilot studies enabled test parameters to be selected for use in the final test protocol that maximized the efficiency and accuracy of data collection.

2.3 Results

2.3.1 Experiment 1: Sensitivity to IPDs conveyed in low-frequency pure tones as a function of reference IPD

Figure 2.1 plots IPD JNDs for pure tones (left column) for a frequency range of 64 Hz to 1024 Hz as a function of reference IPD, for one example subject (AL; Figs. 1a-e). The error bars plot one standard deviation from the geometric mean. Data are plotted for left-leading (negative) and right-leading (positive) reference IPDs for the range ±120°. Repeatability of threshold recordings was generally good. Thresholds were lowest for the 128- and 256-Hz pure tones at reference IPDs below 60°, and were highest for the 64 and 1024-Hz pure tones, and at 90° and 120° reference IPDs for all pure-tone frequencies except for 512 Hz where IPD thresholds were fairly constant across reference IPD. Average data for 6 subjects are plotted in the left column of Figure 2.2(a-e). Error bars plot one geometric standard deviation from the mean. The dashed lines in Figure 2.2(a-e) represent the geometric mean thresholds of the 3 subjects who also took part in the transposed task, selected due to their consistency during experiment 1. The geometric mean threshold data from the 6 subjects shows a similar pattern to the

individual subject data in Figure 2.1, with stimulation at 256 and 128 Hz for smaller reference IPDs, eliciting best performance. A repeated-measures ANOVA was performed on the log geometric mean thresholds for all 6 subjects, with frequency and reference IPD as factors. This revealed significant differences between thresholds as a function of frequency ($F_{4, 20}=24.812$, $p<0.001$) and reference IPD ($F_{4, 20}=42.975$, $p<0.001$), with thresholds improving (decreasing) as a function of increasing frequency and thresholds worsening (increasing) as a function of increasing reference IPD, and a significant interaction of frequency and reference IPD ($F_{16, 80}=5.111$, $p<0.001$), with the effect of reference IPD greater for lower than for higher frequencies. The frequency-dependent elevation in thresholds as a function of reference IPD (Figure 2.2a-e) suggests that ITDs are less-accurately encoded as reference delay increases.

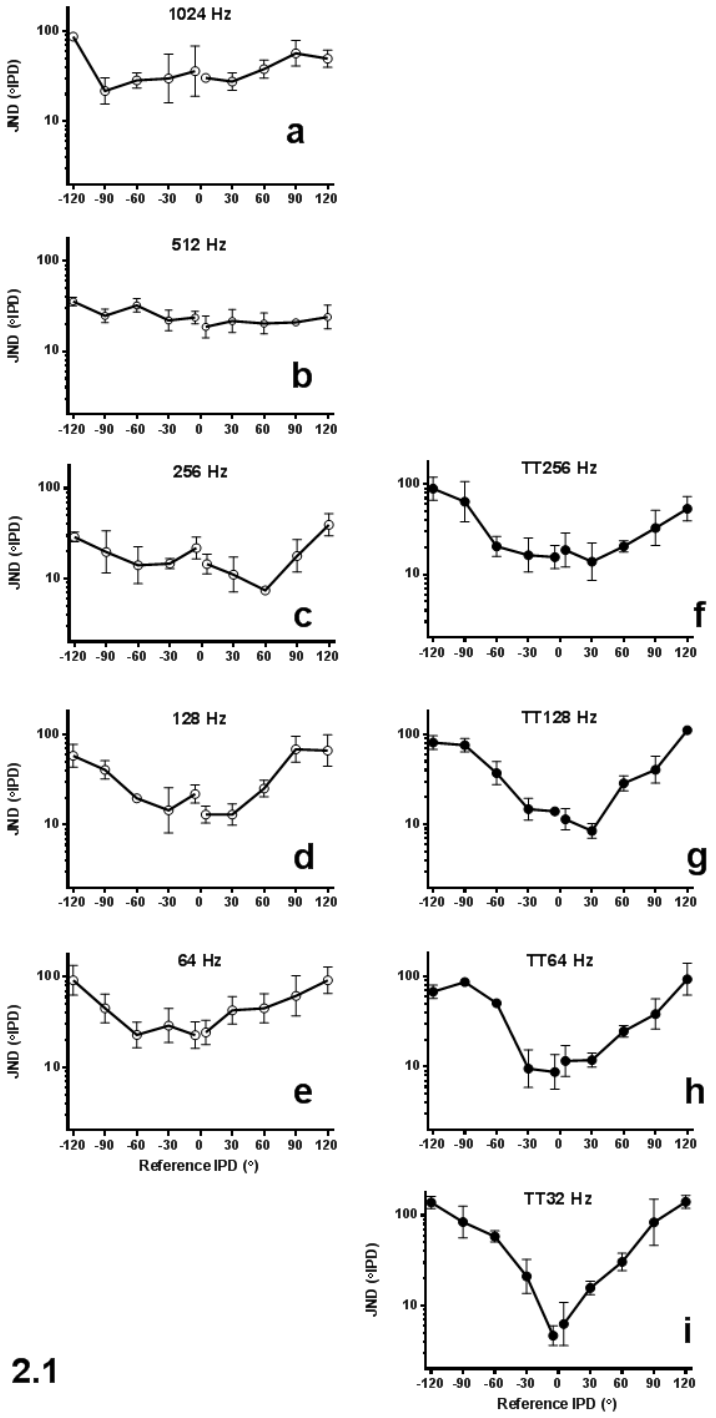


Figure 2.1

IPD JNDs for pure tones (left column) for a frequency range of 64 Hz to 1024 Hz as a function of reference IPD, for one example subject (AL; Figs. 1a-e). The error bars plot one standard deviation from the geometric mean. Data are plotted for left-leading (negative) and right-leading (positive) reference IPDs for the range ±120°. The right column plots IPD JNDs as a function of reference IPD for subject AL (Figs. 2.11f-i) for IPDs conveyed in the envelopes of 4-kHz transposed tones, for modulation frequencies in the range 32 to 256 Hz.

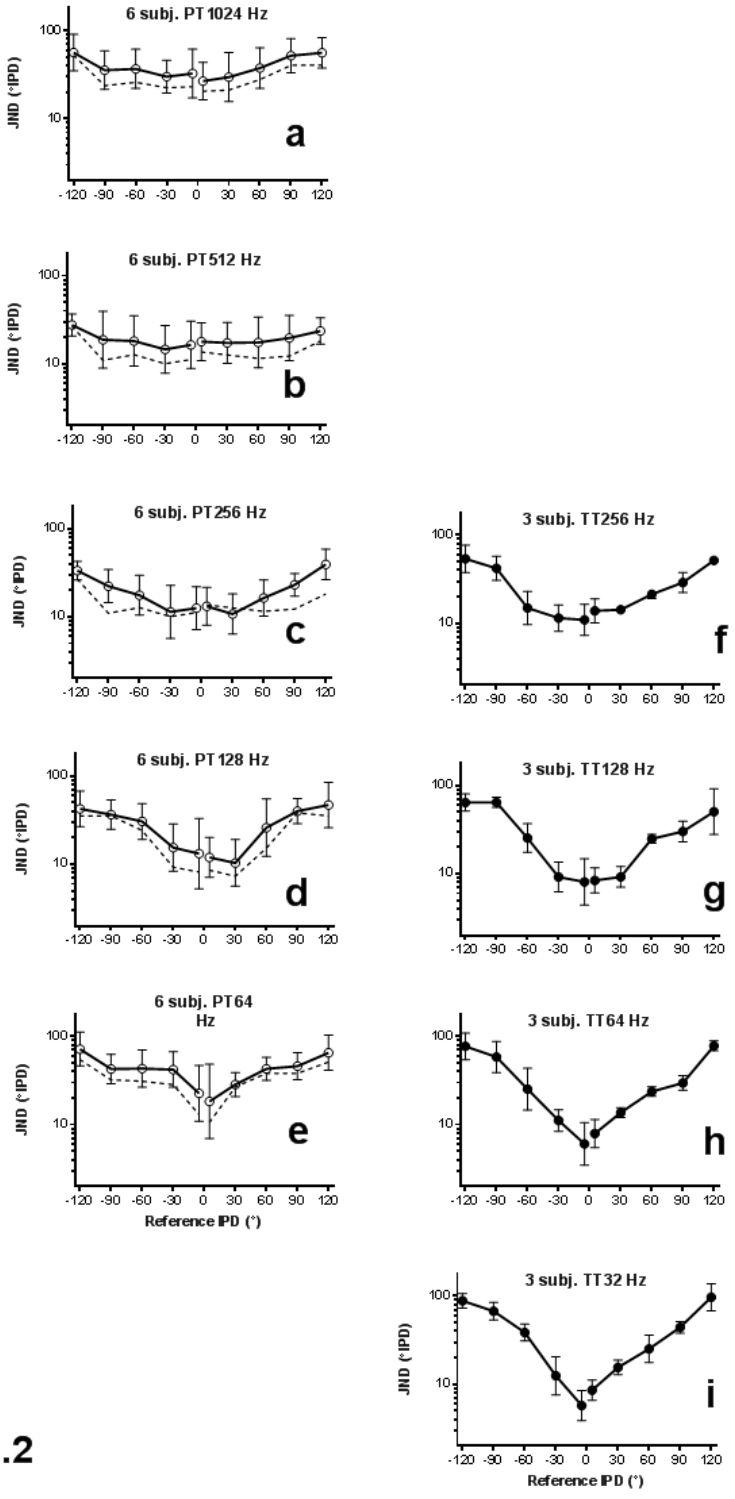


Figure 2.2

Same as Figure 2.1 but average responses for 6 (solid line (a-e) and 3 subjects (f-i). Dashed line in (a-e)represents the average response of the same 3 subjects as in (f-i).

Figure 2.3a displays threshold ITDs for the three highest pure-tone frequencies (256, 512 and 1024 Hz) as a function of reference ITD, where both reference IPDs and threshold IPD are converted to the equivalent ITD. (Left- and right-leading ITD thresholds are averaged and JNDs are shown as a function of the reference ITD). For these frequencies all but two of the reference IPDs assessed correspond to ITDs within the approximate human physiological range of ±670 μs (dashed line). The change in performance as a function of reference ITD is seen to be frequency dependent as despite threshold ITDs at 1024 Hz being lower than at all other frequencies for a reference of zero ITD, performance starts to decline (i.e. thresholds increase) more rapidly for this frequency as reference ITD increases. Across these three frequencies, for a reference ITD of 325 μs (corresponding to reference IPDs of 120°, 60° and 30° IPD, respectively for 1024, 512 and 256 Hz) thresholds at 1024 Hz are the highest for this set of three frequencies. Interestingly, threshold performance appears to actually improve slightly with increasing reference ITD before it starts to decline for the 256 and 512 Hz tones.

The relative decline in threshold sensitivity as a function of reference ITD is plotted for all five pure-tone frequencies in Figure 2.3b, as the ratio of the threshold at each reference IPD to that at zero IPD. This facilitates the comparison of threshold decline across frequency. With a ratio at zero reference equal to 1.0, the relative insensitivity of the lower-, compared to the higher-, frequency tones is evidenced by the slower decline in performance at the lower-frequencies. Even out to reference ITDs of 2600 μs, the decline in performance relative to zero ITD remains less for 64-Hz pure tones than it does for 1024-Hz pure tones at 325 μs.

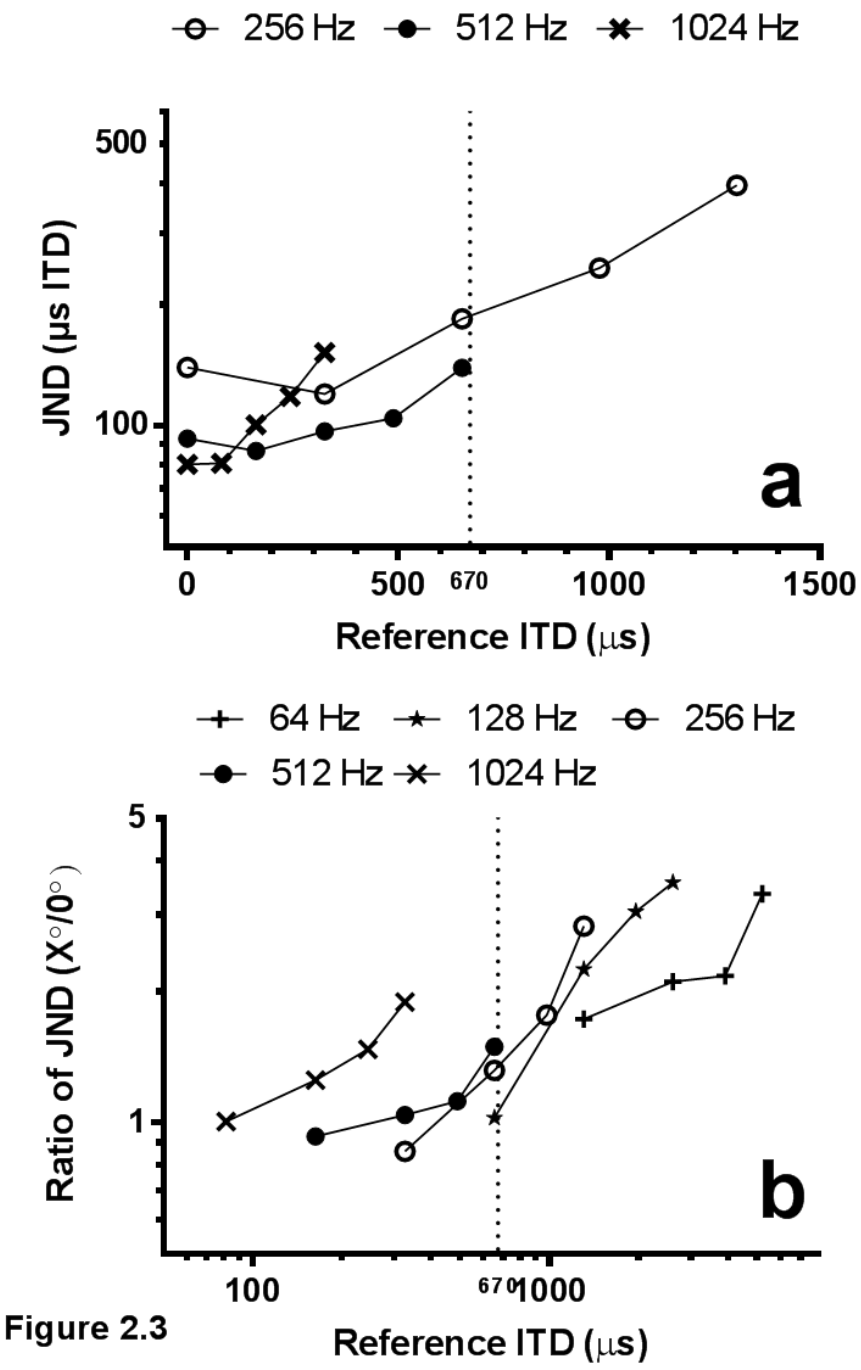


Figure 2.3

Figure 2.3 a) Threshold ITDs as a function of reference delay for the three highest pure-tone frequencies (256, 512 and 1024 Hz), IPDs were converted to the equivalent ITD. b) Relative JNDs for ITD normalised to the JND at zero IPD as a function of reference IPD. The reference IPDs were converted to the equivalent ITD. The dashed line represents an estimation of the human physiological range of ITDs (670 µ*s*).

2.3.2 Experiment 2: Comparison of sensitivity to ITDs conveyed in low-frequency pure tones and in the modulated envelope of high-frequency transposed-tones

The three listeners with the lowest and most reliable thresholds in the pure-tone discrimination task were assessed with respect to their sensitivity to IPDs conveyed in the envelopes of 4-kHz transposed-tones, for modulation frequencies in the range 32 to 256 Hz. The right column of Figure 2.1 plots threshold IPDs for left- and right-leading reference IPDs as a function of reference IPD for example subject AL (Figs. 2.1f-i), and the right column of Figure 2.2 plots threshold IPDs for transposed-tones averaged across all three listeners (Figs. 2.2f-i). At frequencies for which thresholds for both pure tones and transposed-tones were assessed (64, 128 and 256 Hz), thresholds were either similar for reference IPDs close to zero or, in the case of 64 Hz, better for transposed-tones. Thresholds were generally higher for larger reference IPDs for the transposed-tones compared to the pure tones indicating a more rapid worsening of IPD discrimination of the former. For the 32 Hz and 64 Hz transposed-tone especially, thresholds increased considerably as reference IPD increased from zero.

Figure 2.4 compares ITD discrimination thresholds for pure tones and transposed-tones, averaged across the three listeners who took part in both the transposed-tone and pure tone tasks. The mean thresholds of all 6 subjects from the pure tone task are represented by the dashed line. ITD JNDs are plotted as a function of frequency, for five different reference IPDs. In each panel, threshold values are averaged across left-leading and right-leading delays. For a reference of 0° IPD (Figure 2.4a), discrimination thresholds for the lowest frequency for which both stimuli were assessed (64 Hz) shows sensitivity to ITDs conveyed in high-frequency transposed-tones to be better than for pure tones; similar to Bernstein and Trahiotis (2002) mean thresholds for pure tones are about double

those for transposed-tones (for the same 3 subjects) at this frequency. For frequencies of 128 and 256 Hz, thresholds were similar for transposed and pure tones, again consistent with Bernstein and Trahiotis (2002). All three subjects were also able to discriminate reliably ITDs conveyed in a 32-Hz transposed-tone; although absolute sensitivity in terms of ITD was relatively poor (discrimination thresholds were in the order of 600 μs). Shifting the reference IPD away from zero (Figs. b-e) had the effect of elevating discrimination thresholds in a manner dependent on frequency and on whether the ITDs were conveyed in low-frequency pure tones or transposed-tones. For a reference of 30° IPD, thresholds at 32 Hz (transposed) were substantially elevated (to well beyond the ecological range of ITDs). Thresholds at 64 Hz for both transposed-tones and pure tones were also elevated, although to a lesser degree. In contrast, for transposed-tones at 128 and 256 Hz, and for pure tones at 256, 512 and 1024 Hz, thresholds were relatively unaffected by changing the reference IPD to 30°. In fact, for pure tones, thresholds appear to be slightly lower (better) for a reference IPD of 30° than 0°, at 256, 512 Hz. As reference IPD increased beyond this, however, discrimination performance deteriorates considerably; for a reference IPD of 90°, thresholds lower than the upper limit of the human ecological range (dashed line) were limited to a single transposed frequency, 256 Hz. A repeated-measures three-way ANOVA was performed on the log-transformed geometric mean threshold IPD data, with domain (TFS *vs.* envelope), frequency and reference IPD as factors. Significant main effects of frequency ($F_{2,4}$=27.258; p=0.005) and reference IPD ($F_{4,8}$=33.25, p<0.001) were observed; the factor domain was not significant ($F_{1,2}$=2.78 p=0.237). In addition, significant interactions between domain and frequency ($F_{2,4}$=9.004, p=0.03), between frequency and reference IPD ($F_{8,16}$=9.28, p<0.001) and domain and reference ($F_{4,8}$=4.721, p=0.03) were observed, with the effect of reference IPD being greater for transposed-tones than for pure tones, and greater for lower, compared with higher, frequencies.

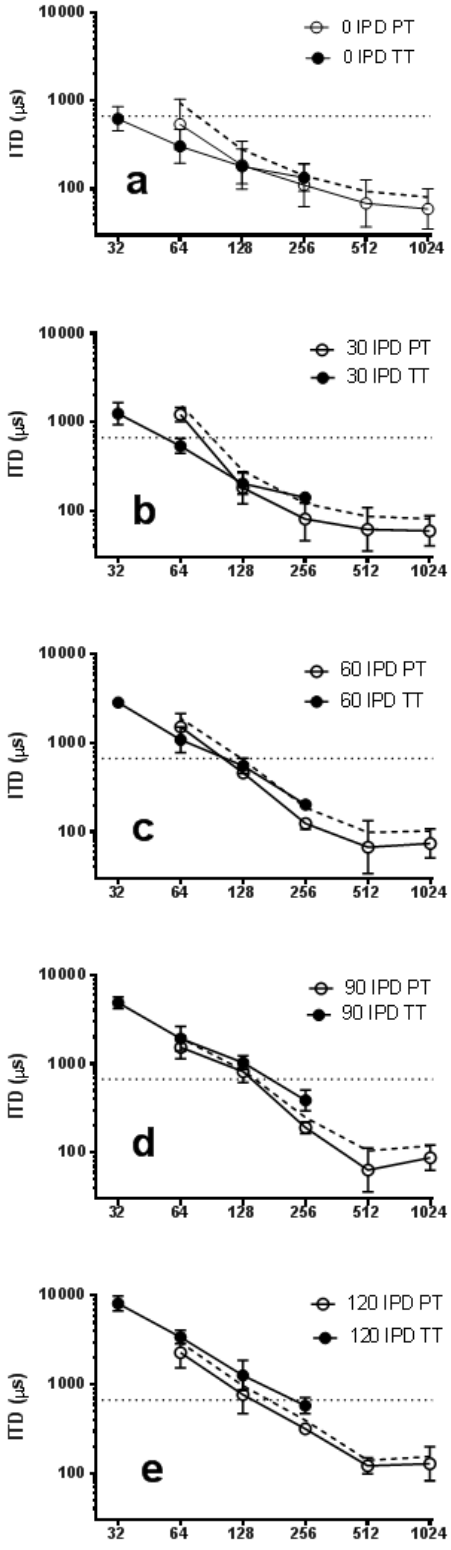


Figure 2.4

Threshold ITDs as a function of frequency for pure tones (open symbols) and transposed tones (filled symbols) for reference IPDs of 0°-120° (a-e). The values are geometric means of ITD JNDs from 6 subjects (dashed lines) and 3 subjects (solid lines) for the pure tone task and for the same 3 subjects in the transposed tone task.

The generally more-rapid decline in performance with increasing reference IPD for transposed-tones compared to pure tones is further explored in Figure 2.5, which plots the ratio of the threshold at each reference IPD to the threshold at a reference IPD of 0°, as a function of reference IPD, comparing data for pure tones and transposed-tones for 64, 128 and 256 Hz (Figs. 2.5a-c) and for all pure-tone (Fig. 2.5d) and transposed-tone (Fig. 5e) frequencies. For the three frequencies at which thresholds in both domains were obtained, significant effects of frequency ($F_{2, 4}$=22.04, p=0.007) and reference IPD ($F_{4, 8}$=20.605, p<0.001) were observed, but no significant effect of domain ($F_{1, 2}$=1.839, p=0.308). In addition there was a significant interaction between frequency and reference ($F_{8, 18}$=7.452, p<0.001) and domain, frequency and reference ($F_{8, 16}$=4.195, p<0.007). Note especially, that despite thresholds for 64-Hz transposed-tones being lower than for 64-Hz pure tones (by nearly a factor of two) for a reference IPD of zero, performance decayed more rapidly for the transposed-tones as reference IPD increased.

For all 5 frequencies at which sensitivity to ITDs conveyed in pure tones were assessed (64, 128, 256, 512 and 1024 Hz), significant effects of frequency ($F_{4,8}$=4.289, p<0.038) and reference IPD ($F_{4,8}$=6.874, p=0.011) were observed, as well as a significant interaction between the two ($F_{16,32}$=3.58, p<0.001) and similarly, for all 4 frequencies at which sensitivity to ITDs conveyed in transposed-tones were assessed (32, 64, 128 and 256 Hz), significant effects of frequency ($F_{3,6}$=22.978, p<0.001) and reference IPD ($F_{4,8}$=52.858, p<0.001) were observed, as well as a significant interaction between the two ($F_{12,24}$=11.652, p<0.001). Performance declined more rapidly as a function of reference IPD for lower, compared to higher, frequencies in both the temporal fine structure and the envelope domains.

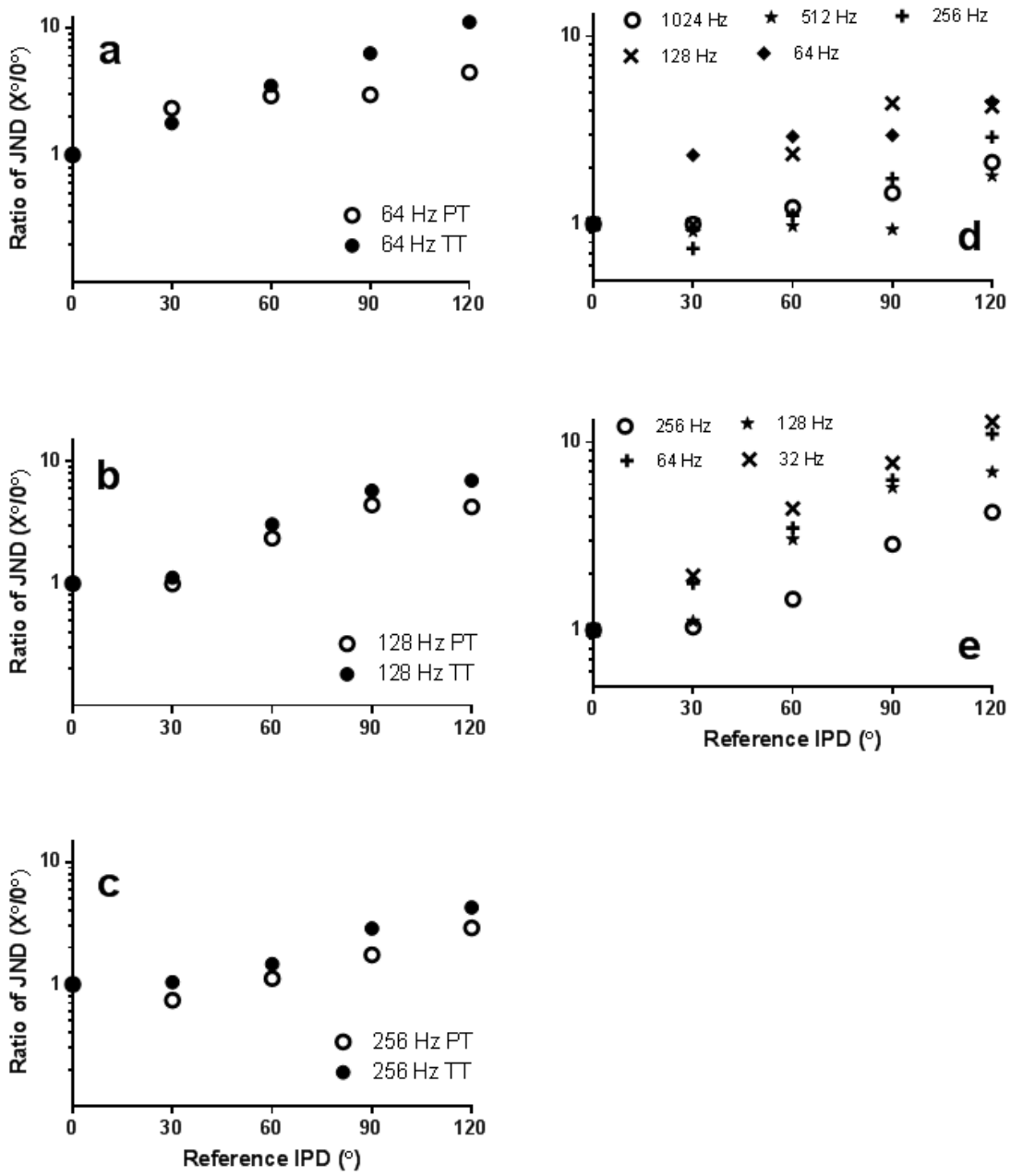


Figure 2.5

a-c) Ratio of threshold ITDs, with respect to threshold at 0° reference IPD, as a function of reference IPD for pure tones (open symbols) and transposed tones (closed symbols) for 64, 128 and 256 Hz. d) Ratio of threshold ITDs as a function of reference IPD for pure tones in the range 64-1024 Hz. e) Ratio of threshold ITDs as a function of reference IPD for transposed tones in the range 32-256 Hz.

2.3.3 Experiment 3: The effect of low-pass masking noise on ITD sensitivity to transposed-tones

A potential explanation for the existence of sensitivity to envelope ITDs conveyed in high-frequency sounds is that it reflects the contribution of distortions generated by the non-linear response of the cochlea (Pressnitzer and Patterson 2001). Such distortions, resolved on the low-frequency region of the basilar membrane, would be expected to reflect the ITD sensitivity conveyed in the low-frequency temporal fine structure (Klumpp and Eady 1956; Zwislocki and Feldman 1956; Nuetzel and Hafter 1976; Henning 1980; Bernstein and Trahiotis 1994) and has been demonstrated to contribute to the perception of the pitch of sounds comprising harmonics lying beyond the presumed upper limit of phase-locking to the temporal fine structure (Pressnitzer and Patterson 2001). Whilst masking distortions with low-pass noise is a common means of reducing their potential contribution to listening tasks, the level of noise required to achieve such masking is usually only assumed and, indeed, can only ever be estimated indirectly. Nevertheless, given the exquisite temporal resolution of the low-frequency auditory pathways, it is essential to assess their potential contribution to coding of envelope ITDs. To this end, the impact of masking noise on envelope ITD discrimination was assessed as a function of both masker level and modulation frequency. Three subjects performed the same ITD discrimination task as in Experiment 2 for 128 and 256 Hz transposed-tones but with varying levels of low-pass masking noise or no low-pass masking noise (see Methods).

Figure 2.6 illustrates the effect of increasing levels of masking noise on threshold ITD, for three subjects with transposed-tones of 128 Hz (open symbols) and 256 Hz (filled symbols) (Figs. 2.6a-c; mean for 3 subjects in Fig. 2.6d) for a reference IPD of 0°. A repeated measures ANOVA, with modulation frequency and noise level as factors revealed no significant effect of frequency

($F_{1,2}$=0.88, p=0.795) or noise ($F_{3,6}$=4.067, p=0.068), and no significant interactions. All three listeners showed relatively good performance (low thresholds) for transposed-tones with no noise for both frequencies and this performance was maintained as the level of the low-pass masker was increased to 80 dB. The data suggest that threshold performance in quiet is determined by sensitivity to ITDs conveyed in the high-frequency modulated envelopes of the transposed-tones, rather than by low-frequency distortions generated by the non-linear response of the cochlea.

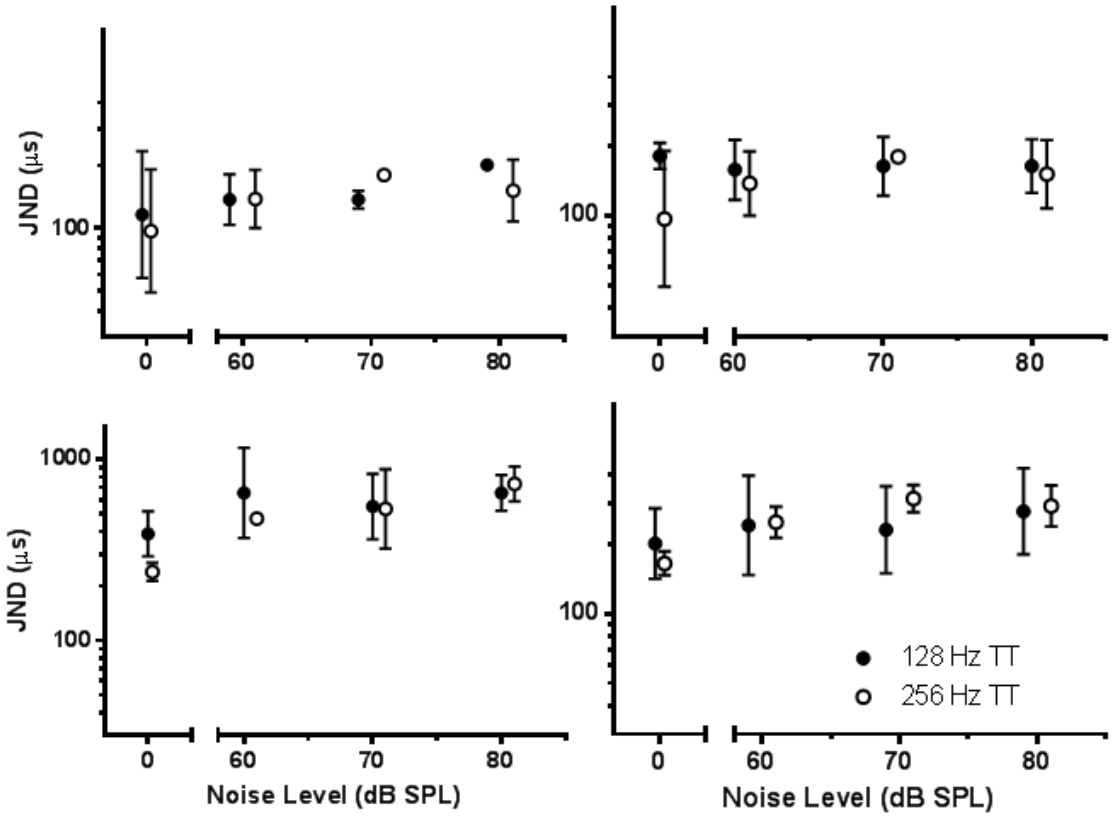


Figure 2.6

a-c) Threshold ITDs from zero reference IPD as a function of level of a 1300 Hz low-pass filtered masking noise for 3 listeners with 128-Hz (filled symbols) and 256-Hz (open symbols) transposed tones. d) Mean threshold from all 3 listeners.

2.4 Discussion

2.4.1 Experiment 1: Lower IPD thresholds at reference IPDs close to 0°

The data from the present experiments enables the construction of a more complete map of human sensitivity to IPDs for reference IPDs up to 120°. One observation that replicates the findings of previous studies such as Mills (1958), is the enhancement of IPD sensitivity for reference IPDs close to 0° azimuth. Increased sensitivity to IPDs at reference IPDs close to 0° is an auditory parallel, in some respects, to the visual fovea. The visual fovea is a part of the macula region of the retina, the centre of which is the foveal pit (Polyak 1941), a depression in the retina that results from a specific arrangement of supporting fibres, ganglion cells and bipolar cells arranged in such a way as to allow for a compact arrangement of sensory cells. While the visual fovea is responsible for sharp central vision, an 'acoustic fovea' could be used to describe an area in auditory space in which IPD discrimination, and sound localisation ability, is greatest.

Most models of binaural processing implement some form of central weighting function in which the distribution of ITD detectors favours ITDs within the physiological range. This is determined by the width of the head and for humans is calculated as approximately 660 μs (Woodworth 1938; Feddersen, Sandel et al. 1957). While the general trend across all acoustic stimuli is for improved IPD thresholds for reference IPDs close to 0° IPD, the manner in which the IPD threshold changes across this range appears to be dependent on stimulus frequency. For the 1024 Hz pure tone stimuli the ITD thresholds measured are consistently below 200 μs even for large reference IPDs. This can be attributed to the fact that all reference IPDs for a 1024 Hz pure tone equate to ITDs well within the human physiological range, a 120° reference IPD for example being

equal to an ITD of 325.5 μs (for an example head width of 23.5cm). As we decrease the pure tone frequency, the reference IPDs used will at some point fall outside of the physiological range and we have observed the effect this has on threshold IPDs. While all reference IPDs for a 1024 Hz and 512 Hz pure tone are within the physiological range, for a 256 Hz pure tone, a reference IPD above 60° falls outside of the physiological range. At 128 Hz a reference IPD above 30° is outside of the physiological range while all of the reference IPDs above 0° for a 64 Hz pure tone are larger than the physiological range. It is possible therefore to interpret the IPD thresholds at which there is a sudden increase, as corresponding to the point that coincides with the reference IPD at which the target stimulus is beyond the physiological range.

.

2.4.2 Experiment 2: ITD sensitivity for transposed-tones and pure tones as a function of reference IPD

Overall, ITD thresholds tend to decrease with increasing frequency, but the rate at which this occurs varies according to the reference IPD. Factoring out absolute sensitivity, a clear and systematic relationship is evident between the extent to which thresholds increase as a function of reference ITD and the domain in which those ITDs are conveyed.

The low-pass characteristics of envelope discrimination performance (Bernstein and Trahiotis 2002) argues for the existence of a low-pass filter in high-frequency auditory channels that limits the ability to exploit temporal information conveyed in the temporal envelopes of sounds. With regard to envelope ITDs, Bernstein and Trahiotis (2002) demonstrated the existence of such a filter, with ITDs conveyed in transposed-tones higher than 256 Hz virtually undetectable. Additionally, the filter cut-off frequency appears progressively more low-pass with increasing centre frequency, with discrimination performance steadily deteriorating between 4 and 10 kHz. Since

peripheral filters broaden (in Hertz terms) with increasing centre frequency, and therefore might be expected to represent higher modulation rates more effectively, an envelope filter of central origin is plausible. A similar filter is evident in the monaural domain, where modulation-rate discrimination performance also falls with increasing centre frequency, suggesting a common origin for low-pass temporal filtering in the auditory brain.

The ability to exploit ITDs conveyed in the envelopes of high-frequency sounds has potential interest for the field of cochlear implantation. Currently, most cochlear-implant stimulation strategies convey only information concerning the rectified stimulus envelope of the sounds, discarding the temporal fine structure. Further, with electrical stimulation favouring more basal, high-frequency regions of the cochlea, any temporal information conveyed by a CI device is likely to excite auditory nerve fibres projecting to central auditory neurons that may be less-temporally precise in their neural representation of sound. With the advent of bilateral cochlear stimulation promising the potential of restoring binaural hearing, a critical factor in the design of digital sound processors will be how they deliver the relatively low modulation rates at which high-frequency temporal information can be accessed by implant listeners (or indeed, by normal-hearing listeners in which sounds have been high-pass filtered.

Experiment 1 examined the effect of reference delay on IPD (or, equivalently, ITD) discrimination thresholds when IPDs were conveyed in the temporal fine structure of low-frequency sounds. Threshold performance was found to be dependent on reference IPD, with thresholds increasing with increasing reference IPD. Models of binaural processing usually incorporate some form of central-weighting function, with more detectors tuned to ITDs near zero, often evenly distributed across the human physiological range, and a decreasing number of detectors beyond this

(Colburn 1977; Stern, Zeiberg et al. 1988). Such a model might be expected to account for the decline in performance with increasing reference IPD although several factors generally not considered in these models are evident in the present data such as the more rapid decline in performance with increasing reference ITD for higher vs. lower frequencies. For those frequencies (256 to 1024 Hz) at which absolute performance was best (and virtually identical for a reference delay of zero), performance declined more rapidly with increasing reference ITD, even within the physiological range of ITDs, for the higher frequencies. Although the relative decline in performance with increasing reference delay was ultimately greater over the 120° range of reference IPDs at lower compared to higher frequencies (Figure 2.3), this was not apparent within the physiological range. Indeed, for 128-Hz tones, where absolute performance at all reference delays was relatively poor compared to 256, 512 and 1024 Hz, thresholds at a reference ITD of 651 μs (corresponding to 30° IPD) were still numerically better than at zero reference. The decline in performance within and beyond the physiological range of ITDs appears to be strongly frequency dependent.

Experiment 2 replicated and extended the findings of Bernstein and Trahiotis (2002), demonstrating the correspondence of ITD sensitivity conveyed in pure tones and transposed-tones for frequencies up to 256 Hz. One conclusion from that study was that by rendering as similar as possible the peripheral representation of temporal information in low- and high-frequency auditory channels converging onto the binaural integrator, binaural sensitivity is comparable across the low-frequency and high-frequency domains. This equivalence has previously been hypothesized by Colburn and Esquissaud (1976). The caveat to this is the existence of a relatively lower-pass filter in the envelope domain. Given that monaural peripheral representations of temporal information are unchanged when the reference IPD is altered, the relatively-reduced discrimination performance when the reference IPD of the envelope,

compared to that of the TFS, is increased suggests a fundamental difference in the processing of binaural temporal information across the tonotopic gradient. Given the different neural pathways through which such temporal information is conveyed, it seems likely that limitations in the conveyance of temporal information in high-frequency channels can account for this difference. One plausible explanation is that although small reference ITDs are equally-well represented, in terms of the number or fidelity of delay channels close to zero, the number or quality of delay channels declines more rapidly with increasing interaural delay in the high-frequency, modulation domain compared to the low-frequency, fine-structure domain.

Other possible mechanisms to account for the difference in performance in low- and high-frequency binaural channels include differences in intrinsic neural mechanisms contributing to sensitivity to interaural delays. Griffin et al (2005) reported generally less acute responses of individual midbrain auditory neurons to ITDs conveyed in the transposed envelopes of high-frequency (>2 kHz) tones compared with low-frequency pure-tones (although the best responses were comparable), with ITD sensitivity limited to rates below about 300 Hz. The possibility that temporal information, despite being identical at the output level of the auditory nerve fibres is nevertheless processed to a different degree of fidelity post-synaptically is also suggested by Sayles and Winter (2008). They demonstrated that neurons across the tonotopic gradient in the cochlear nucleus, the first neural stage of processing beyond the auditory nerve fibres, represent temporal information with less accuracy for increasing sound frequency, despite that temporal information likely being strongly conserved across the tonotopic gradient at the output level of the peripheral filters. The capacity of neurons to convey temporal information, particularly under conditions of reverberation, declined systematically as a function of the characteristic frequency of the neurons they recorded, suggesting that the temporal acuity of the action potentials converging onto binaural neurons from each ear would depend on

from where along the tonotopic gradient they received their input. Nevertheless, it remains the case that an ITD-specific degradation of performance between low- and high-frequency auditory channels is required to account for the differences in performance observed with increasing reference delay. Low-frequency ITD information is generally thought to converge onto neurons in the medial superior olive specialized for temporal processing capabilities (Grothe and Sanes 1994), whereas ITD information conveyed in the envelope of high-frequency sounds may be processed by the, possibly less temporally-precise, neurons of the lateral superior olive (Joris and Yin 1995), offering a plausible explanation for the differences observed for ITD sensitivity conveyed in low-frequency sounds and in the modulated envelopes of high-frequency sounds.

2.4.3 ITD sensitivity of transposed-tones as a function of masker intensity

Experiment 3 demonstrated that low-pass masking noise for 128- and 256-Hz transposed-tones does not appear to significantly alter threshold performance. Threshold performance at 128- and 256-Hz modulation rates is therefore likely driven by temporal information conveyed in high-frequency auditory channels, even in the absence of low-pass masking noise designed to remove the capacity of listeners to exploit cochlear-generated distortions in ITD tasks. Bernstein and Trahiotis (2003) reported no effect of low-pass masking noise on lateralisation performance for 256-Hz transposed-tones, and Dietz et al. (2013) reported no effect of low-pass masking noise on threshold ITDs for 128-Hz SAM tones at 60 dB signal level. Both of these studies suggest that low-pass masking noise is not an absolute requirement when assessing ITD sensitivity conveyed in the envelope of high-frequency modulated sounds, at least for frequencies equal or below 256 Hz and at moderate signal levels.

The ability to exploit ITDs conveyed in the temporal fine structure and envelopes of high-frequency sounds has potential interest for the field of cochlear implantation. Most CI stimulation strategies only convey information about the rectified stimulus envelope of the sounds, discarding the temporal fine structure completely. Recent technological innovations, however, aim to retain information about the temporal fine structure, raising the possibility that temporal information, including from apical, low-frequency regions of the cochlea, might be exploited to enhance binaural hearing in CI listeners. Currently, with electrical stimulation favouring more basal, high-frequency regions of the cochlea, any temporal information conveyed by a CI device is likely to excite auditory nerve fibres projecting to central auditory neurons likely to be less-temporally precise in their neural representation of sound. To this end, a critical factor in determining future CI user outcomes will be the extent to which the relatively low modulation rates at which high-frequency temporal information can be accessed by implant listeners (or indeed, by normal-hearing listeners in which sounds have been high-pass filtered) for binaural processing.

3 The influence of the envelope waveform on binaural tuning of inferior colliculus neurons

3.1 Introduction

The electrophysiology experiments conducted as part of this thesis provided a link between the human behavioural findings for ITD sensitivity and the neurophysiological processes that underlie them. By using complementary stimuli it was hoped that components of the system that enable the mammalian brain to locate accurately sound sources could be further elucidated. Given that CI users must make use of envelope based information due to the nature of CI digital sound processing, by designing experimental stimuli that manipulate this component of an acoustic signal, preferred envelope shapes (in terms of single neuron discrimination thresholds) may provide a basis with which to further develop our understanding of the fundamental processes occurring in the brain of a CI user during signal processing. The psychophysical component of this thesis discussed how sensitivity to ITD varies with frequency and reference IPD as well as how information required to produce a neural code for ITDs must be carried in both the envelope and fine structure of an acoustic waveform. The electrophysiology component explores the possible neural correlates of ITD sensitivity as well as ways in which ITD sensitivity could be enhanced through manipulations of the envelope structure of an acoustic waveform.

Bernstein and Trahiotis (2002) measured the JNDs of ITDs in transposed-tones and SAM tones in order to assess the auditory system's ability to utilise envelope ITDs to lateralise high-frequency sounds. Transposed-tones have a distinct Pause phase between modulation cycles and Attack and Decay segments that are steeper than those of SAM tones of equivalent modulation frequencies (see Appendix i), Figure 3.28A). The use of the transposed-tone resulted in ITD JNDs that were between 100 and 150 μs better than those obtained for SAM

tones at all modulation frequencies. Bernstein and Trahiotis (2009) developed further stimuli that allowed the Pause duration and the Attack and Decay steepness to be independently manipulated creating stimuli that had parameters between those of the SAM tones and tranposed-tones. The results showed that a graded increase in Pause duration an decrease in Attack and Decay steepness results in a graded decrease in ITD JNDs.

Studies that assess the effect of envelope shape on ITD JNDs often have co-varying elements which led Klein Hennig et al. (2011) to conduct their psychophysical experiments in a manner that aimed to clarify further the role of the different envelope components that featured in the existing studies of envelope based binaural sensitivity. This was achieved by manipulating independently the duration of the four distinct envelopes segments; Attack, Sustain, Decay and Pause. The results demonstrated that with a short Attack duration (1.3 ms) and a long Decay duration (18.8ms) ('Damped' stimulus) mean ITD JND, across 6 subjects, was 114 μs. Reversing the signal in the time domain, such that the Attack and Decay durations were switched ('Ramped' stimulus), resulted in a mean ITD JND of just under 400 μs. Existing models of binaural interaction do not clearly explain this as their performance often depends on the spectrum of the pre-processed signal (Bernstein and Trahiotis 1996; Dietz, Ewert et al. 2012), and in the absence of any peripheral adaptation, the spectrum of the two signals is identical (Figure 3.28C). A significant motivator of the neurophysiology experiments presented in this chapter is this large observed difference in ITD sensitivity with the two temporally asymmetric envelope shapes of the Ramped and Damped stimuli.

The aim of the experiments presented in this chapter was to investigate the neural responses to ITDs conveyed in the envelopes of high-frequency tones. Recordings were obtained from single neurons in the inferior colliculus (IC) of the guinea pig (*Cavia Porcellus*). Eighteen different envelope shapes were

created by manipulating the durations of the four segments; Attack, Sustain, Decay and Pause. Each envelope was applied as a modulator to a pure tone carrier with a frequency equal to a specific neuron's characteristic frequency. The response characteristics of four example neurons were analysed in detail and are presented first including a comparison of phase-locking to both a SAM tone and a pseudo-square-wave (PSW) modulated tone, a detailed exploration of ITD-rate-functions to all eighteen parameters and a comparison of the raster plots of the responses to the four envelope shapes: SAM, PSW, Ramped and Damped. The SAM tone and PSW are ideal comparators due to the difference in Attack and Decay segment duration that each has at equal modulation frequency.

Following this, the neuron population characteristics were analysed in detail and are presented with specific reference to the Ramped and Damped parameters including first spike latencies, normalised ITD-rate-functions and ITD JNDs. The final section of the electrophysiological experiments explores the ITD JNDs of the neuron population in response to all eighteen envelope shape parameters. The responses are arranged over eight comparisons including modulation frequency, Pause duration and Attack duration as these have previously been found to influence ITD sensitivity the greatest. We hypothesised that by measuring neural ITD JNDs for single cells within the inferior colliculus it would be possible to begin to establish a neural correlate for the ITD JNDs observed in the existing psychophysical data (Bernstein and Trahiotis 2002; Klein Hennig, Dietz et al. 2011). We predict that in response to increasing Pause duration and decreasing Attack duration specifically, the two factors that have the greatest impact on behavioural thresholds, we would also be able to measure some form of change in the neural responses to the envelope ITDs. Existing electrophysiological data regarding phase-locking to the envelope of high frequency stimulation, whereby above around 250 Hz modulation frequency phase-locking decreases (Dreyer and Delgutte 2006), allowed us to also hypothesise that there would be some decrease in neural ITD sensitivity in

response to increasing the modulation frequency of our own envelope shape parameters.

3.2 Methods

3.2.1 Animal and Surgery

All experiments were performed in accordance with the Animal (Scientific Procedures) Act of 1986 of Great Britain and Northern Ireland. Single-neuron recordings were made from the right inferior colliculus of 15 adult tri-coloured guinea pigs under urethane anaesthesia (20% solution, 1.5 g/kg dosage, IP). Administered subcutaneously at the beginning of each experiment were Rimadyl (50 mg/ml solution, 2 mg constant dosage), a non-steroidal anti-inflammatory drug that also acts as an analgesic, Buprenorphine (0.3 mg/ml solution, 0.05 mg/kg dosage) an opioid with which supplementary doses were administered when an experiment continued beyond 8 hours and Colvasone (2 mg/ml solution, 2 mg/kg dosage), a corticosteroid with an anti-inflammatory action. Lignol, a local anaesthetic containing lignocaine and lidocaine hydrochloride was administered locally before any surgical incision. A tracheal cannula was inserted and core temperature was maintained close to 37°C with a heating blanket (Harvard Apparatus, Kent, UK). Animals were placed in a sound-attenuating chamber (IAC, Winchester, UK) and held in a stereotaxic frame with hollow ear speculae (modified from model 1730, David Kopf Instruments, Tujunga, CA). Before positioning the animal, the tragus was cut to obtain clear access to the tympanic membrane. A craniotomy was performed in order to expose the cortex overlying the IC, and the covering dura was removed. The bullae were vented by insertion of cannulae and sealed with Vaseline in order to equalize air pressure in the middle ears. A parylene-coated tungsten microelectrode [1-2 MΩ; World Precision Instruments, Sarasota, FL] was positioned above the IC and advanced ventrally using a micromanipulator from

outside the recording chamber (model SM-5, Luigs & Neumann, Ratingen, Germany). At the end of each experiment, animals were administered a terminal dose of pentobarbital sodium (200 mg/ml, IP).

3.2.2 Stimuli production and presentation

Sounds were produced using Tucker Davis Technologies (TDT, Alachua, FL) digital signal processing hardware. TDT Brainware, Real Time Processor Visual Design Studio (RPvdsEx), and system III hardware were used to generate the 18 different envelope shapes (48.828-kHz sampling rate). Stimuli were generated and scaled such that their peak voltages were at 5V of the digital-to-analogue converters (DACs). The outputs were attenuated to achieve the desired level for the experiments using PA5 (system III) modules (TDT). Sounds were delivered by Etymotic ER-4S headphones inserted into the hollow ear speculae with the common reference wire separated in order to abolish crosstalk.

3.2.3 Spike Recording

Electrical signals from the electrode were current-amplified by a head-stage (RA16AC, TDT) and further amplified and digitised by a preamplifier (TDT Medusa RA16PA) at a 25 kHz sampling rate. The signal was conducted by a fibre-optic cable to the RX5 base station for filtering (300-3500 Hz). Spike data were passed from the RA16 base station to TDT Brainware and spikes were selected, according to manually-adjusted spike characteristics, to ensure data were analysed from a single neuron.

3.2.4 Neuron Isolation and Characterisation

A binaural search stimulus consisting of ongoing 50 ms presentations of diotic pure tones were used to isolate neurons by eliciting measurable electrical

activity. A first approximation of the neurons response threshold and characteristic frequency (CF) were made. The threshold and CF were confirmed by recording a frequency-level response followed by an ITD sensitivity test run. The confirmed CF of the neuron was used as the carrier frequency for the envelope modulation. If the CF was sufficiently high such as to preclude ITD sensitivity being carried in the signal fine structure, as well as the neuron being readily responsive to ongoing stimulation and sensitive to changes in ITD, the neuron was then presented with the full envelope shape stimulation protocol. A total of 172 IC neurons were isolated with 71 of these found to match the search criteria required for stimulation by the full parameter matrix.

3.2.5 Envelope Shape Stimulation Protocol

Each cycle of the stimulus envelope was constructed from four segments, identical to those used by Klein-Hennig et al. (2011); (1) A Pause segment with zero amplitude (2) an Attack segment identical to the rising portion of a squared sinusoid (3) a sustain segment with a maximum amplitude and (4) a Decay segment identical to the falling portion of a squared sinusoid. The envelope shape parameters are presented in Table 1 and are illustrated in Figure 3.0. The duration of each of the envelope segments is provided as well as the modulation frequency. The stimuli were constructed with a carrier frequency equal to the neuronal characteristic frequency and were 100% modulated. The ITD was applied after multiplying the carrier with the envelope resulting in a full waveform shift. The ‘fine’ recording range was between ±2 ms and consisted of responses to 25 evenly spaced ITDs. The ‘coarse’ recording range extended over a maximum set of ITDs between -8.33 ms and +28.33 ms and 25 intervals, varying as a function of parameter modulation period.

Parameter Number	Pause (ms)	Attack (ms)	Sustain (ms)	Decay (ms)	Period Duration (ms)	Modulation Frequency (Hz)	Referred to as
1	0	12.5	0	12.5	25	40	40 Hz SAM
2	11	1.5	11	1.5	25	40	40 Hz PSW
3	0	1.5	22	1.5	25	40	
4	4	1.5	18	1.5	25	40	
5	8	1.5	14	1.5	25	40	
6	18	1.5	4	1.5	25	40	
7	8	1.5	0.5	15	25	40	Damped
8	8	15	0.5	1.5	25	40	Ramped
9	0	1.5	0	1.5	3	333.3333333	
10	0	2.5	0	2.5	5	200	
11	0	5	0	5	10	100	
12	2	1.5	2	1.5	7	142.8571429	
13	4	1.5	4	1.5	11	90.90909091	
14	0	1.5	4	1.5	7	142.8571429	
15	8	1.5	4	1.5	15	66.66666667	
16	8	1.5	0.5	1.5	11.5	86.95652174	
17	8	15	0.5	15	38.5	25.97402597	
18	8	5	0.5	15	28.5	35.0877193	

Table 1. Parameters of the envelope modulation stimuli used in the experiments.

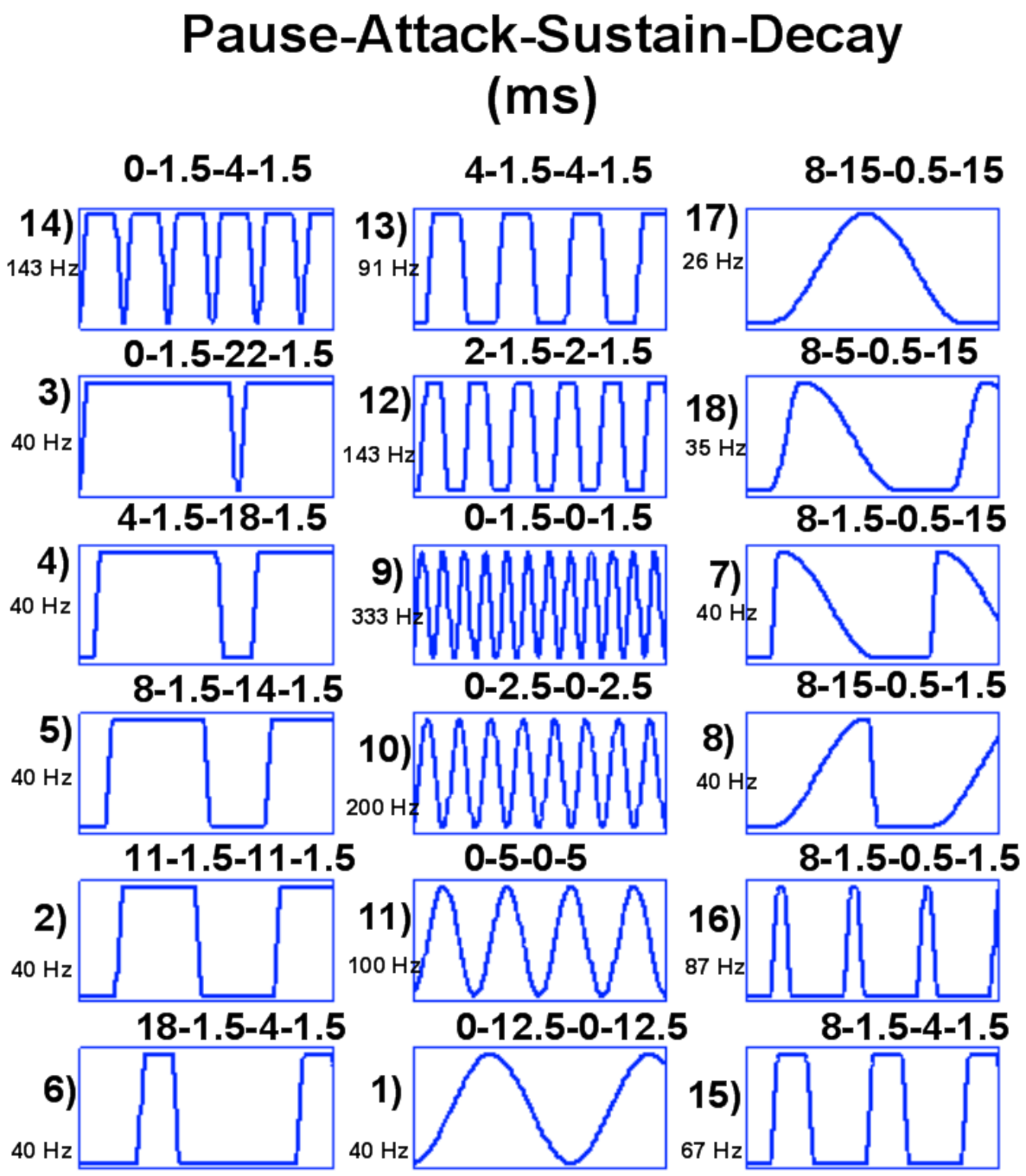


Figure 3.0 1-18) 18 Envelope shapes with respective envelope segment durations and corresponding modulation frequency.

The stimulus duration was always an integer multiple of the cycle duration (starting and stopping in the modulation trough) and as close to 1 s as possible (990-1010 ms). The recording duration was 1100 ms with the first 10 ms and the last 90 ms recording a neurons spontaneous activity and offset response. The inter stimulus interval was at least 300 ms. Responses to 18 different envelope shapes with an average of 38 different ITDs were recorded between 4 and 10 times (mean = 7.8), dependent on how well the isolation of the neuron was maintained. The recording duration lasted between 56 - 140 minutes for each isolated neuron. The stimulus presentation order was chosen so as to record conditions with identical envelope shapes en bloc. Within each block, the ITD was successively increased. Thereafter, for the next run, the presentation order was reversed. Stimuli were presented with a constant maximum amplitude approximately 20 dB above the pure-tone threshold at the CF of the neuron. The maximum output of the system was approximately 110 dB SPL.

3.2.6 Data Analysis

Latency

Neuron latencies were calculated for the 71 neurons from which recordings were made in response to the full stimulation protocol. Three latencies were calculated for each neuron; the shortest latency irrespective of envelope shape and ITD (absolute latency), the latency in response to parameter 7 (1.5 ms Attack and 15 ms Decay – 'Damped') at zero ITD and the latency with respect to parameter 8 (15 ms Attack and 1.5 ms Decay – 'Ramped') at zero ITD. The absolute latency was calculated by first establishing the spontaneous firing rate of the particular neuron. This was measured as the neuron's mean spike rate in the first 10 ms of the overall 1100 ms recording prior to the stimulus onset. Following this, the cell's PSTH (with 1 ms bins) was utilised in order to establish the time point at which the cell's mean spike rate increased above 3

standard deviations of the cell's spontaneous rate. The time between this and stimulus onset (10 ms into the recording) was taken as the absolute latency.

The latency with respect to the Damped and Ramped envelope shapes at zero ITD were calculated by first isolating the spike times in response to these parameter conditions. As a first spike latency was recorded for each neuron in response to each envelope shape for each repeat recording, the time stamps of the first three spikes were extracted in order to provide an accurate representation. It was necessary to gather these multiple first spike latency time stamps across repeats in order to exclude spike times that corresponded to either pre-stimulus responses such as spontaneous activity as well as first spikes responses that were a result of stimulation by envelope modulations beyond the first. This was found to be common amongst multiple repeat recordings from the same neuron to the same envelope shape and inclusion would have reduced the accuracy of the first spike latency calculation. By calculating the mean of the spike time stamps recorded that occurred after the stimulus onset and within one modulation cycle of the stimulus envelope, minus the Pause duration preceding the Attack phase of the stimulus, the resulting time stamp of the cell onset latency with respect to a particular parameter was established.

Raster Plot

For 4 example neurons, raster plots were constructed of the spike responses to parameter 1 (40 Hz SAM tone), parameter 2 (40 Hz PSW), parameter 7 (40 Hz Damped) and parameter 8 (40 Hz Ramped). The time stamp of every spike collected across 8 repetitions of the stimuli were plotted as a function of envelope ITD for the coarse and fine recordings separately.

Phase Locking

For the same 4 example neurons, spike response times were analysed with reference to both Parameter 1 (40 Hz SAM) and Parameter 2 (40 Hz PSW). A frequency distribution plot consisting of 25, 1 ms bins was used to represent the modulation period duration of the 40 Hz SAM and the 40 Hz PSW envelope shapes. All spikes in response to the signal with 0 ITD were allocated to the appropriate time bin. This allowed the spike responses to be assigned to a particular phase of the stimulus envelope and facilitated a comparison between the spread of spike responses to both the 40 Hz SAM and the 40 Hz PSW.

ITD-rate-functions

Sensitivity to ITDs was studied by calculating the mean number of spikes that occurred in response to a particular envelope ITD. Separate functions for both the 'coarse' and 'fine' recordings were made. Any spikes occurring in response to the first 2 modulation cycles were excluded in order to eliminate any onset effects. The 'fine' recordings were used to calculate both the standard separation value (D) and the JND of ITD.

Standard Separation (D)

A discrimination index is required in order to set the criteria from which to extrapolate the ITD JNDs from the present cell population. The index *D*, the Standard Separation (Sakitt 1973) has the following properties that make it ideal as part of a discriminability index in the present context: i) it is independent of assumptions about the underlying response distributions ii) it is independent from the units used for rating task responses iii) it allows the comparison of different neurons from the same task and the same neurons with different tasks iv) it is be affected by changes in variance of responses. D is calculated as:

$$D_{(ITD1, ITD2)} = \frac{R_{(ITD2)} - R_{(ITD1)}}{\sqrt{SD_{(ITD1)} \; SD_{(ITD2)}}}$$

Where $R_{(ITD2)}$ and $R_{(ITD1)}$ are the mean spike responses in response to two different envelope ITDs and $SD_{(ITD1)}$ and $SD_{(ITD2)}$ are the standard deviations of the respective response distributions. A random rating would give D = 0 and perfect discrimination would produce an infinite D. The neuron population characteristics (Section 3.3.2) are defined using the criterion of D = 1, 1.5 and 2 while the stringent criteria of D = 2 is used for the population analysis with D = 1 also applied for the analysis of SAM tone modulation frequency and Pause duration i) (Section 3.3.3).

JND of ITD

In order to calculate the ITD JND for a particular cell and envelope shape, the ITD relative to two recorded values at which the selected D value is reached is calculated. Within the physiological range for the guinea pig, the smallest distance between two ITDs that reached the D criteria was taken as the ITD JND. This point will occur between two test ITDs (for the *fine* recordings, responses were recorded at 0.167 ms ITD intervals) and so linear interpolation is used to calculate the ITD JND.

3.3 Results

3.3.1 Single Neuron Characteristics

A detailed characterisation of 4 neurons from 4 different guinea pigs is initially presented in order to demonstrate the range of envelope parameters and possible

response characteristics prior to presenting the population analysis of 71 neurons in response to all 18 parameters. Each of the single neuron characterisations is divided in to seven sections i) Frequency Response Area (FRA) ii) pooled PSTH iii) neuron latency analysis iv) SAM tone vs PSW phase locking comparison v) SAM tone vs PSW rate ITD function comparison vi) detailed exploration of ITD rate functions for all 18 parameters and vii) raster plots of responses to 4 envelope shapes (SAM, PSW, Ramped and Damped).

Neuron 203 (6) in Figure 3.1 had a CF of 4.2 kHz as depicted by the FRA response plot (Figure 3.1A). A pooled PSTH for this neuron (Figure 3.1B) displays the spike responses for all 18 envelope shapes that were presented in 1s stimulations. This pooled response makes it possible to observe the presence of an onset effect due to the additive effect of the pooled response and manifests here as a large number of spikes at the beginning of the PSTH. As this overall onset effect was common and would vary when observed in response to each envelope shape, for this study the rate ITD functions were assembled after discarding any spike that occurred in the first 60 ms or last 100 ms of the 1.1 s recording time. This would discard spontaneous firing occurring in the first 10 ms and last 90 ms of the recording as well as spike responses to the first 50 ms and final 10 ms of stimulation in order to ensure the removal of any onset or offset activity from the rate level functions which may be drown out the response to the ongoing signal. Figure 3.1C represents the first 70 ms of the PSTH for neuron 203 (6) and was utilised to calculate the latency of this neuron as 24 ms as described in the methods section. The dashed line depicts the stimulus onset and the solid line depicts the first point at which the spike rate increased above three standard deviations of the spontaneous rate as calculated from the first 10 ms of the PSTH. Figure 3.1D displays the distribution frequency of spikes in response to parameter 1, a 40 Hz SAM tone and parameter 2, a 40 Hz pseudo square wave (PSW) at 0 ITD. The 25 ms period of both envelope shapes was divided in to 25 1 ms time bins. In response to the

PSW, responses peak at a higher level and at an earlier phase as measured from the start of the Attack phase. Figure 3.1E and 3.1F depict the rate-level-functions in response to the 40 Hz SAM and the 40 Hz PSW respectively, with the coarse and fine recordings represented in part i) and ii) respectively. The shaded region in the coarse recording of the rate ITD function represents the ±2 ms region in which the part ii) fine recording of the rate ITD function is recorded from. The shaded region in the fine recording represents the ± 330 μs physiological range of guinea pig ITDs (Sterbing, Hartung et al. 2003). The spike rate at each ITD was recorded 8 times with the mean and the smoothed three-point-average function depicted by the solid and dashed line respectively. The error bars represent one standard deviation from the mean spike-rate of the eight repeats. The 40 Hz SAM and the 40 Hz PSW were chosen as they provide typical examples of some of the main differences in envelope shape being explored. The 40 Hz SAM has a long 12.5 ms Attack-phase and a 0 ms Pause-phase while the 40 Hz PSW has a very short, 1.5 ms Attack-phase and a long 11 ms Pause-phase. Figure 3.1E and 3.1F shows how these differences in envelope shape result in both a difference in spike rate at any given ITD as well as a difference in the rate of change in response levels across ITDs. The difference between the maximum and minimum firing rate for ITDs within the physiological range (Figure 3.1E/Fii)) is 6sp/s for the 40 Hz SAM and 29sp/s for the 40 Hz PSW suggesting a greater sensitivity to changes in ITD when the signal envelope shape is a PSW rather than sinusoidal. In order to assess the statistical significance of the ITD sensitivity for 40 Hz SAM and the 40 Hz PSW, the change in neural firing rates across ITDs within the physiological range were evaluated by a non-parametric ANOVA. The result confirmed the greater change in mean spike rate for the 40 Hz SAM ($p = < 0.0001$) compared to the 40 Hz PSW ($p = 0.0146$).

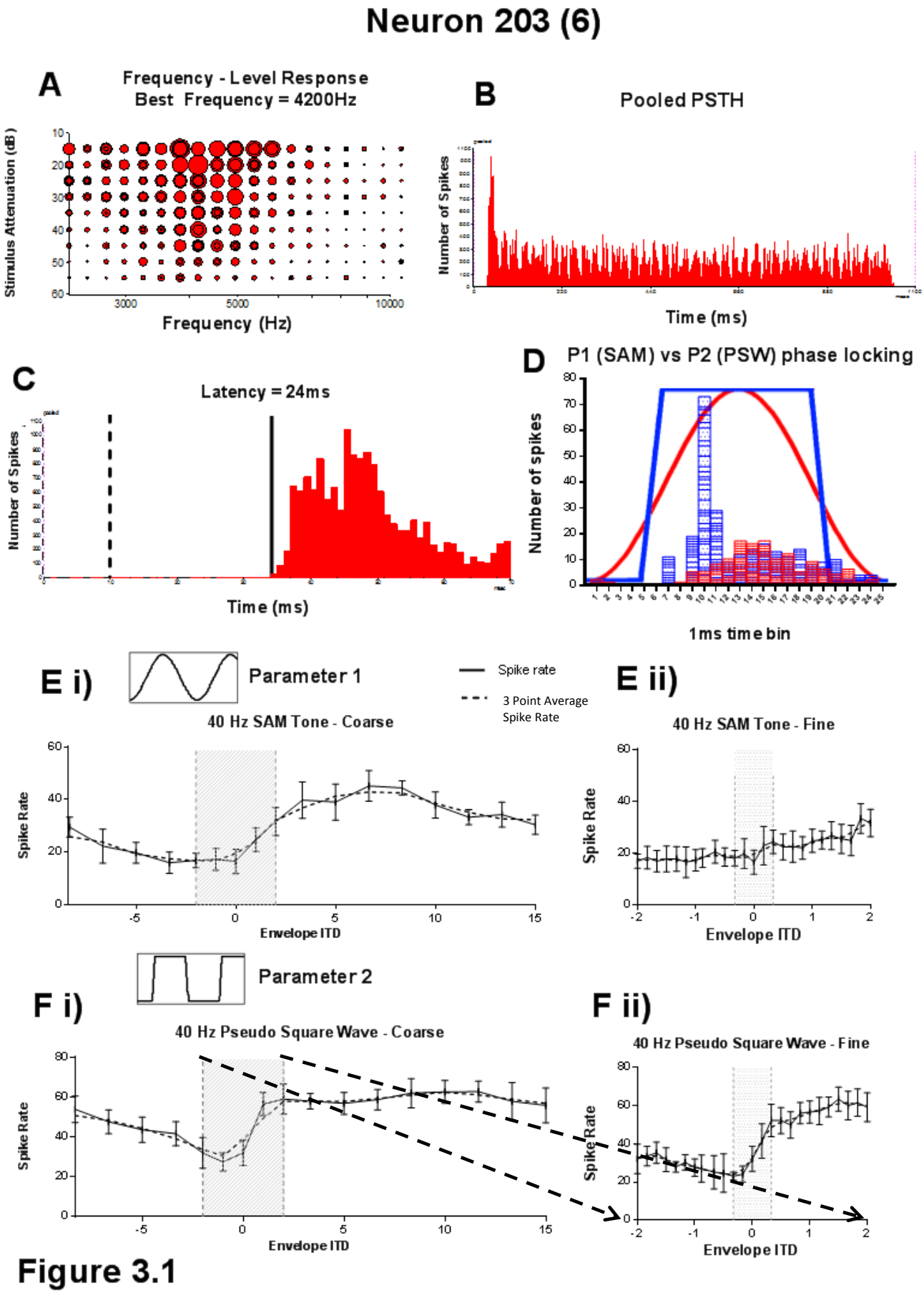

Neuron 203 (6)
A
Frequency - Level Response
Best Frequency = 4200Hz
Stimulus Attenuation (dB)
Frequency (Hz)
B
Pooled PSTH
Number of Spikes
Time (ms)
C
Latency = 24ms
Number of Spikes
Time (ms)
D
P1 (SAM) vs P2 (PSW) phase locking
Number of spikes
1ms time bin
E i)
Parameter 1
Spike rate
3 Point Average Spike Rate
40 Hz SAM Tone - Coarse
Spike Rate
Envelope ITD
E ii)
40 Hz SAM Tone - Fine
Spike Rate
Envelope ITD
F i)
Parameter 2
40 Hz Pseudo Square Wave - Coarse
Spike Rate
Envelope ITD
F ii)
40 Hz Pseudo Square Wave - Fine
Spike Rate
Envelope ITD

Figure 3.1

Figure 3.1G-J are raster plots of the responses of neuron 203 (6) to 4 envelope shape parameters, each with a 40 Hz frequency modulation 1) SAM 2) PSW 3) Damped and 4) Ramped. Each plot displays both the coarse and the fine recordings with the ITD rate function illustrated below the relevant raster plot. Each raster plot depicts all of the spikes collected over 8 repeats. Phase-locking to the envelope of the stimulus is evident in all plots represented by the increased number of spikes occurring at regular intervals across the duration of the 1 s stimulation protocol consisting of 40 modulation cycles. The phase-locking is seen to be ITD sensitive and each plot illustrates the neurons ITD tuning characteristics with either an increase or decrease in spike responses at particular envelope ITDs relative to others. Phase-locking appears most robust in response to the PSW (Figure 3.1G - parameter 2) and the Damped envelope shapes (Figure 3.1I - parameter 7). Of the 8 raster plots presented in Figure 3.1G-J, only the fine recordings of responses to the SAM tone (Figure 3.1H - parameter 1) and the Ramped envelope shape (Figure 3.1J - parameter 8) do not show a distinct decrease in both the neuron responses and phase locking robustness in a region close to 0 ITD.

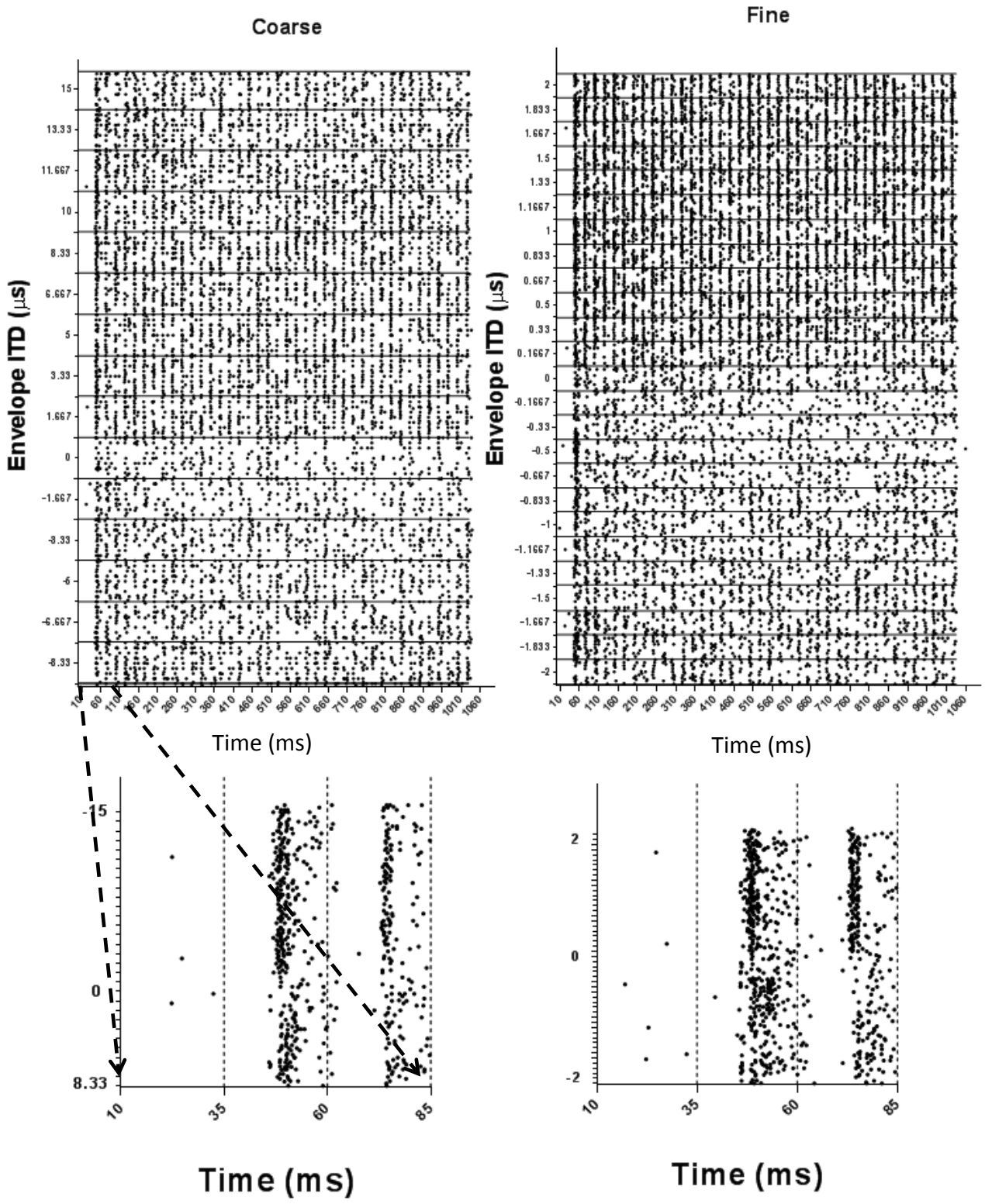

203 (6): 40 Hz PSW
Coarse
Fine
Envelope ITD (μs)
Envelope ITD (μs)
Time (ms)
Time (ms)
Time (ms)
Time (ms)

Figure 3.1 G

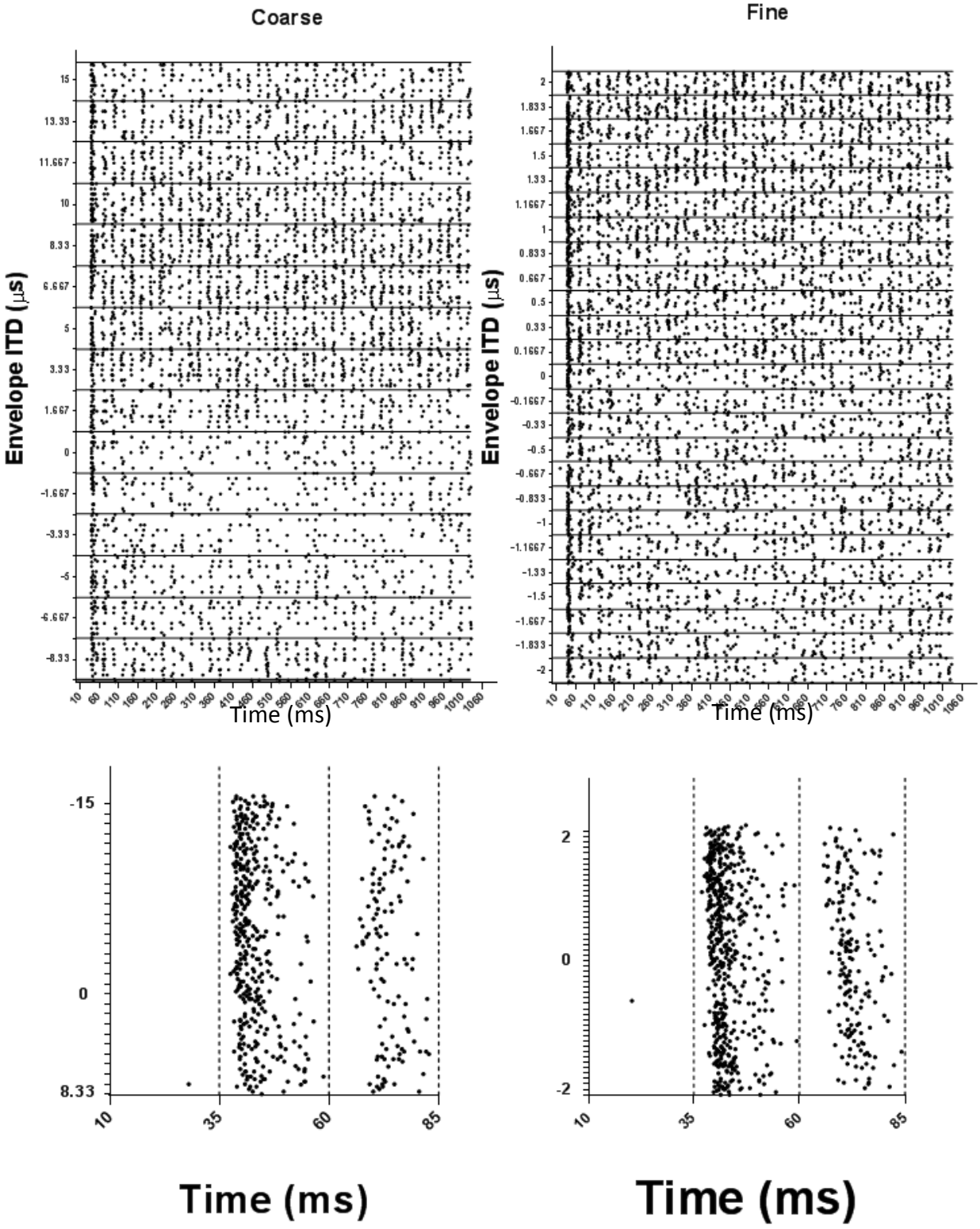

203 (6): 40 Hz SAM Tone
Coarse
Fine
Envelope ITD (µs)
Envelope ITD (µs)
Time (ms)
Time (ms)
Time (ms)
Time (ms)

Figure 3.1 H

203 (6): 40 Hz 'Damped'

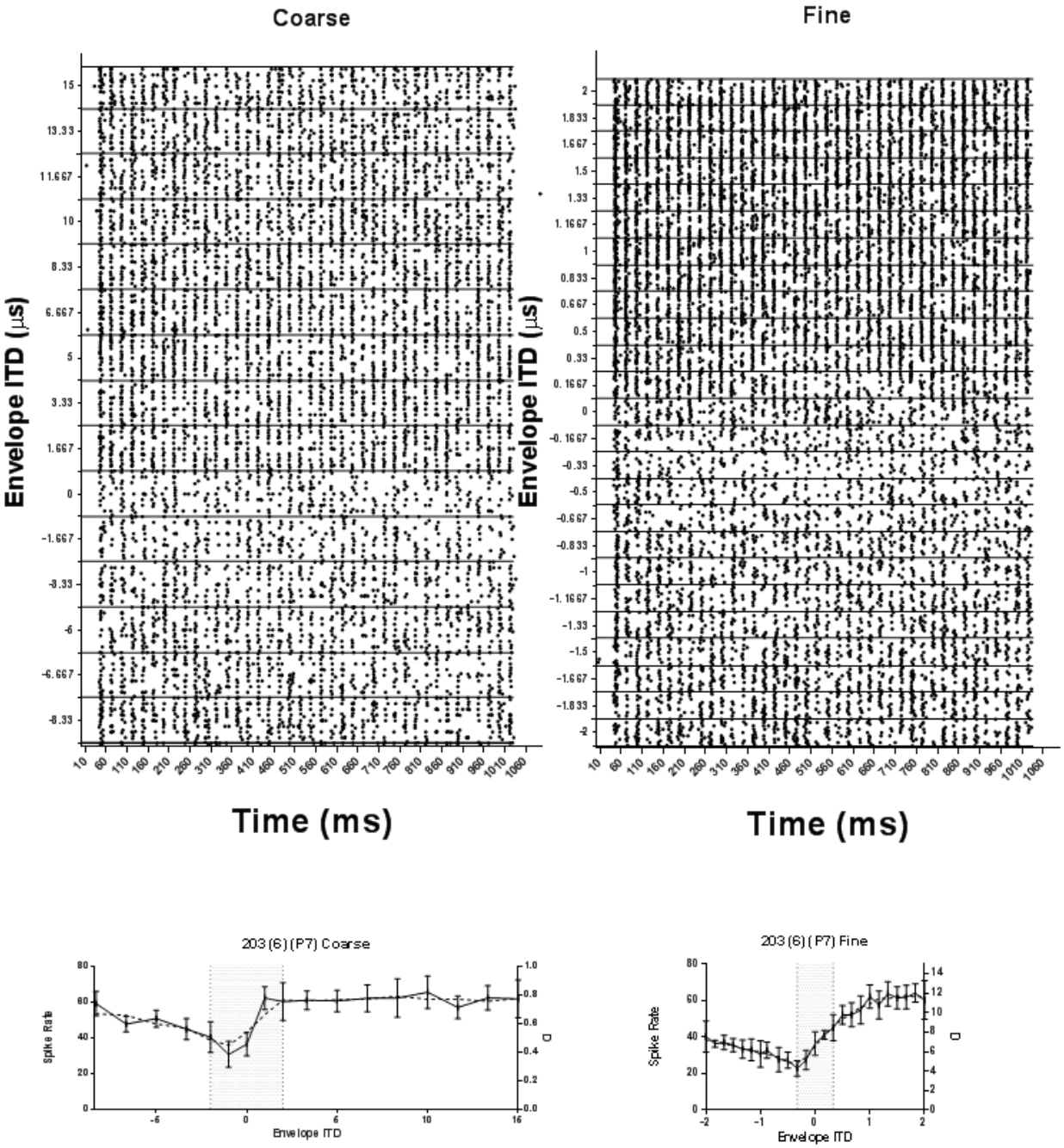


Figure 3.1 I

203 (6): 40 Hz 'Damped'

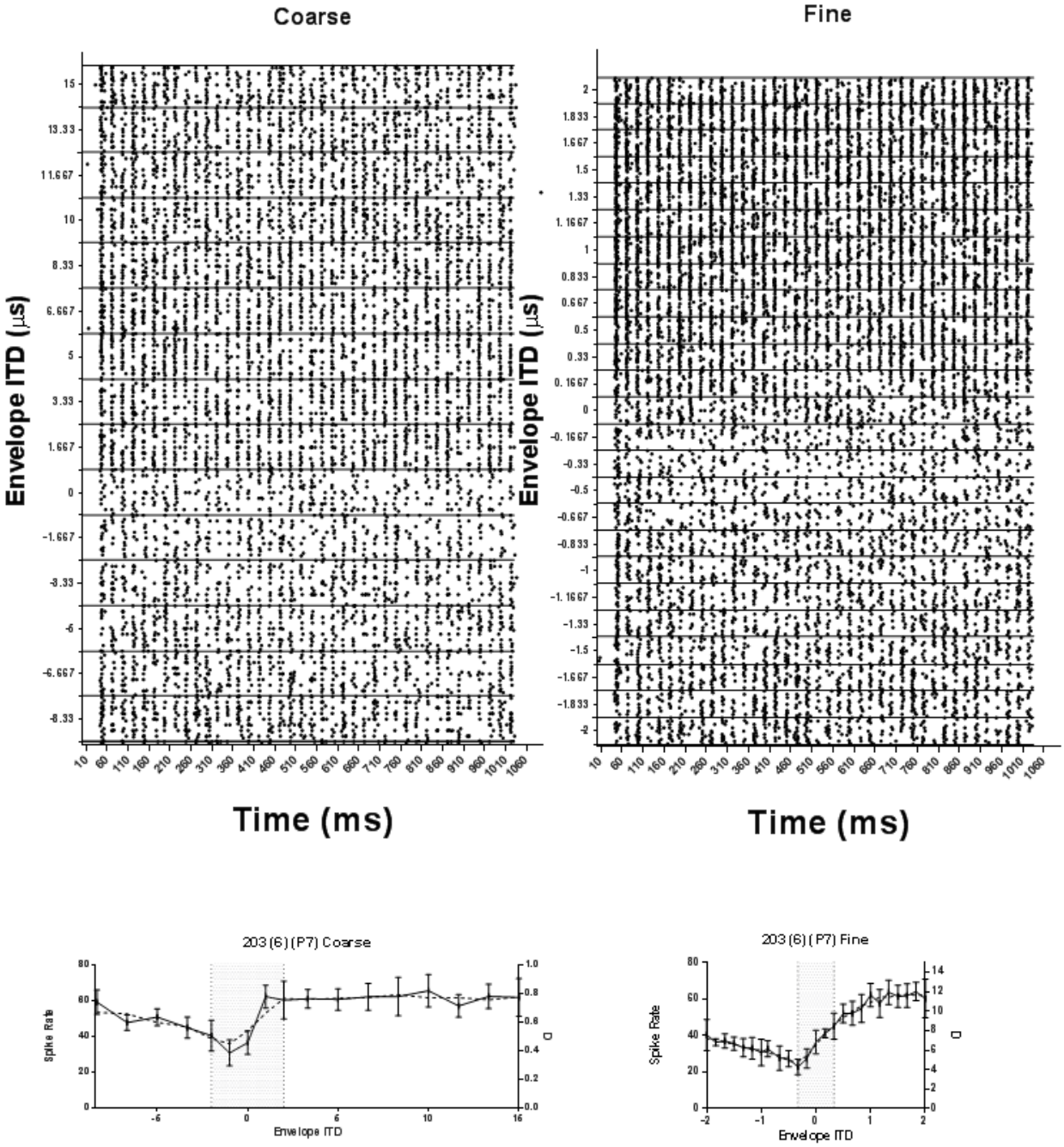


Figure 3.1 I

203 (6): 40 Hz 'Ramped'

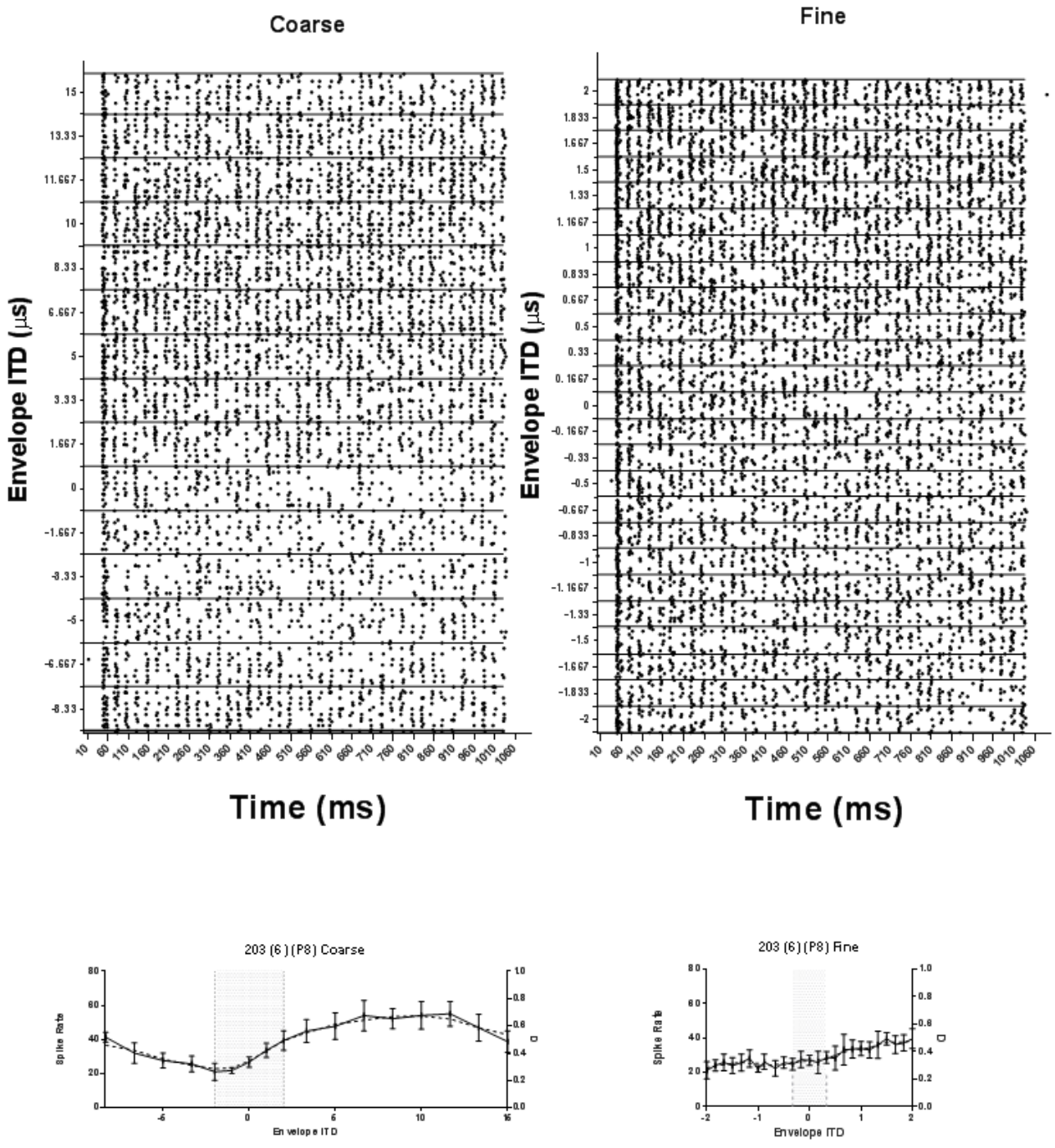


Figure 3.1 J

Figure 3.1. Neuron 203 (6) A) rate vs level response, BF = 4200 Hz, threshold = -65 dB re maximum system output. B) Pooled PSTH to 1.1s recording of 18 envelope shape parameters. C) The first 70ms of the PSTH for neuron 203 (6), used to calculate the latency at 24 ms. The dashed line depicts the stimulus onset and the solid line depicts the first response onset. D) Distribution frequency of spikes in response to a 40 Hz SAM tone and a 40 Hz PSW, at 0 ITD. E) and F) rate-level-functions in response to the SAM tone and PSW respectively. Coarse and fine recordings represented in part i) and ii) respectively. The shaded region in the coarse recording of the rate ITD function represents the ±2 ms region in which the part ii) fine recording of the rate ITD function is taken from. The shaded region in the fine recording represents the ± 330 µs physiological range of guinea pig ITDs. The error bars represent one standard deviation from the mean spike-rate of eight repeats. (G-J) Raster plots for recordings of 40 Hz PSW, 40 Hz SAM, Damped and Ramped responses respectively. Each band on the ordinate shows spike times in response to 8 repeat presentations of the corresponding envelope ITD. Left panels: Coarse recording between -8.33ms and 15ms ITD. Right panels: Fine recording between -2ms and 2ms ITD. G) and H) lower panels: Response to first 85 ms of stimulus. Dashed lines = period duration of first 3 complete envelope modulations. I) and J) lower panels: rate-level-functions in response to Damped and Ramped stimuli respectively. Coarse and fine recordings represented left and right respectively. The shaded region in the coarse recording represents the ±2 ms region in which the fine recording of the rate ITD function is taken from. The shaded region in the fine recording represents the ± 330 µs physiological range of guinea pig ITDs. The error bars represent one standard deviation from the mean spike-rate of eight repeats

The 18 envelope shape parameters have been divided in to 8 separate plots of rate ITD functions. This facilitates the comparison and observation of how manipulating each of the four envelope segments (Attack, Sustain, Decay, Pause) influences the rate ITD functions of a particular IC neuron. Figure 3.2A-D and 3.2E-H display the 8 plots of rate ITD functions for neuron 203 (6). Figure 3.2A displays the influence of SAM tone modulation frequency on rate ITD functions. The SAM tone modulation frequencies used were 40 Hz (parameter 1), 100 Hz (parameter 11), 200 Hz (parameter 10) and 333 Hz (parameter 9). In order to assess if the mean spike rates (from 8 repeat recordings) in response to different ITDs within the physiological range are significantly different from one another, a non-parametric ANOVA test was utilised with ($p = < 0.05$) indicating significance. While there was no significant difference in mean spike rate for ITDs within the envelope of a 333 Hz SAM tone ($p = 0.2825$) there was for the lower frequency modulation rates of 200 Hz ($p = 0.0332$), 100 Hz ($p = 0.0004$) and 40 Hz ($p = 0.0146$).

Figure 3.2B: The influence of PSW Pause duration on rate ITD functions. The Attack, Sustain and Decay duration were all kept constant at 1.5 ms, 4 ms and 1.5 ms respectively. The Pause durations used were 0 ms (parameter 14 – 143 Hz modulation frequency), 4 ms (parameter 13 – 91 Hz modulation frequency), 8 ms (parameter 15 – 67 Hz modulation frequency) and 18 ms (parameter 6 – 40 Hz modulation frequency). Only the rate ITD function for a PSW with a 0 ms Pause is not strongly modulated. The non-parametric ANOVA test confirmed this ($p = 0.0719$ for parameter 14 and $p = < 0.0001$ for parameters 13, 15 and 6).

Figure 3.2C: The influence of Attack duration on rate ITD functions. The Sustain, Decay and Pause duration were all kept constant at 0.5 ms, 15 ms and 8 ms respectively. The Attack durations used were 1.5 ms (parameter 7 – 40 Hz modulation frequency), 5 ms (parameter 18 – 35 Hz modulation frequency) and 15 ms (parameter 17 – 26 Hz modulation frequency). The non-parametric

ANOVA test indicated that the shortest Attack phase of 1.5 ms for the Damped envelope shape results in the greatest difference in spike rates across ITDs ($p = < 0.0001$) with the 5 ms of parameter 18 still significant ($p = 0.0016$) but the 15 ms Attack phase of parameter 17 reduced the difference in spike rates across ITDs such that it is not significant ($p = 0.3662$) as well as reducing the overall spike rate compared to parameters 7 and 18.

Figure 3.2D: The influence of the temporally asymmetric envelope shapes of the Damped and the Ramped envelope shapes on rate ITD functions. The Sustain and Pause durations were kept constant at 0.5 ms and 8 ms respectively. The Attack durations used were 1.5 ms (Damped) and 15 ms (Ramped) with the same durations used for the Decay phase but for the alternative parameter. The non-parametric ANOVA test indicated that the Attack phase of 1.5 ms and Decay of 15 ms for the Damped envelope shape resulted in a highly significant difference in spike rates across ITDs ($p = < 0.0001$) whereas the temporally reversed the Ramped envelope shape produced a difference in spike rates that was not significant ($p = 0.4458$), also reducing the overall spike rate compared to parameters 7.

Figure 3.3E: The influence of PSW modulation frequency on rate ITD functions. The rate ITD functions are generally steeper than for the SAM tone modulation frequencies in Figure 3.2A. For a modulation frequency of 333 Hz (parameter 9), the PSW stimulus is identical to the SAM. The remaining PSW modulation frequencies are 143 Hz (parameter 12), 91 Hz (parameter 13) and 40 Hz (parameter 2). Increasing the modulation rate from 40 Hz to 333 Hz reduced the response rate of the neuron. The non-parametric ANOVA test indicated that while there was no significant difference in mean spike rate for ITDs within the envelope of a 333 Hz PSW tone ($p = 0.2825$) there was for the lower modulation frequency rates of 143 Hz ($p = 0.0005$), 91 Hz ($p = < 0.0001$) and 40 Hz ($p = < 0.0001$).

Figure 3.3F: The influence of PSW duty cycle duration at a constant 40 Hz modulation frequency. The Attack and Decay durations were held constant at 1.5 ms whilst the sustain duration was decreased and the Pause duration increased proportionally. The Pause durations were 0 ms (parameter 3), 4 ms (parameter 4), 8 ms (parameter 5), 11 ms (parameter 2) and 18 ms (parameter 6). The response rate of the neuron increases with a reduction in duty cycle i.e. with an increase in Pause duration and decrease in sustain duration. The non-parametric ANOVA test indicated no significant difference in mean spike rate for the longest duty cycle of parameter 3 ($p = 0.9678$) but there was for the shorter duty cycles of parameter 4 ($p = 0.0003$), parameter 5 ($p = 0.0013$) parameter 6 ($p = < 0.0001$).and parameter 2 ($p = < 0.0001$).

Figure 3.3G: The influence of increasing the Pause duration from 0 ms to 8 ms between lobes of a SAM tone. The SAM tone frequency was decreased from 40 Hz (parameter 1) to 26 Hz (parameter 17) by the inclusion of the 8 ms Pause. The response rate of the neuron increases with the inclusion of the Pause duration. The non-parametric ANOVA test indicated that although the response rate is increased with the presence of the Pause, this had the effect of reducing the difference in mean spike rates across ITD for parameter 17 ($p = 0.3662$) in comparison to parameter 1 ($p = 0.0146$).

Figure 3.3H: The influence of increasing the Pause duration from 0 ms to 8 ms once again, but in this case between lobes of two envelope shapes with very short duty cycles of 3 ms (parameter 9) and 3.5 ms (parameter 16). Increasing the Pause time resulted in a decrease in modulation frequency from 333 Hz (parameter 9) to 87 Hz (parameter 16). The response rate of the neuron again increases with the inclusion of the Pause duration as well as greatly increasing the difference in mean spike rates across ITD as seen in the non-parametric ANOVA test results for parameter 9 ($p = 0.2825$) compared to parameter 16 ($p = < 0.0001$).

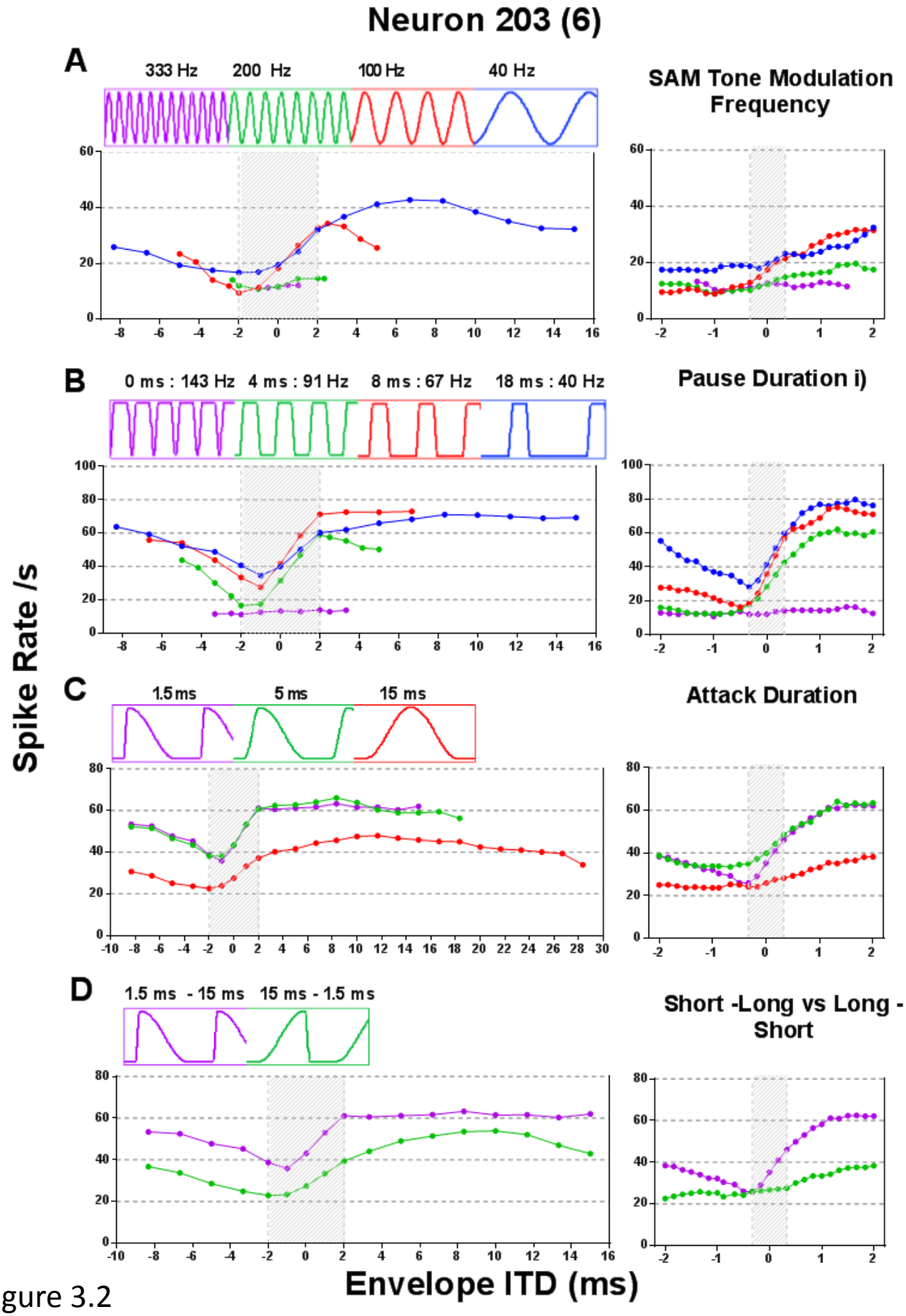

Neuron 203 (6)
A
333 Hz
200 Hz
100 Hz
40 Hz
SAM Tone Modulation Frequency
B
0 ms : 143 Hz
4 ms : 91 Hz
8 ms : 67 Hz
18 ms : 40 Hz
Pause Duration i)
C
1.5 ms
5 ms
15 ms
Attack Duration
D
1.5 ms - 15 ms
15 ms - 1.5 ms
Short -Long vs Long - Short
Spike Rate /s
Envelope ITD (ms)

Figure 3.2

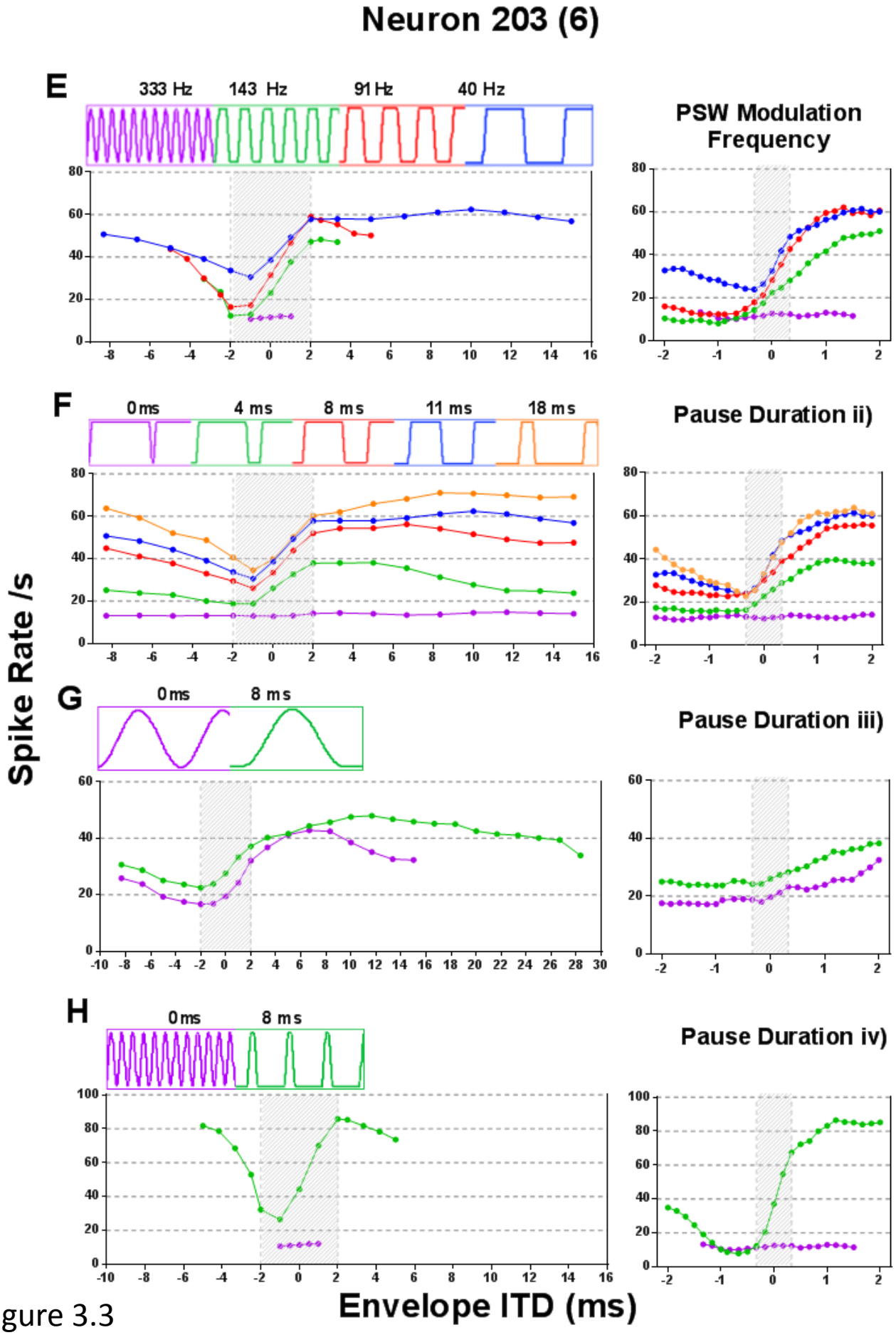

Neuron 203 (6)
E
333 Hz
143 Hz
91 Hz
40 Hz
PSW Modulation Frequency
F
0ms
4 ms
8 ms
11 ms
18 ms
Pause Duration ii)
G
0ms
8 ms
Pause Duration iii)
H
0ms
8 ms
Pause Duration iv)
Spike Rate /s
Envelope ITD (ms)

Figure 3.3

Figure 3.2 the influence of A) 40, 100, 200 and 333 Hz SAM tone modulation frequency on rate ITD functions. B) PSW 0, 4, 8 and 18 ms Pause duration on rate ITD functions. The Attack, Sustain and Decay duration were all kept constant at 1.5 ms, 4 ms and 1.5 ms respectively. C): 1.5, 5 and 15 ms Attack duration on rate ITD functions. The sustain, Decay and Pause duration were all kept constant at 0.5 ms, 15 ms and 8 ms respectively. D) the temporally asymmetric envelope shapes of the Damped (1.5 ms Attack/15 ms Decay) and Ramped (15 ms Attack/1.5 ms Decay) envelopes on rate ITD functions. The sustain and Pause durations were kept constant at 0.5 ms, 15 ms and 8 ms respectively.

Figure 3.3 the influence of E) 40, 91, 143, 333 Hz PSW modulation frequency on rate ITD functions. F) PSW duty cycle duration at a constant 40 Hz modulation frequency. The Attack and Decay durations were held constant at 1.5 ms whilst the sustain duration was decreased and the Pause duration increased to 0, 4, 8, 11 and 18 ms maintaining a total modulation period of 25 ms. G) 0 and 8 ms Pause duration between lobes of a SAM tone. The SAM tone frequency was decreased from 40 Hz to 26 Hz. H) 0 and 8 ms Pause duration between lobes of two envelope shapes with 3 ms duty cycles. The modulation frequency was decreased from 333 Hz to 87.

A further three neurons (204 (5), 205 (18) and 206 (13)) have been characterised in detail in order to display the range of responses measured from the ICs of four different guinea pigs. Neuron 204 (5) in Figure 3.4 had the highest CF of these 4 neurons at 10.4 kHz (Figure 3.4A) and displays a strong onset response followed by the lowest mean spike rate across 8 repeats of these 4 neurons (Figure 3.4B). Figure 3.4C displays the latency of this neuron at 23 ms. Figure 3.4D illustrates

that this neuron responds very little in response to a 40 Hz SAM tone while having a clear response to the PSW across a narrow phase range. Figure 3.1E and 3.1F display how in response to parameter 1 this neuron has very little ITD rate modulation with the occurrence of few spikes while parameter 2 elicits a larger response as well as a clear response modulation across ITD. A non-parametric ANOVA confirmed this with a significant difference in mean spike rate for parameter 2 ($p = < 0.0001$) compared to parameter 1 ($p = 0.8069$).

Raster plots (Figure 3.4G-J) show the responses of neuron 204 (5) to 4 envelope shape parameters, each with a 40 Hz frequency modulation 1) SAM 2) PSW 3) Damped and 4) Ramped. Each plot displays both the coarse and the fine recordings with the ITD rate function illustrated below the relevant raster plot. Each raster plot depicts all of the spikes collected over 8 repeats. Phase-locking to the envelope of the stimulus is evident for the 40 Hz PSW and the Damped stimulus but not to the 40 Hz SAM tone and the Ramped stimulus. The phase-locking where present is ITD sensitive. Phase-locking appears most robust at large positive ITDs for the PSW and the Damped envelope shapes. The response rate is very low in response to the SAM tone and the Ramped stimulus.

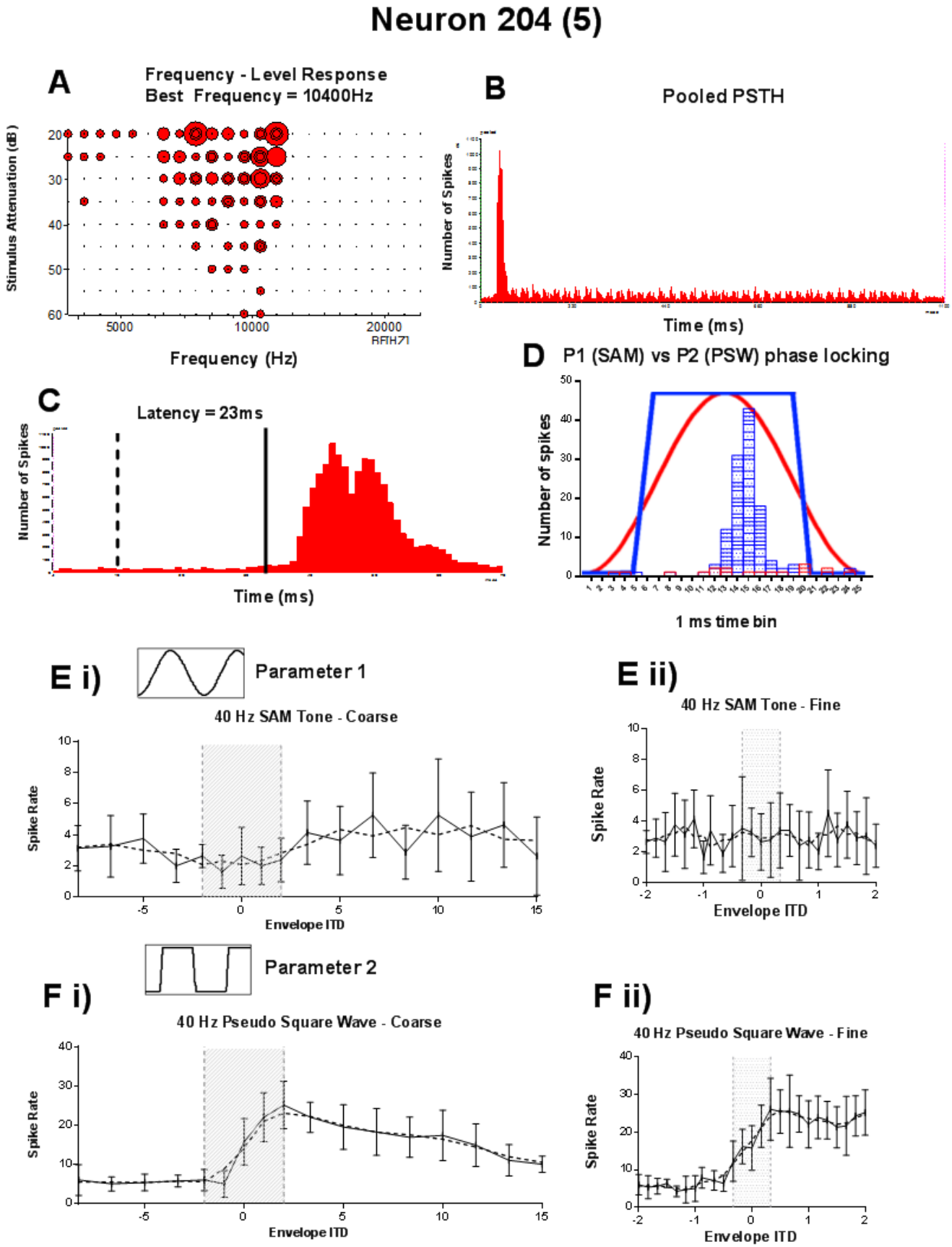
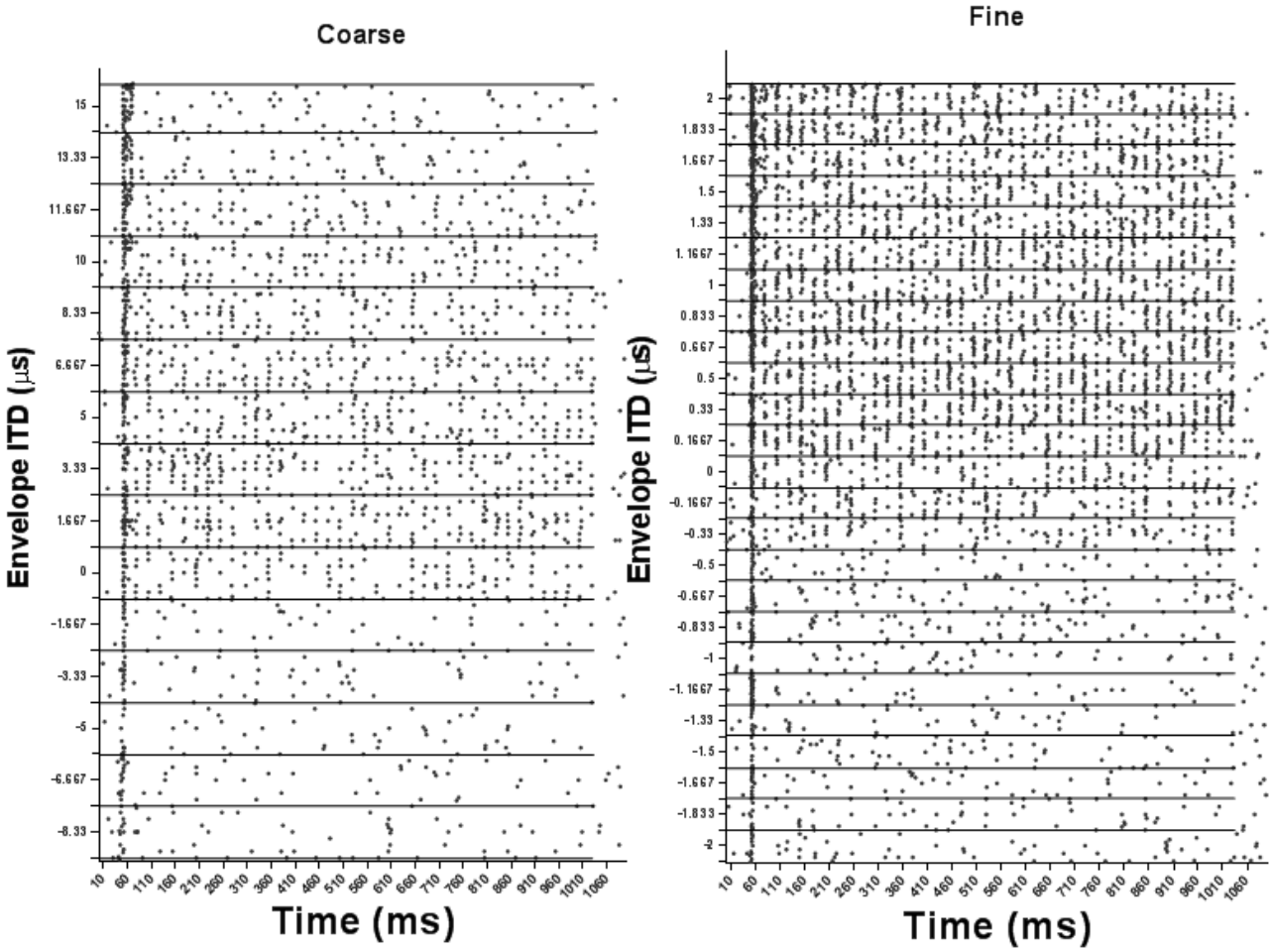
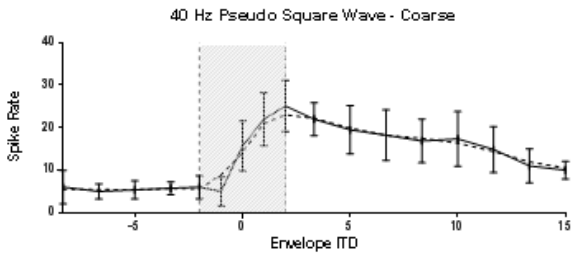

Neuron 204 (5)
A
Frequency - Level Response
Best Frequency = 10400Hz
Stimulus Attenuation (dB)
Frequency (Hz)
B
Pooled PSTH
Number of Spikes
Time (ms)
C
Latency = 23ms
Number of Spikes
Time (ms)
D
P1 (SAM) vs P2 (PSW) phase locking
Number of spikes
1 ms time bin
E i)
Parameter 1
40 Hz SAM Tone - Coarse
Spike Rate
Envelope ITD
E ii)
40 Hz SAM Tone - Fine
Spike Rate
Envelope ITD
F i)
Parameter 2
40 Hz Pseudo Square Wave - Coarse
Spike Rate
Envelope ITD
F ii)
40 Hz Pseudo Square Wave - Fine
Spike Rate
Envelope ITD

Figure 3.4

204 (5): 40 Hz PSW





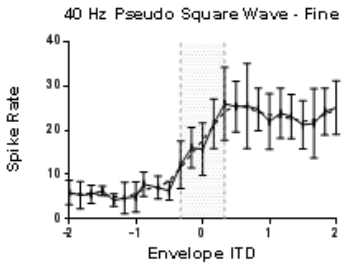
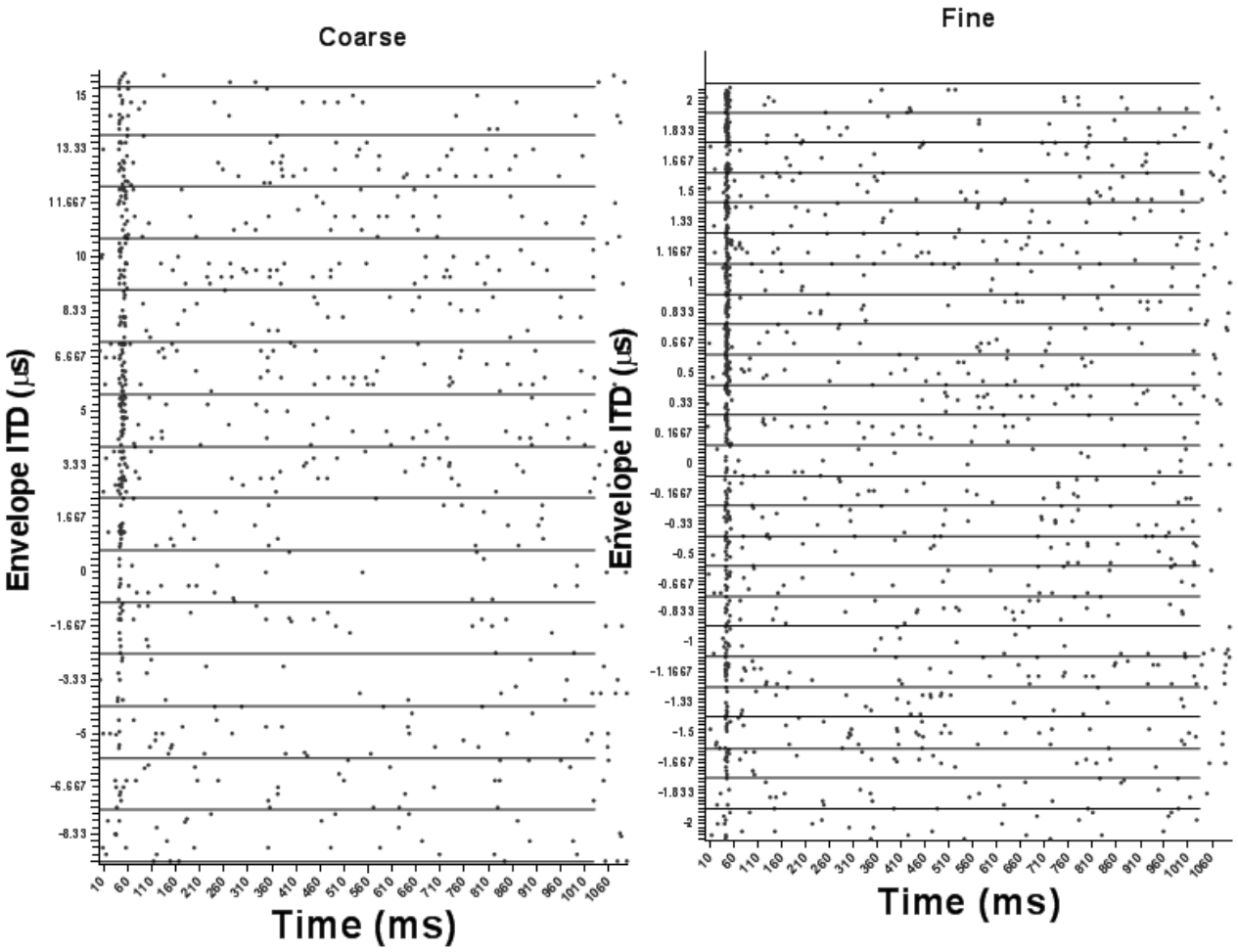
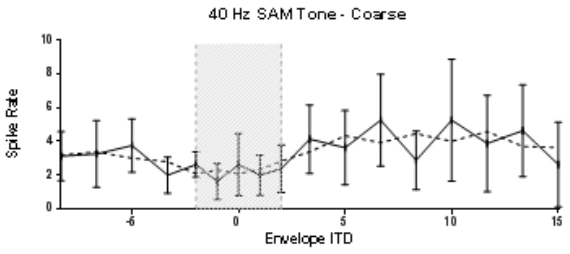


Figure 3.4 G

204 (5): 40 Hz SAM Tone





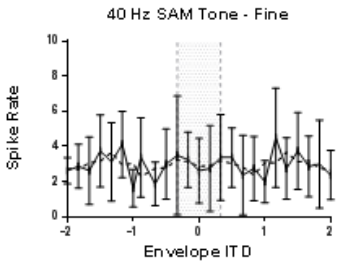
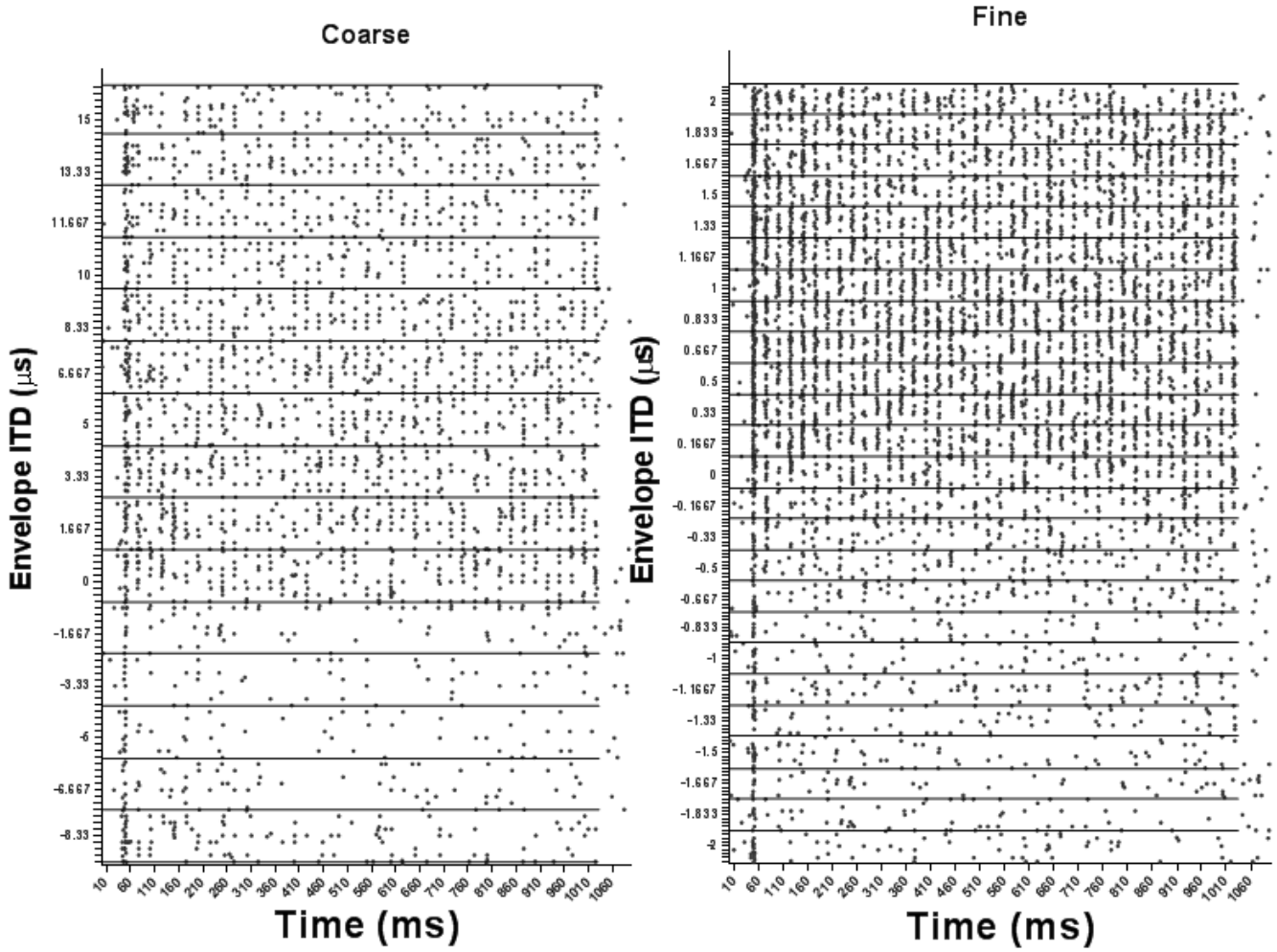


Figure 3.4 H

204 (5): 40 Hz 'Damped'



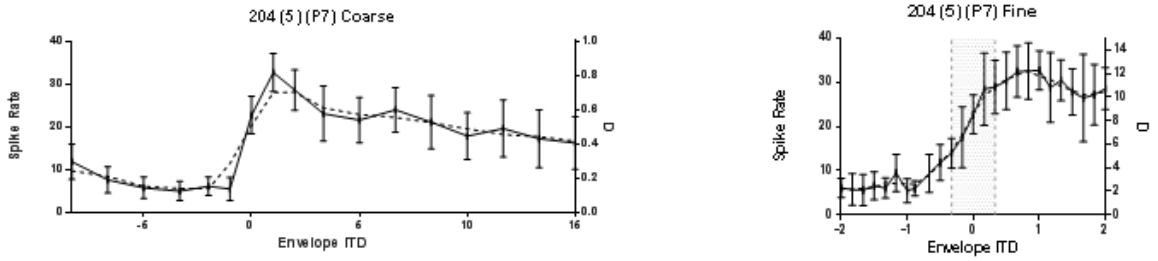


Figure 3.4 l

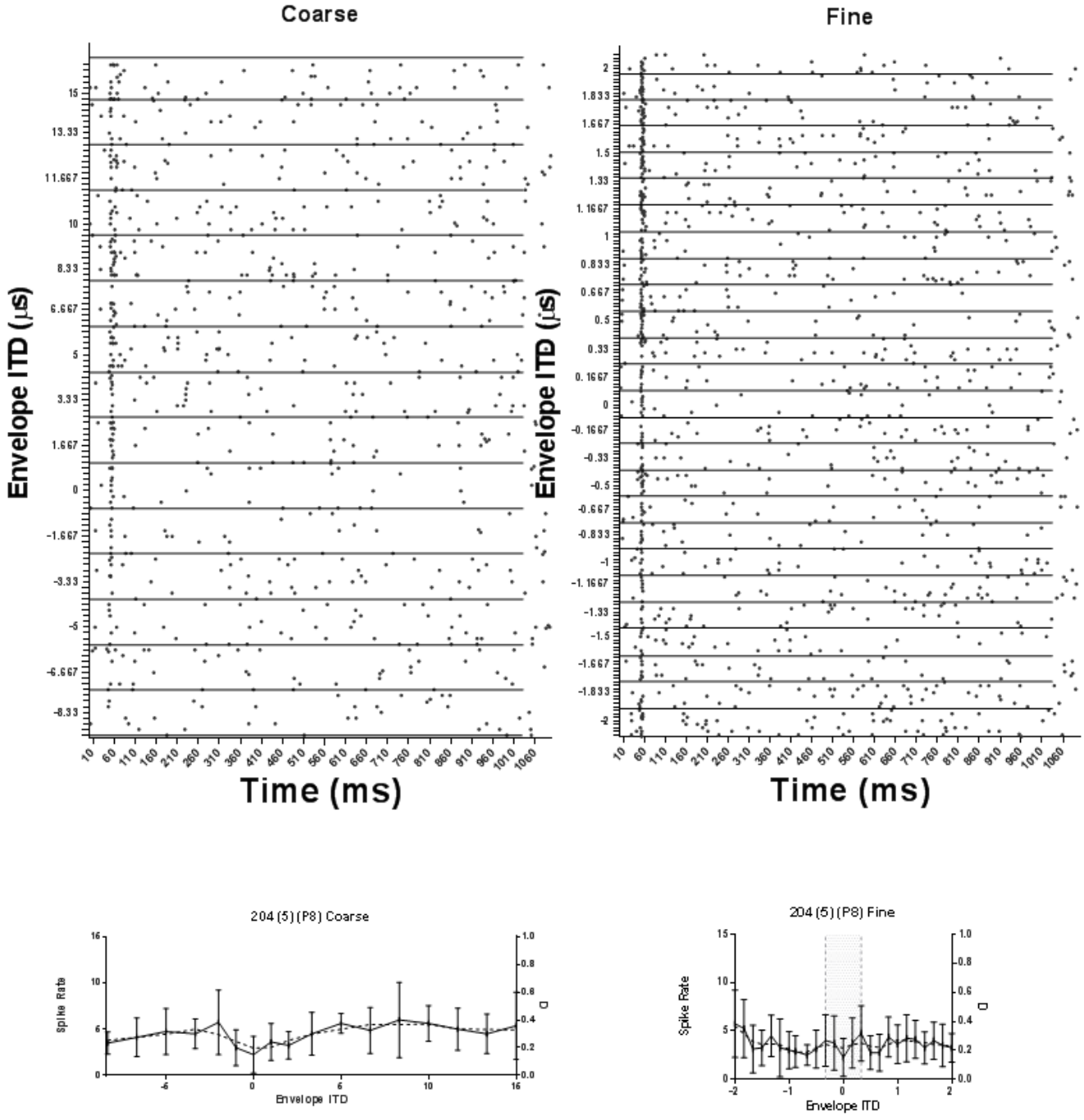


Figure 3.4 J

Figure 3.4. Neuron 204 (5) A) rate vs level response, BF = 10400 Hz, threshold = -55 dB re maximum system output. B-J) same format as Figure 3.1 excluding onset raster plots.

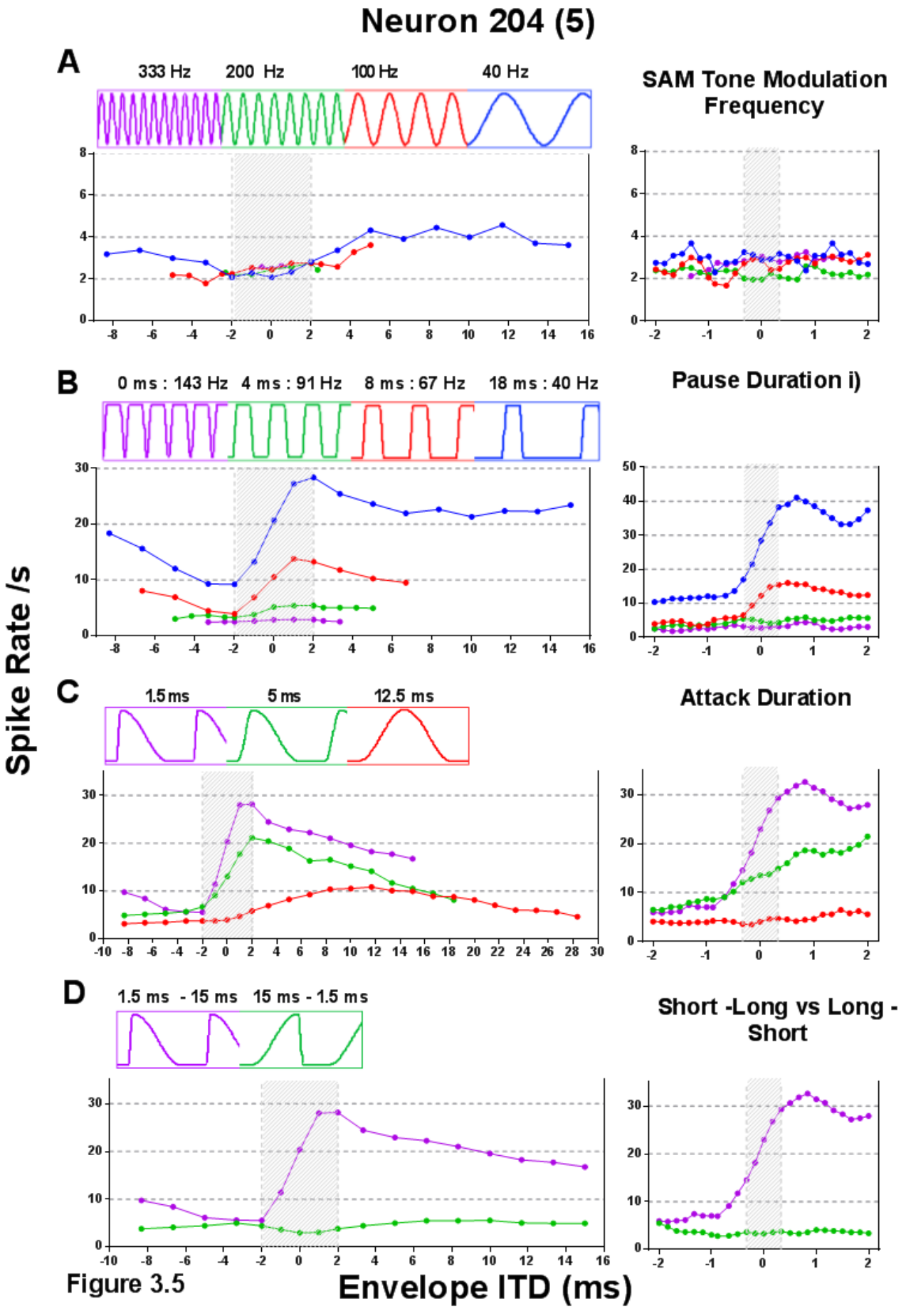


Figure 3.5

Figure 3.5 A-D) same format as Figure 3.2

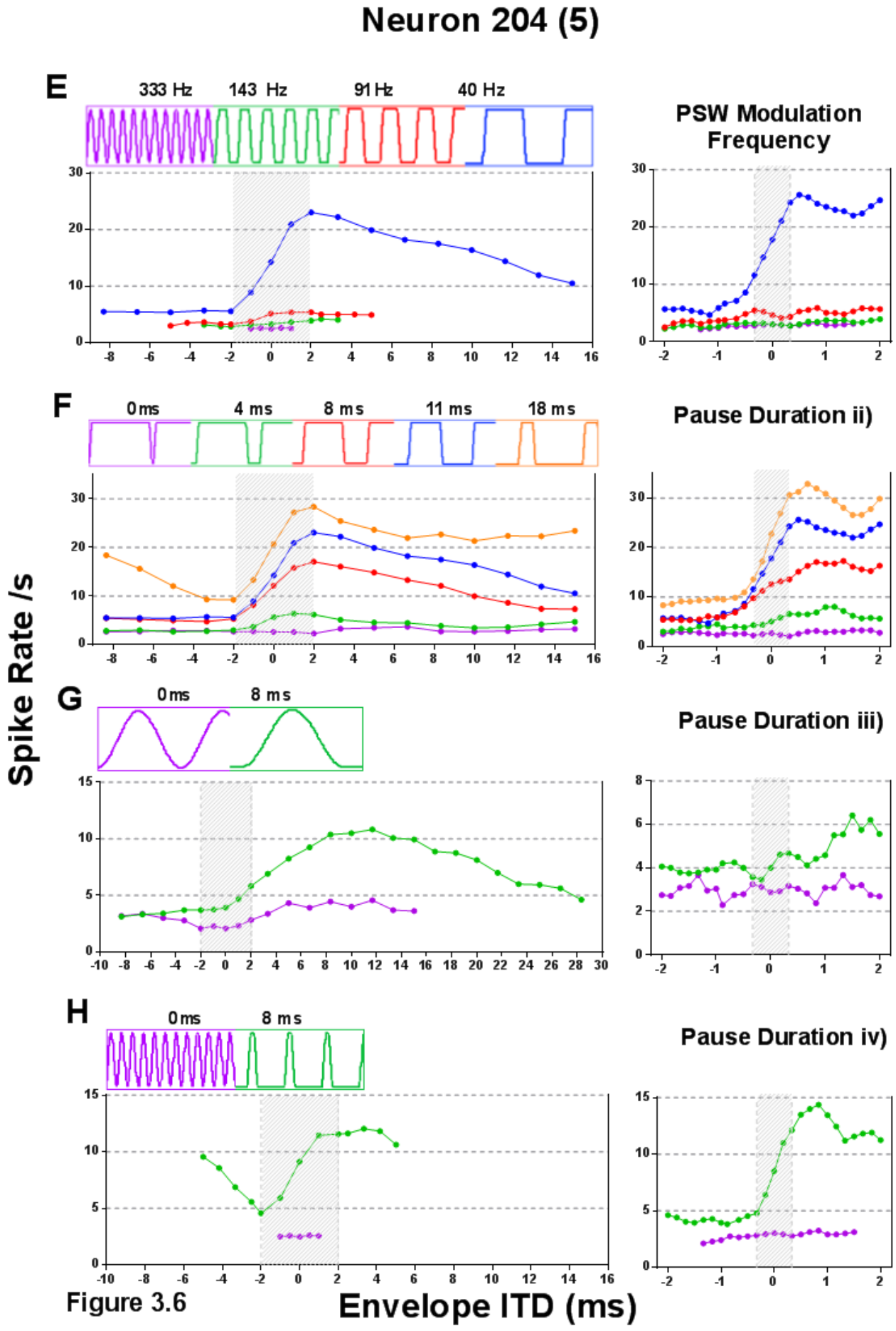


Figure 3.6

Figure 3.6 E-H) same format as Figure 3.3.

Neuron 205 (18) in Figure 3.7 had a CF of 5650 Hz (Figure 3.7A) with a clear onset response (Figure 3.7B) and a first spike latency of 23 ms (Figure 3.7C). Figure 3.7D illustrates that this neuron responds at both the peak and the trough of both the 40 Hz SAM tone and the PSW but to a greater degree for the latter. In response to parameter 1 this neuron has very little ITD rate modulation (Figure 3.7E) while parameter 2 elicits a larger response as well as a modulation across ITD (Figure 3.7F). A non-parametric ANOVA confirmed this with a significant difference in mean spike rate for parameter 2 ($p = 0.0055$) compared to parameter 1 ($p = 0.2579$).

Neuron 206 (13) in Figure 3.10 had a CF of 3150 Hz (Figure 3.10A), a strong onset response (Figure 3.10B) and first spike latency of 25 ms (Figure 3.10C). Figure 3.10D illustrates that this neuron responds maximally at the peaks of the 40 Hz SAM tone and to a greater degree in response to the peak of the PSW. In response to parameter 1 and parameter 2 this neuron does show some degree of ITD tuning across the range of ITDs assessed (Figure 3.10E and 3.10F) although with higher spike rates for parameter 2. A non-parametric ANOVA established that neither parameter 1 ($p = 0.5274$) nor parameter 2 ($p = 0.6561$) elicited a significant difference in mean spike rate across the physiological range of ITDs.

The detailed spike response comparisons for neurons 204 (5), 205 (18) and 206 (13) can be seen in Figures 3.5, .6, .8, .9, .11 and .12. Distinctive observations include the lack of sensitivity neuron 204 (5) has for any of the SAM tone frequencies (Figure 3.5A) and that it is highly unresponsive to the ‘Ramped’ envelope shape of parameter 8 but highly sensitive to changes in ITD for the temporally reversed ‘Damped’ envelope shape of parameter 7 (Figure 3.5D). Increasing the modulation frequency of the PSW reduces the response rate of neuron 205 (18) (Figure 3.9E) while decreasing the duty cycle of the PSW increases the response rate with the shortest duty cycle producing a very large sensitivity to ITDs (Figure 3.9F). Neuron 206 (13) responds similarly to the four

largest duty cycles of the PSW followed by a decrease in response rate but an increase in ITD sensitivity for the shortest duty cycle (Figure 3.12F). All four neurons exhibit an increase in response rate and ITD sensitivity when comparing the Ramped and Damped envelope shapes (Figure 3.2-3.11D). Only neuron 203 (6) is sensitivity to ITDs conveyed by the 143 Hz PSW with none of these 4 neurons sensitive to ITDs conveyed by the 333 Hz PSW (Figure 3.3-3.12E). Amongst the envelope shapes employed, increasing the Pause duration almost invariably results in one or both of an increase in response rate and an increase in ITD sensitivity as demonstrated in Figure 3.3-3.12F.

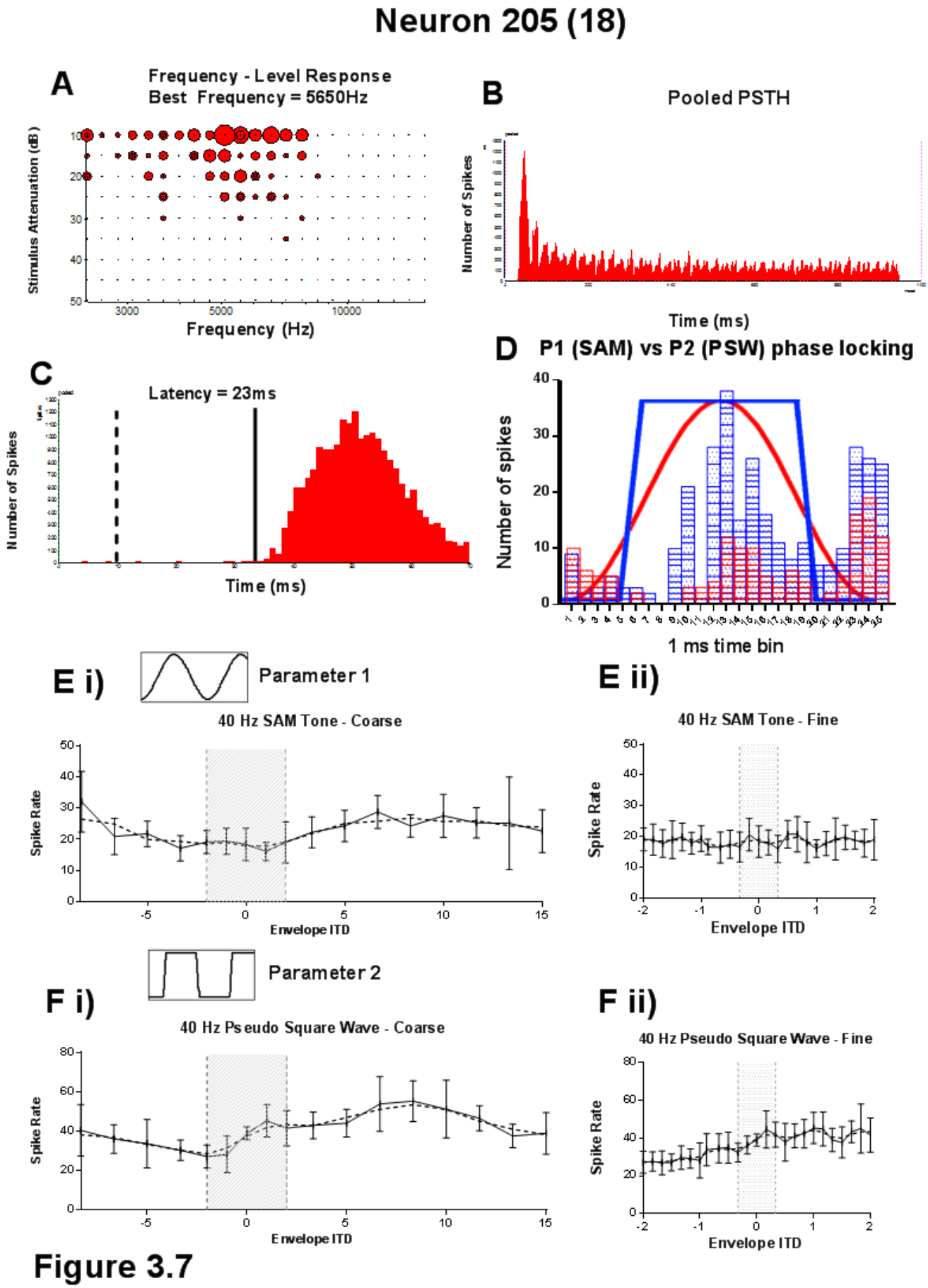

Neuron 205 (18)
A
Frequency - Level Response
Best Frequency = 5650Hz
Stimulus Attenuation (dB)
Frequency (Hz)
B
Pooled PSTH
Number of Spikes
Time (ms)
C
Latency = 23ms
Number of Spikes
Time (ms)
D
P1 (SAM) vs P2 (PSW) phase locking
Number of spikes
1 ms time bin
E i)
Parameter 1
40 Hz SAM Tone - Coarse
Spike Rate
Envelope ITD
E ii)
40 Hz SAM Tone - Fine
Spike Rate
Envelope ITD
Parameter 2
F i)
40 Hz Pseudo Square Wave - Coarse
Spike Rate
Envelope ITD
F ii)
40 Hz Pseudo Square Wave - Fine
Spike Rate
Envelope ITD

Figure 3.7

205 (18): 40 Hz PSW

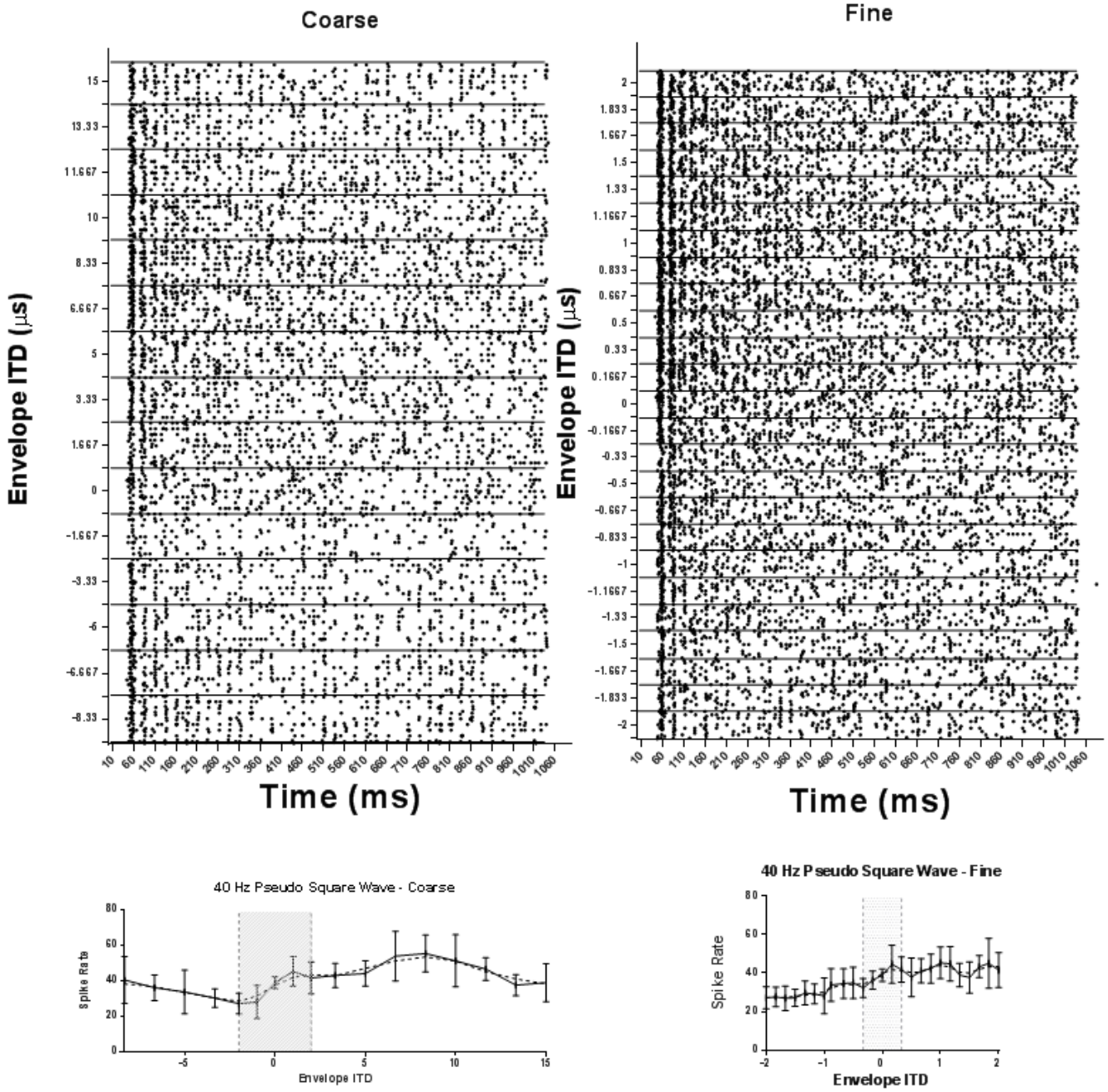


Figure 3.7 G

205 (18): 40 Hz SAM Tone

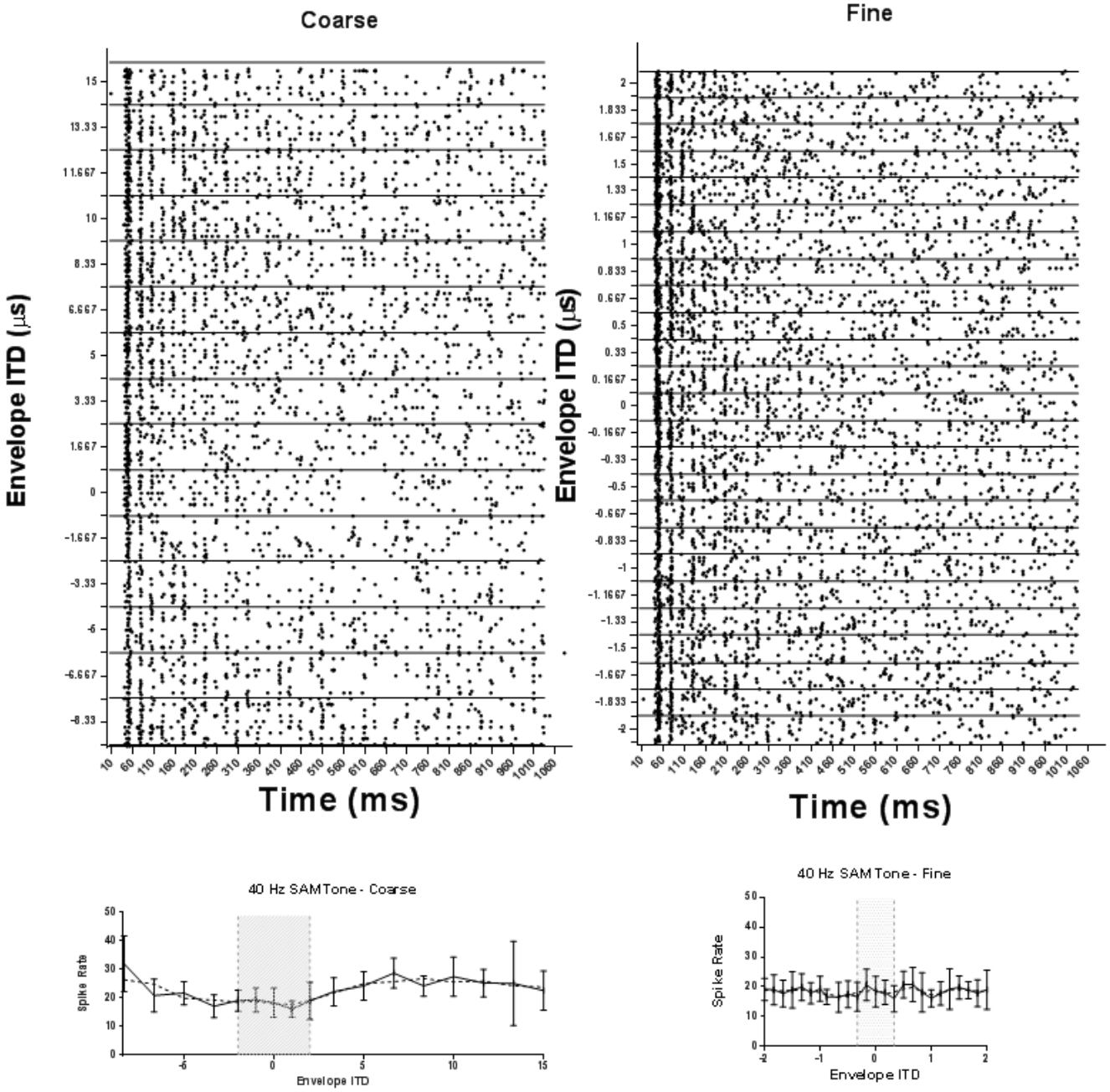
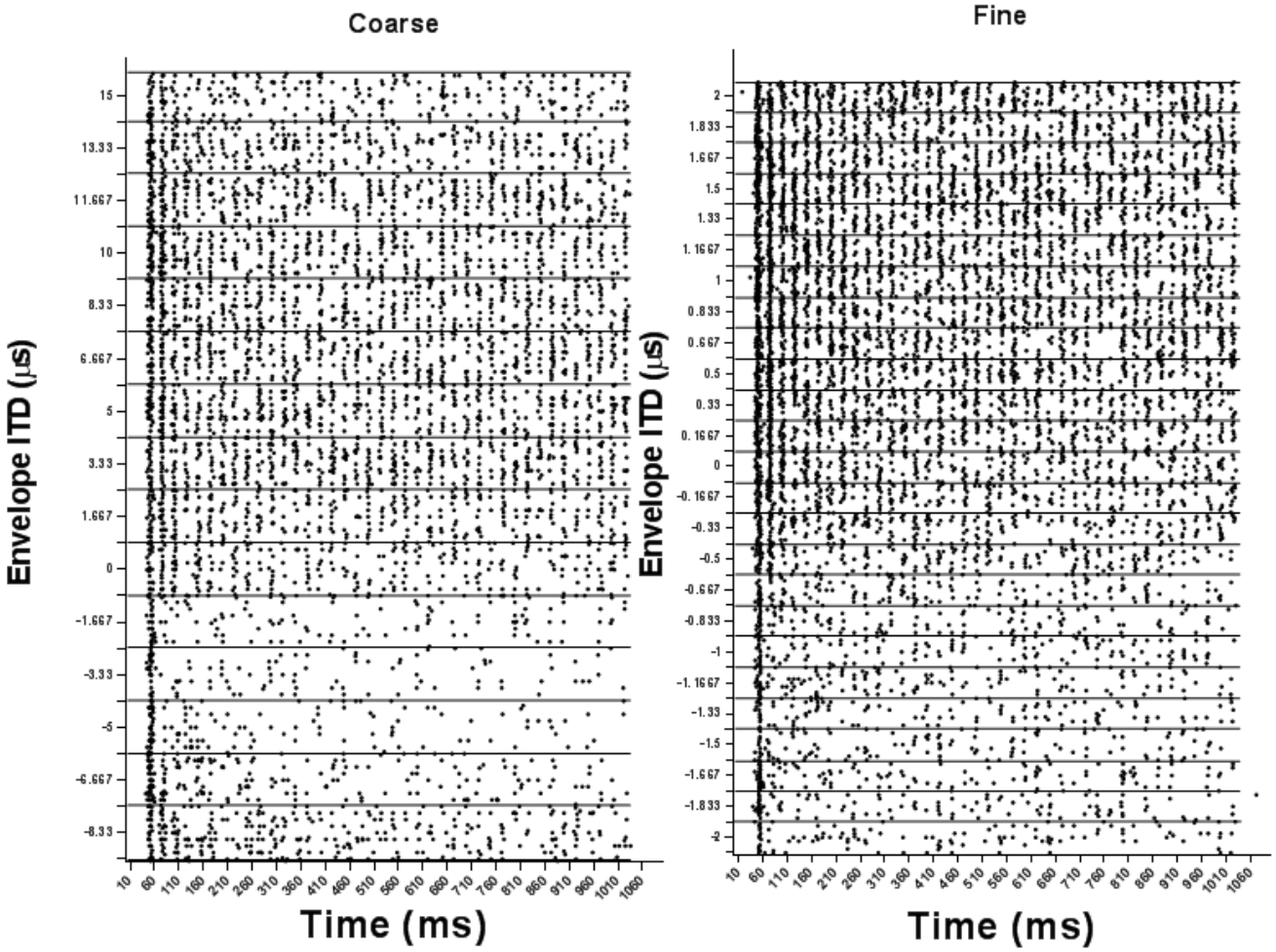
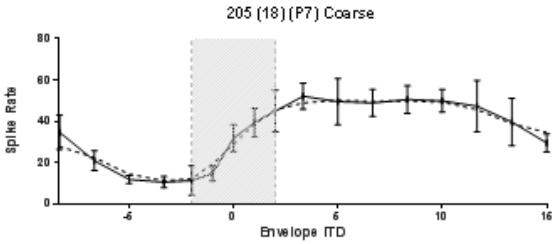


Figure 3.7 H

205 (18): 40 Hz 'Damped'





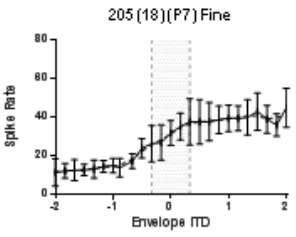


Figure 3.7 I

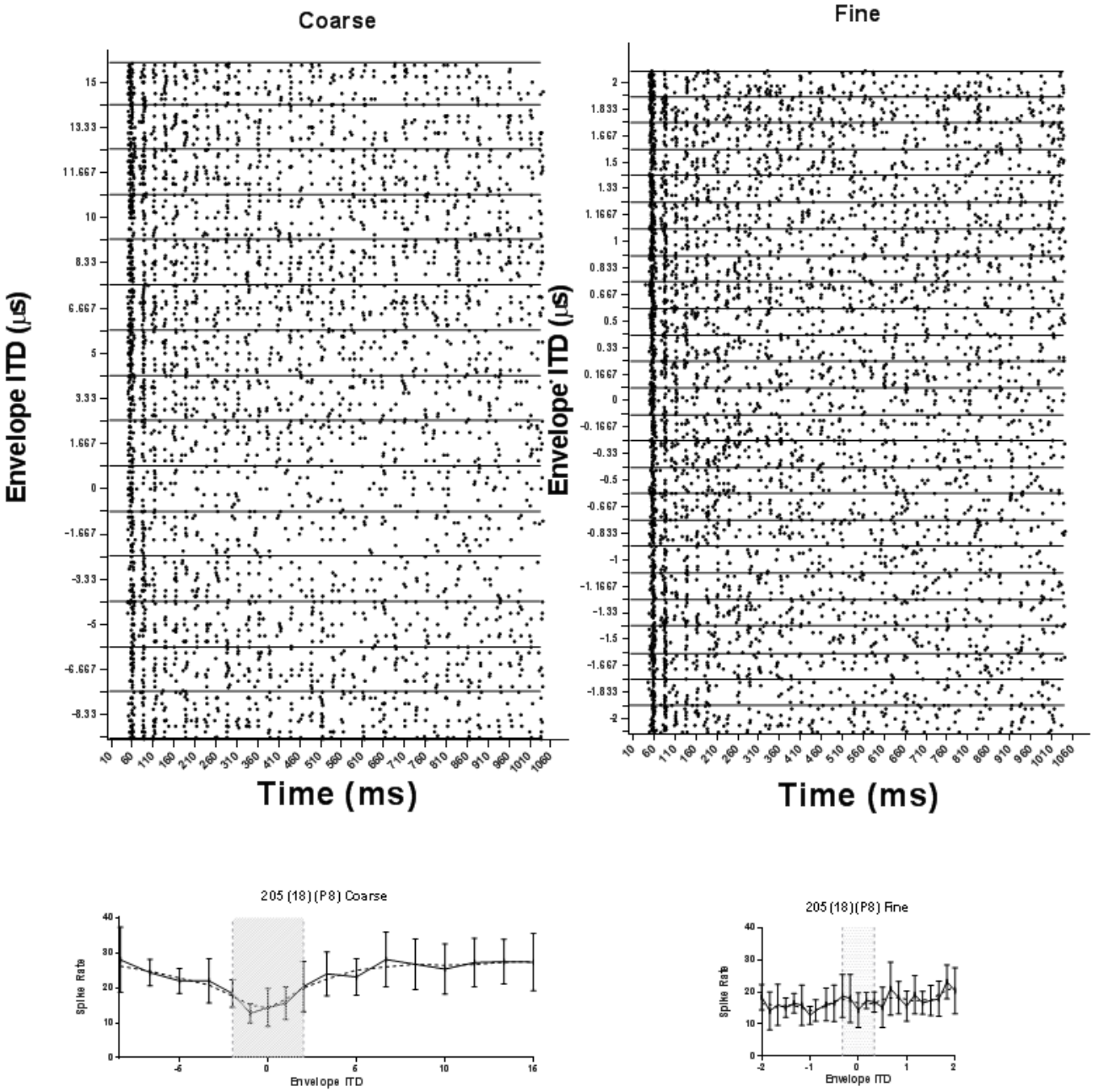


Figure 3.7 J

Figure 3.7. Neuron 205 (18) A) rate vs level response, BF = 5650 Hz, threshold = -35 dB re maximum system output. B-J) same format as Figure 3.1 excluding onset raster plots.

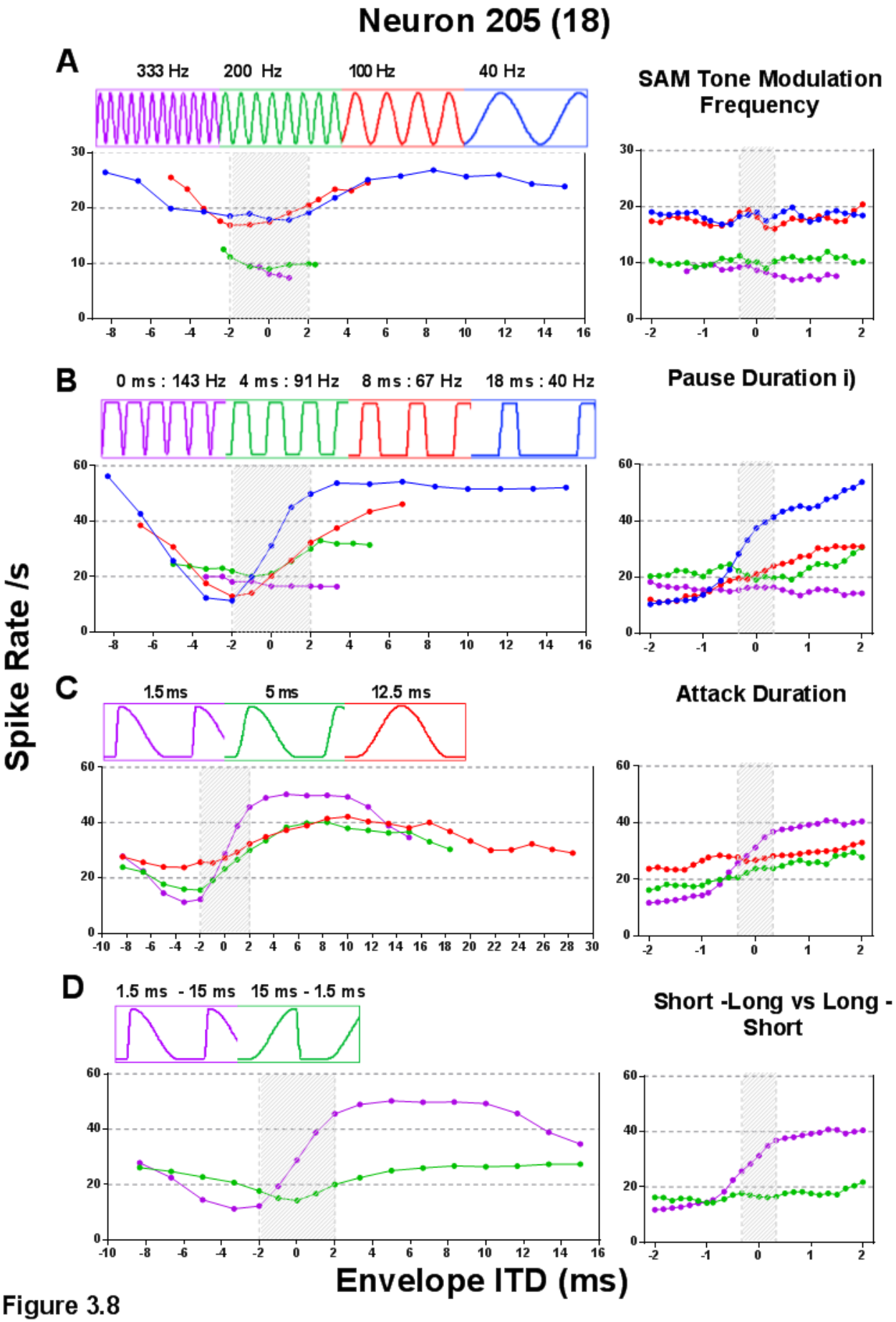


Figure 3.8

Figure 3.8 A-D) same format as Figure 3.2

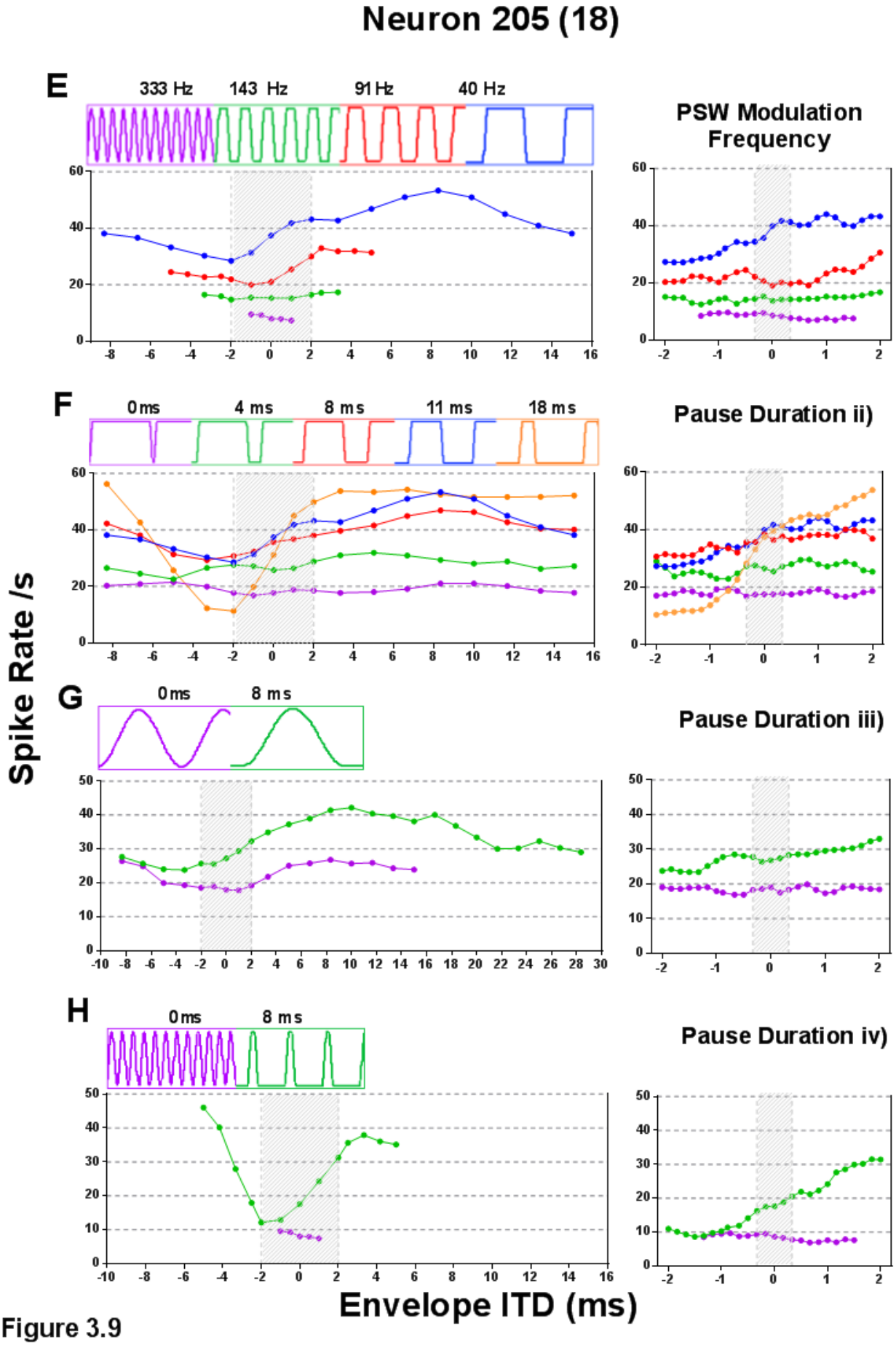


Figure 3.9

Figure 3.9 E-H) same format as Figure 3.3.

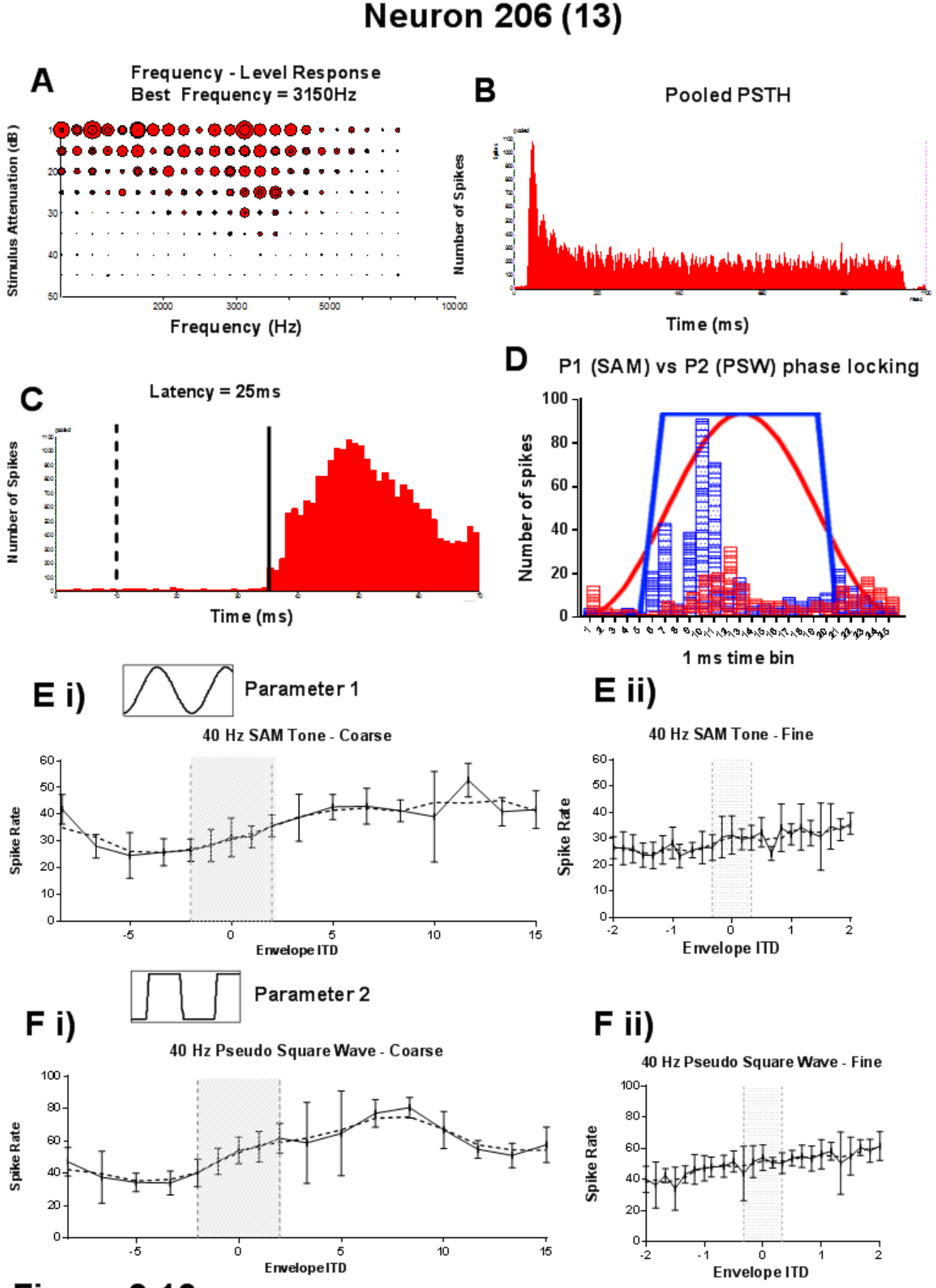

Neuron 206 (13)
A
Frequency - Level Response
Best Frequency = 3150Hz
Stimulus Attenuation (dB)
Frequency (Hz)
B
Pooled PSTH
Number of Spikes
Time (ms)
C
Latency = 25ms
Number of Spikes
Time (ms)
D
P1 (SAM) vs P2 (PSW) phase locking
Number of spikes
1 ms time bin
E i)
Parameter 1
40 Hz SAM Tone - Coarse
Spike Rate
Envelope ITD
E ii)
40 Hz SAM Tone - Fine
Spike Rate
Envelope ITD
F i)
Parameter 2
40 Hz Pseudo Square Wave - Coarse
Spike Rate
Envelope ITD
F ii)
40 Hz Pseudo Square Wave - Fine
Spike Rate
Envelope ITD

Figure 3.10

206 (13): 40 Hz PSW

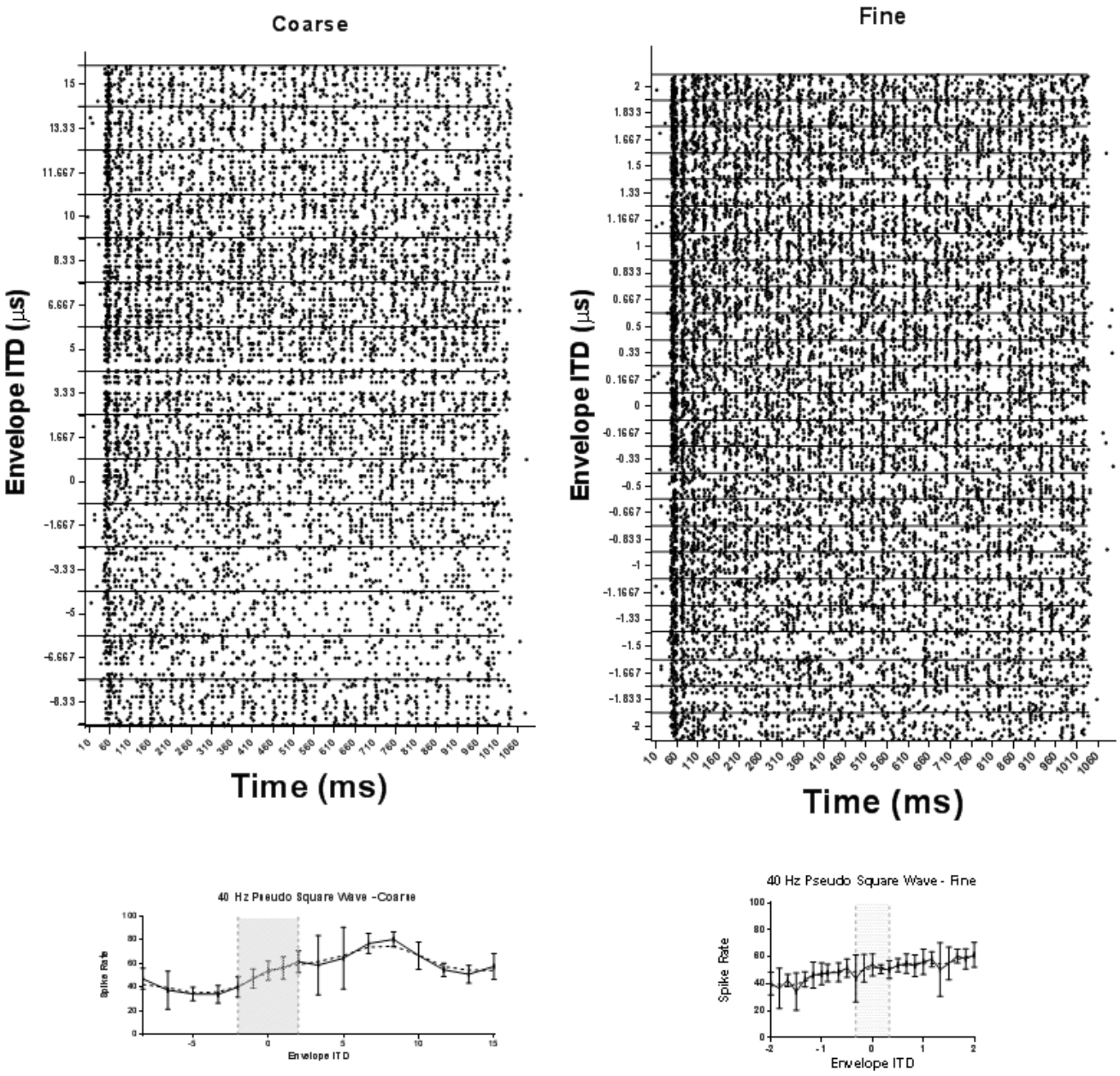


Figure 3.10 G

206 (13): 40 Hz SAM Tone

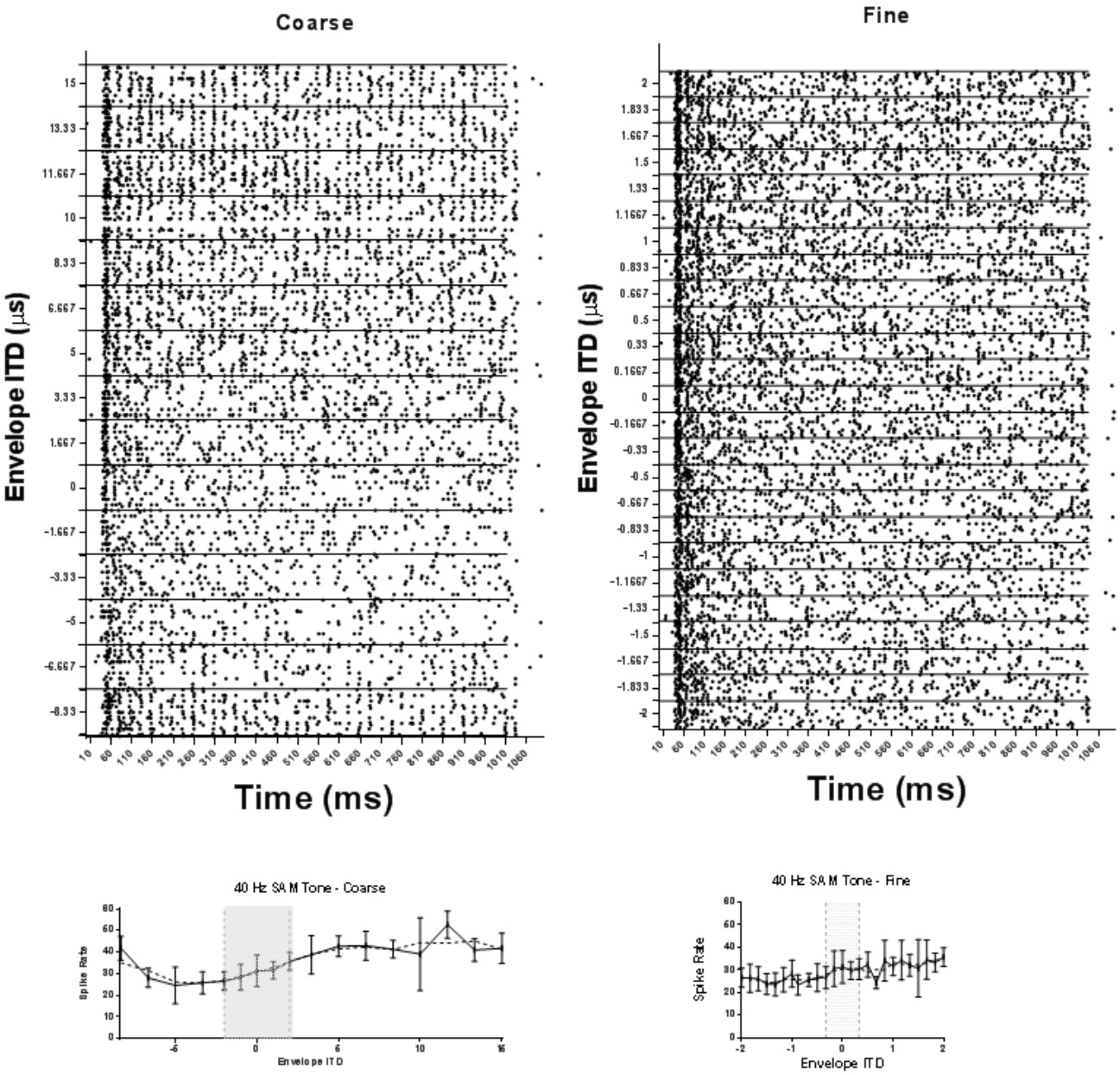


Figure 3.10 H

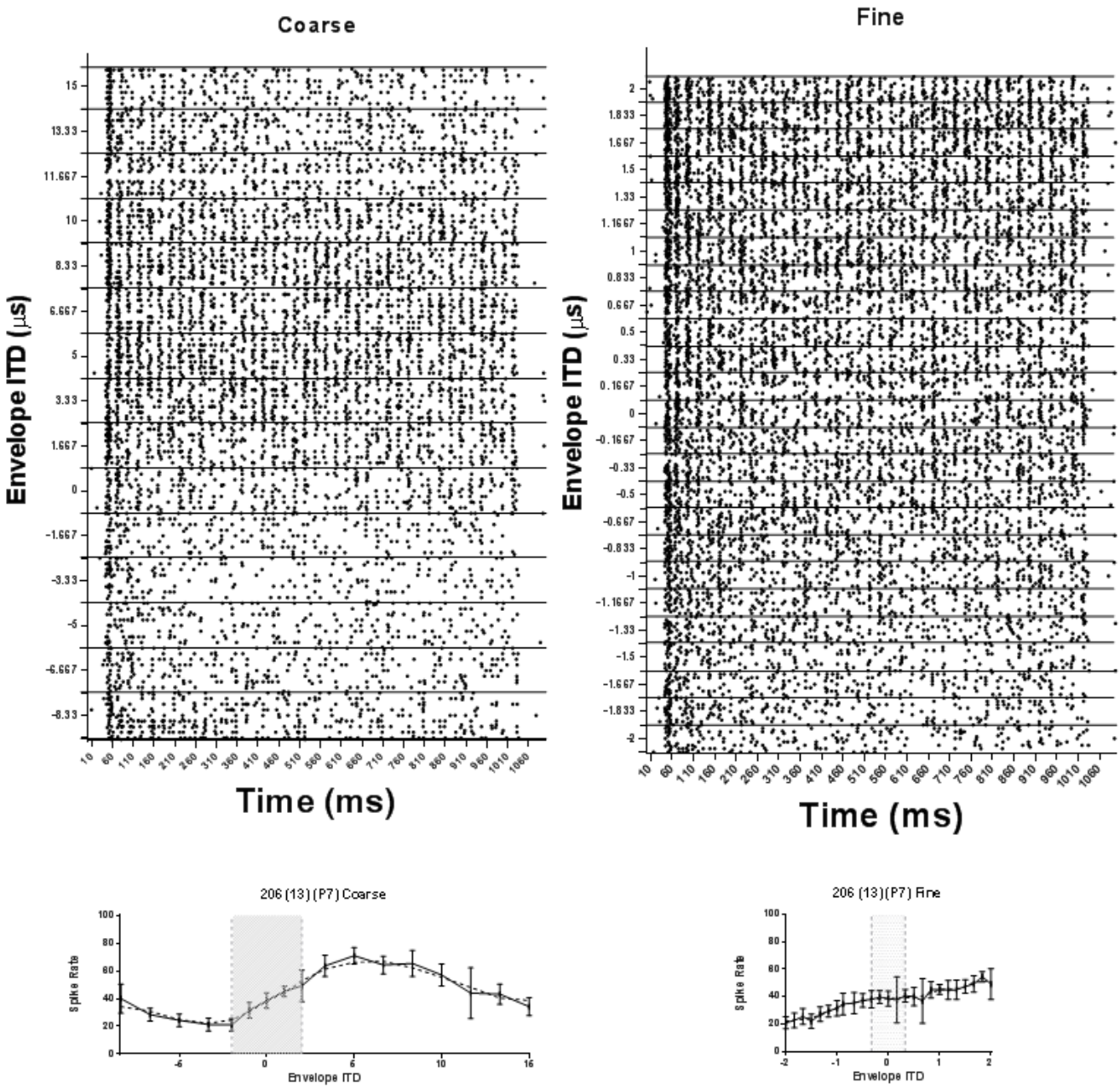

206 (13): 40 Hz 'Damped'
Coarse
Fine
Envelope ITD (μs)
Time (ms)
206 (13) (P7) Coarse
206 (13) (P7) Fine
Spike Rate
Envelope ITD

Figure 3.10 l

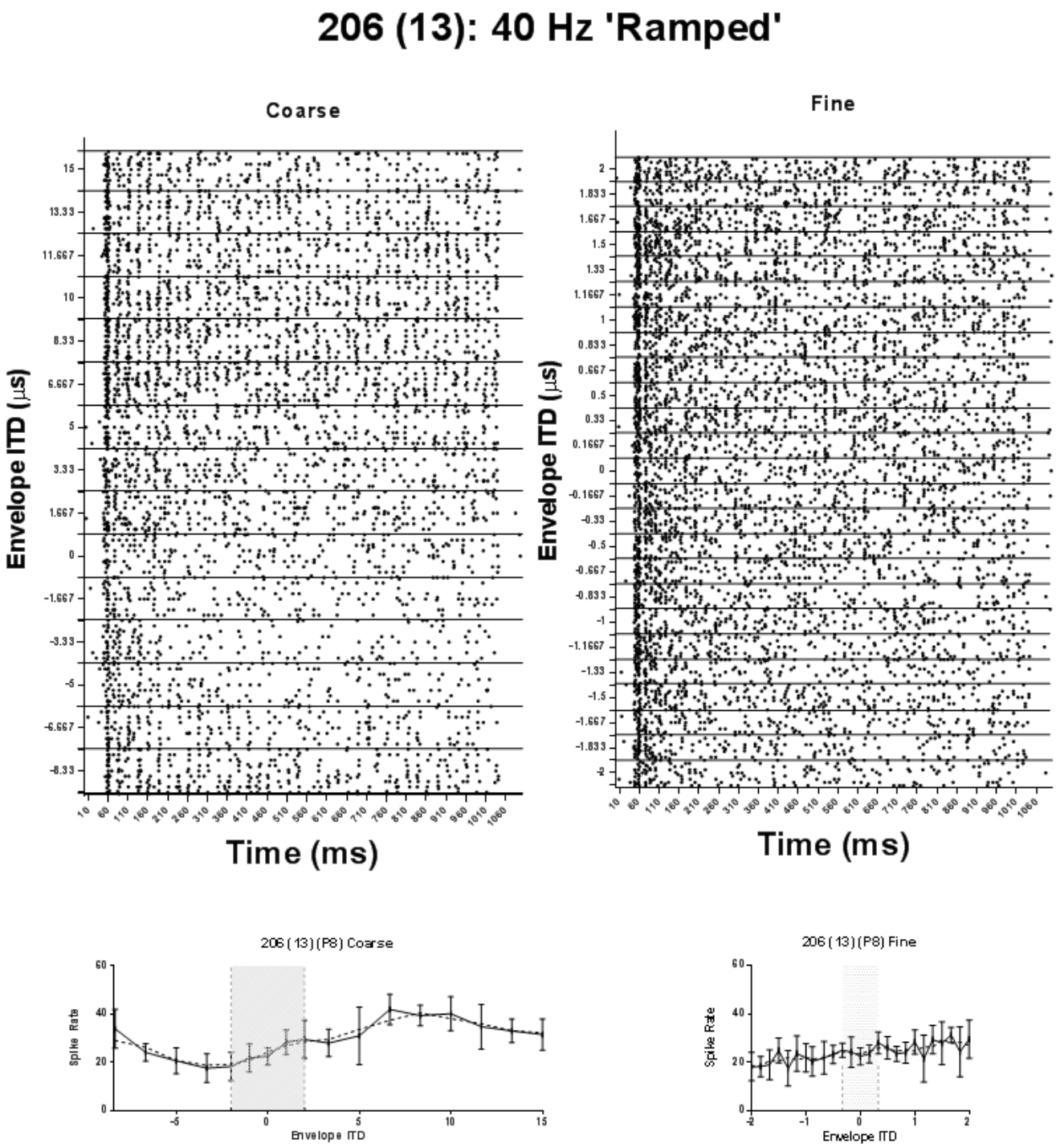


Figure 3.10 J

Figure 3.10. Neuron 206 (13) A) rate vs level response, BF = 3150 Hz, threshold = -40 dB re maximum system output. B-J) same format as Figure 3.1 excluding onset raster plots.

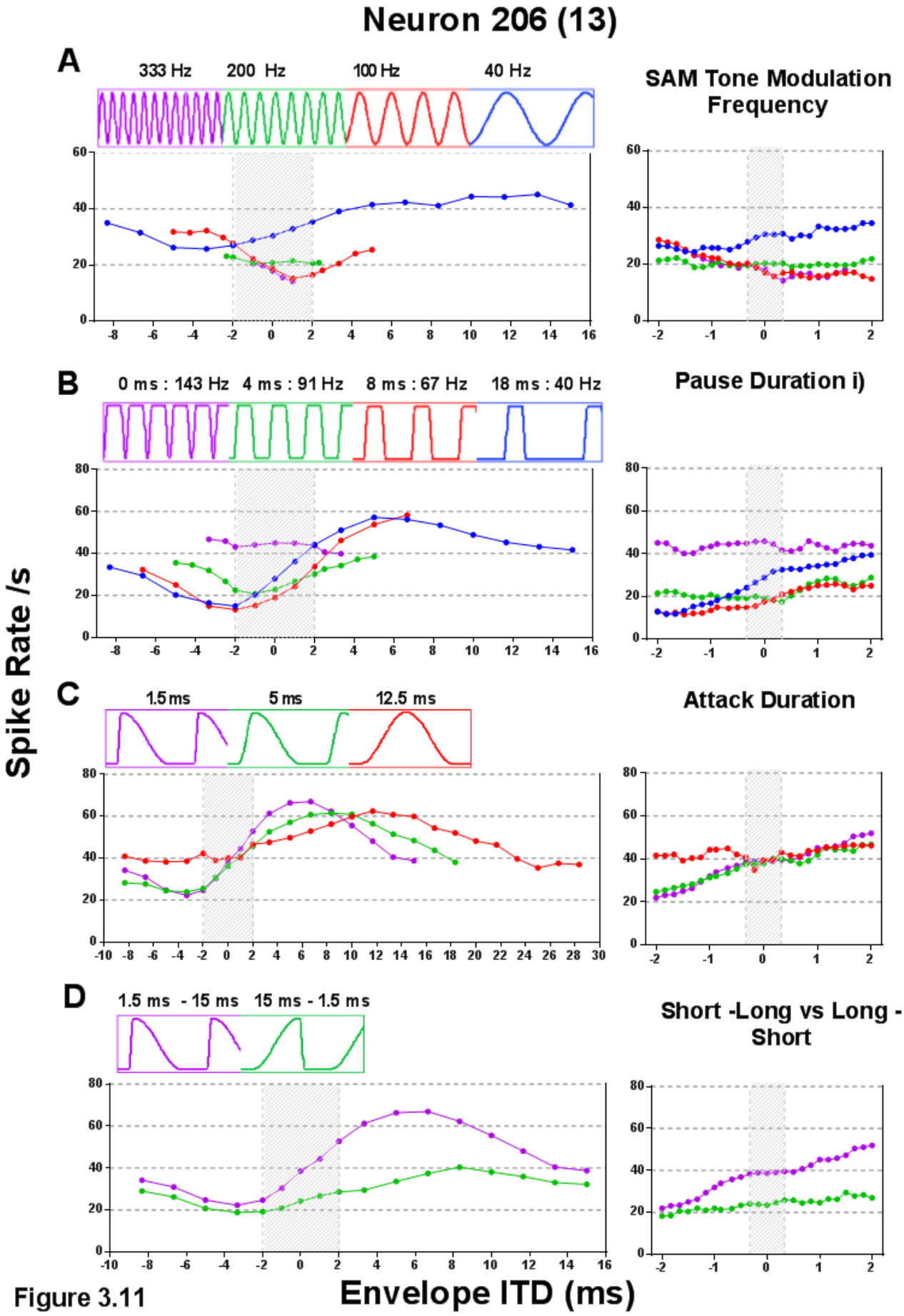


Figure 3.11

Figure 3.11 A-D) same format as Figure 3.2

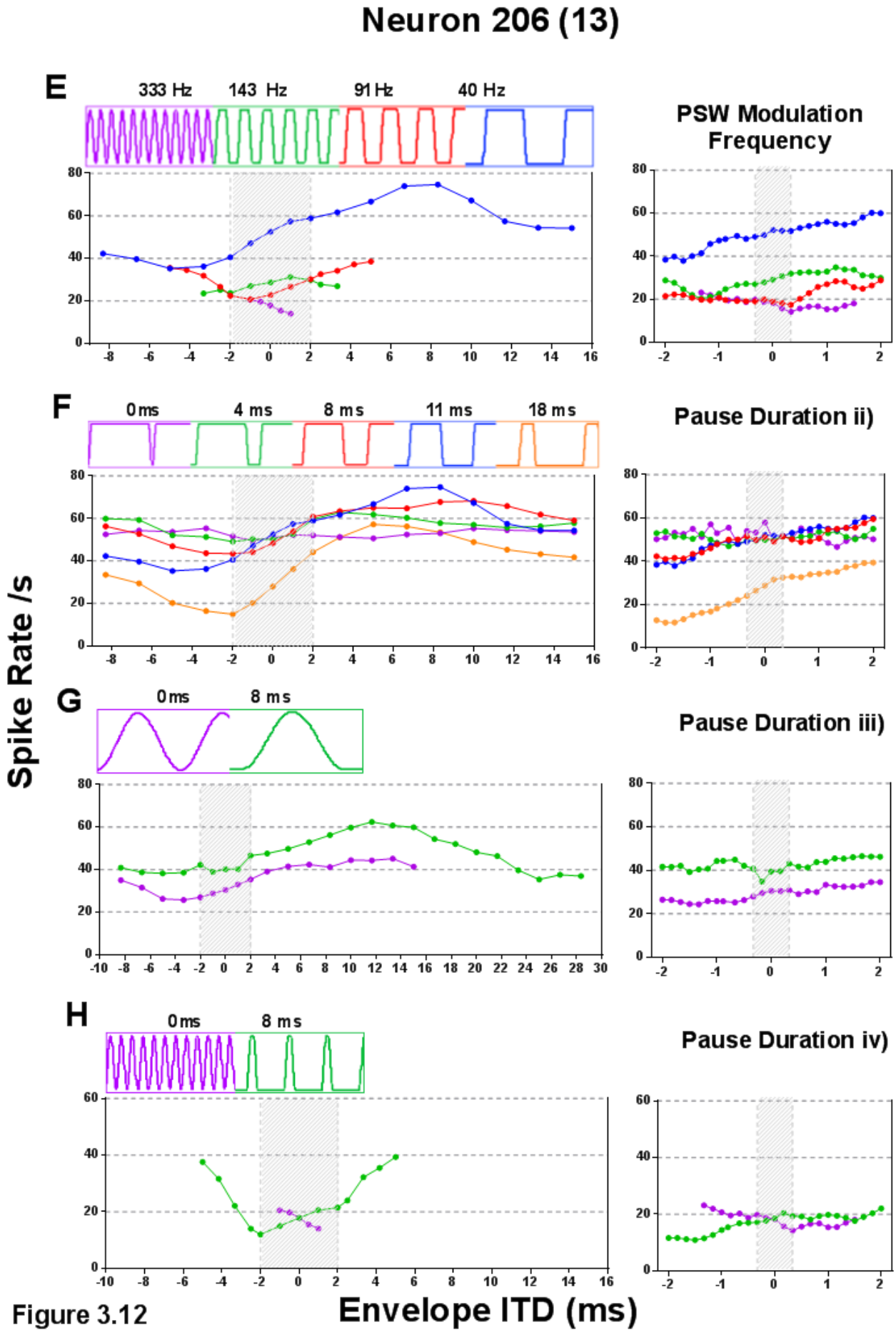


Figure 3.12

Figure 3.12 E-H) same format as Figure 3.3.

3.3.2 Neuron Population Characteristics – Damped vs Ramped

Responses to 18 unique envelope shapes were obtained from the right IC of 15 tri-coloured guinea pigs. A total of 172 IC neurons were isolated with 71 of these found to match the search criteria required for stimulation by the full parameter matrix. The search criteria included assessment of the characteristic frequency, spontaneous firing rate, firing rate in response to a stimulus with varying ITD and the minimum response threshold. After a neuron had met the required criteria the characteristic frequency (CF) and response threshold were measured and recorded. Post hoc analysis of each neuron included stimulus response latencies, rate ITD functions and ITD JNDs. The CFs of the 71 IC neurons were ≥1.4 kHz (1.4 – 12.2 kHz, median = 4.3 kHz) with 14 neurons (19%) having CFs occurring in the range of 2 - 3 kHz (Figure 3.14B). The stimulus response latencies were ≥14 ms (14 – 42ms, median = 25ms) (Figure 3.14A) in agreement with existing studies of first spike latency (Tan, Wang et al. 2008; Zohar, Shackleton et al. 2011)

The effect of envelope ITD on neuronal spike rates was measured for each of the 18 envelope shape parameters. The temporally asymmetric envelope shapes of parameter 7 (‘Damped’) and 8 (‘Ramped’) were selected so as to compare JNDs for ITD for IC neurons in detail prior to a broader population analysis with the full parameter matrix. The Damped envelope shape consists of a short Attack phase and a long Decay phase while the Ramped envelope shape shows the reverse. Note that these stimuli have identical spectra, differing only in the temporal order of the fast and slow components of their envelope. In human psychophysical assessments this difference has been shown to result in an increase of ITD JND by a factor of about 4, with envelope ITD JNDs increasing from approximately 100μs to 400μs (Klein Hennig, Dietz et al. 2011).

Latency

Figure 3.13 plots the neuron latency and best frequency for each of the 71 neurons in response to the stimulation protocol containing all 18 envelope shapes. A Pearson correlation analysis indicates no correlation between neuron latency and neuron best frequency (P=0.9627). Figure 3.14A illustrates that for this neuron population the distribution of latencies is centred around 26 ms with 19 (27%) neurons having a latency between 25 - 27 ms. In addition, 65 (93%) of the neurons in this population have a latency within the range of 21 - 33 ms.

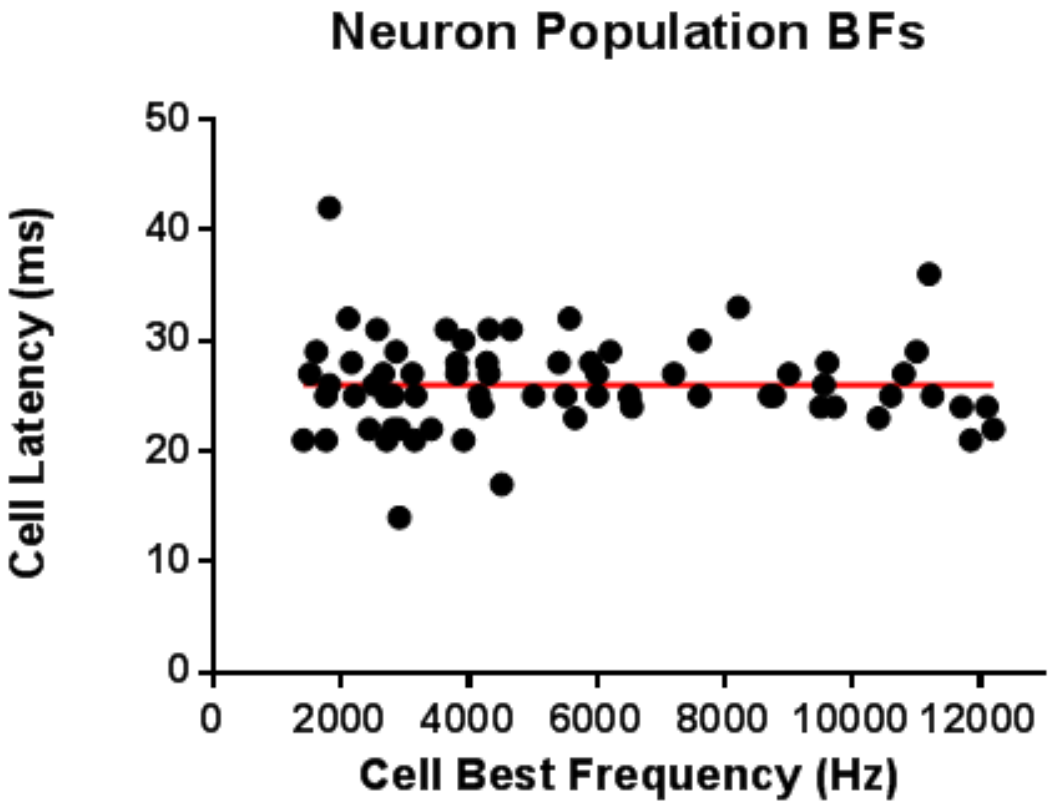


Figure 3.13 Neuron latency as a function of best frequency for each of the 71 neurons in response to the stimulation protocol containing all 18 envelope shapes. Best frequencies ≥1.4 kHz (1.4 – 12.2 kHz, median = 4.3 kHz) Pearson correlation analysis (red line) indicates no significant correlation (P=0.9627)

In order to assess the degree to which the first spike latency is dependent on the stimulus envelope shape for the current neuron population, Figure 3.15A plots the first spike latency for all 71 neurons, when the responses to the Damped and Ramped envelope shape are considered independently. For both the Damped and Ramped envelope shape there is a general trend but no significant correlation between first spike latency and best frequency ($P = 0.0534$ and $P = 0.1281$ respectively). The Damped envelope shape elicits a mean first spike latency of 35.25 ms with a standard deviation of 10.48 ms and the Ramped envelope shape a mean of 44.10 ms and a standard deviation of 13.74. These differences may be due to the fact that the short 1.5ms Attack-phase of the Damped envelope shape results in the neurons reaching an internal threshold criteria for activity more rapidly than with the long 15ms Attack-phase of the Ramped envelope shape. Figure 3.15B plots the first spike latency in response to the Damped and Ramped envelope shapes for a within-neuron comparison. The line-of-best-fit remains above the dashed line of unity, confirming that the Damped envelope shapes first spike latency is in general smaller than that of the Ramped envelope shape. Figure 3.14 C and D displays the frequency distributions of first spike latencies in response to the Damped and Ramped parameters respectively supporting the findings presented in Figure 3.15.

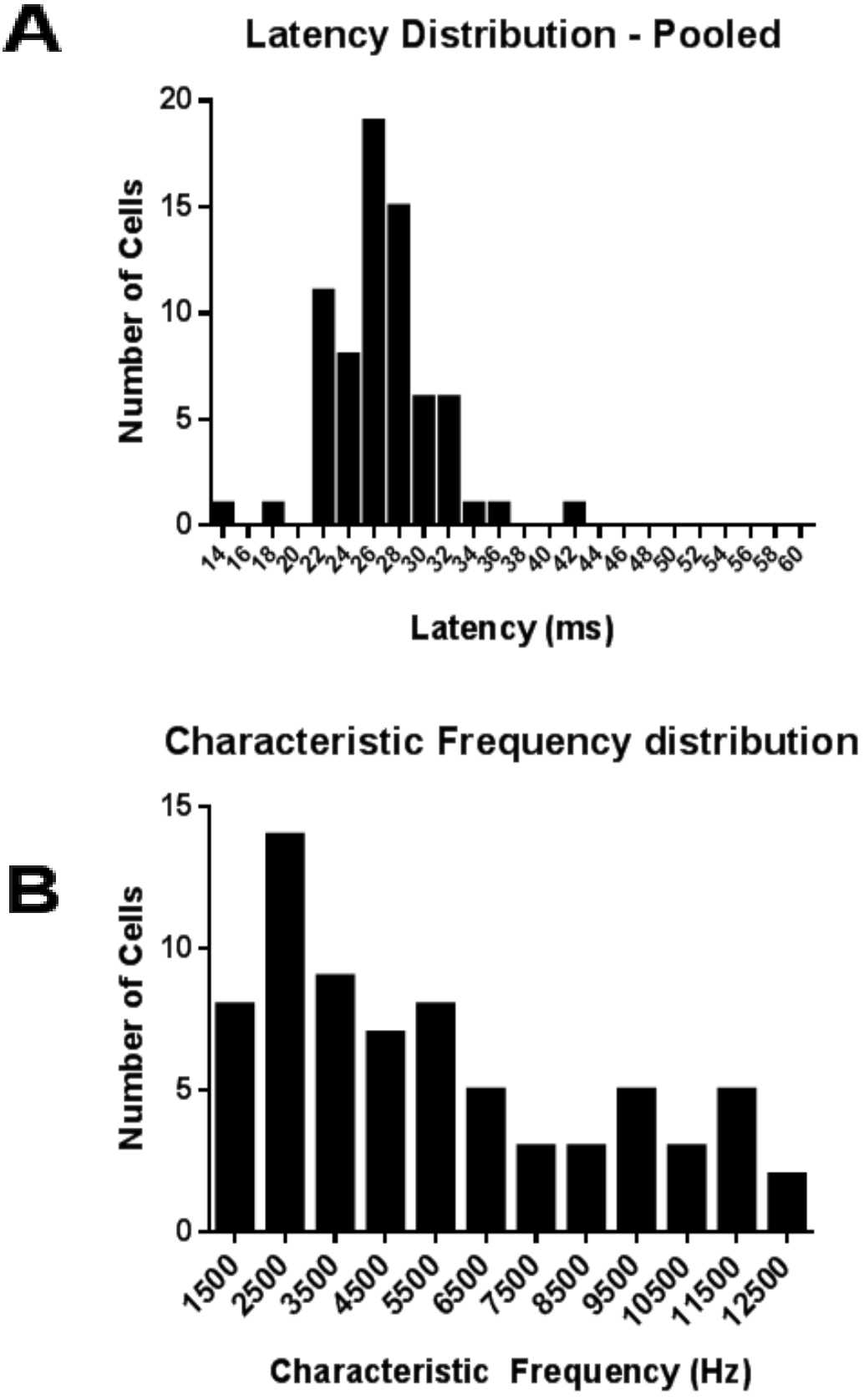
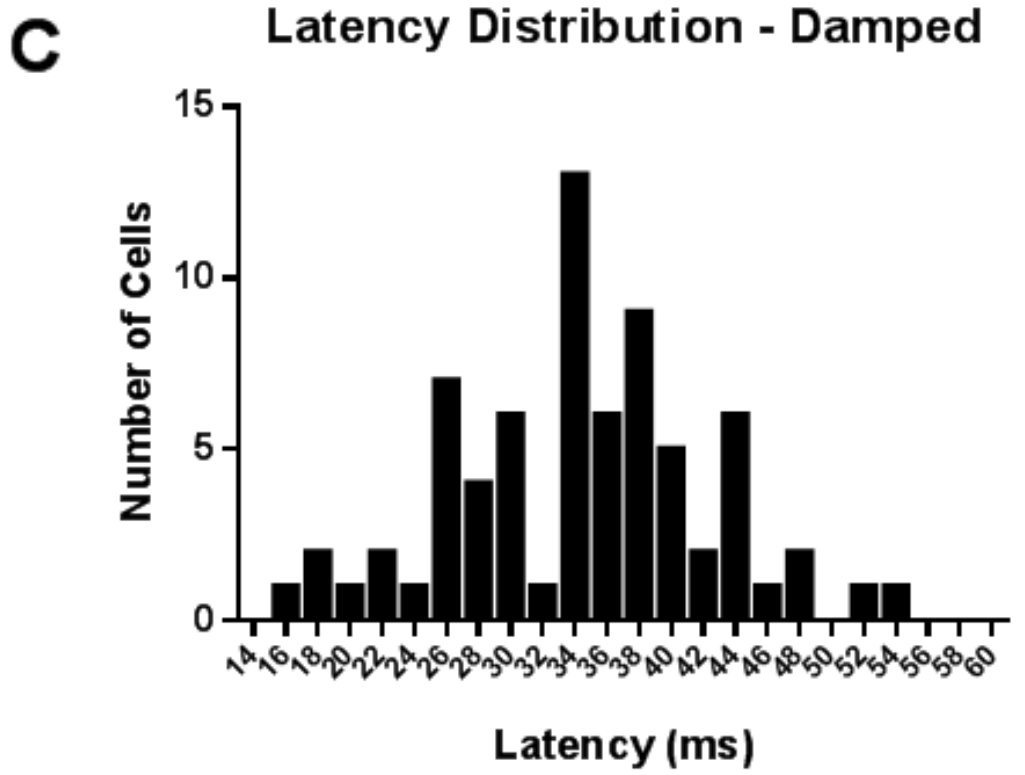


Figure 3.14 A and B) see next page for legend.



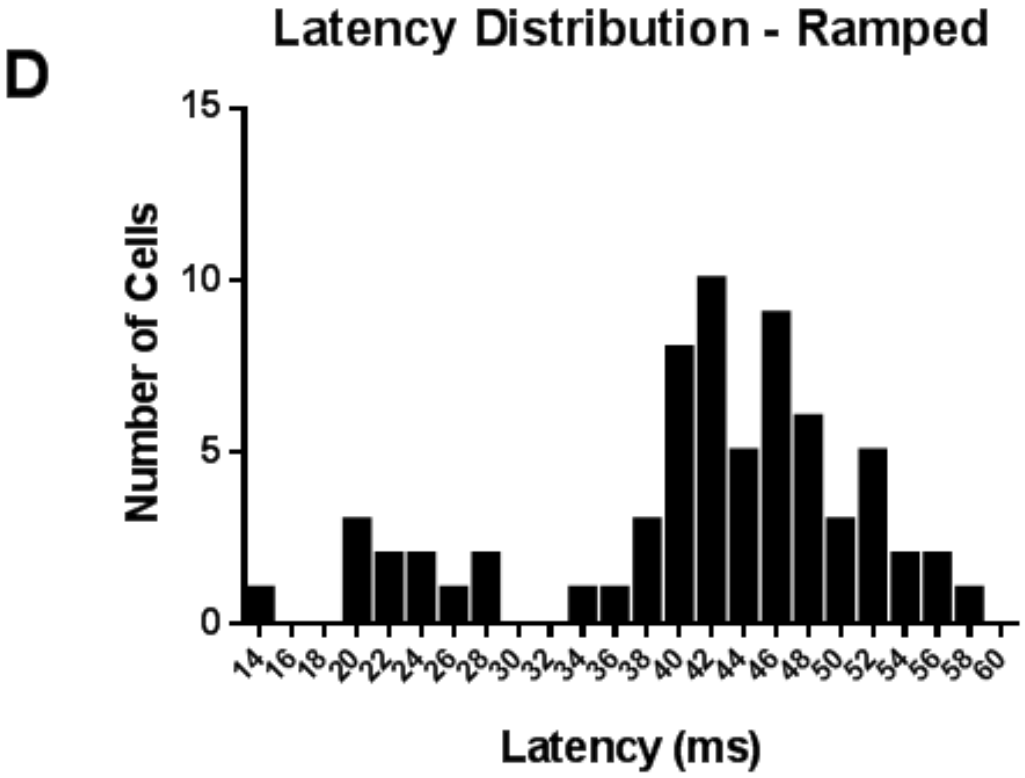


Figure 3.14 A) Frequency distribution of pooled first spike latencies ≥14 ms (14 – 42ms, median = 25ms) B). Frequency distribution of characteristic frequencies. C) and D) latency distributions of first spike latencies in response to the Damped and Ramped parameters respectively

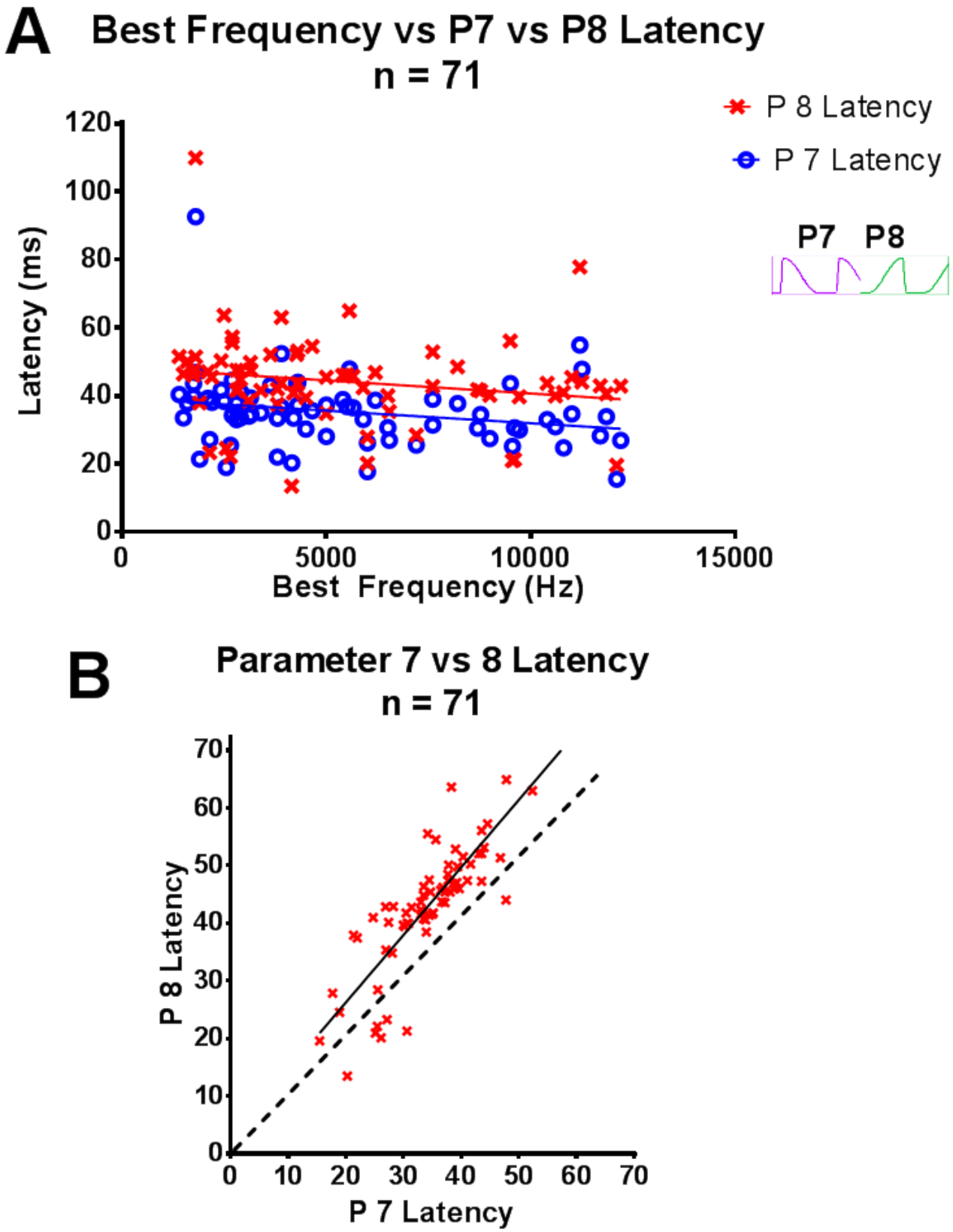


Figure 3.15

Figure 3.15 A) first spike latency for all 71 neurons. Red = Ramped, Blue = Damped. B) within-neuron comparison of first spike latency for Ramped (P8) and Damped (P7)

ITD-rate-functions

The ITD rate functions for the Ramped and the Damped envelope shapes are comprised of spike responses to envelope ITDs from -8.33 to +15ms. Between ± 2ms ITD the response rate was recorded at 25 evenly spaced ITDs contributing to the 'fine' recording. The 'coarse' ITD rate functions were created using recordings at 17 different envelope ITDs between -8.33 and +15ms. In order to examine the shapes of ITD rate functions in response to the Ramped and the Damped envelope shapes, the functions were smoothed using three-point averaging followed by normalising to the maximum and minimum firing rates. The normalised rate ITD functions for all 71 neurons for both the fine and coarse recordings can be seen in Figure 3.16. In red is the average normalised response for all 71 neurons. The ± 330 μs physiological range of guinea pig ITDs (Sterbing, Hartung et al. 2003) is indicated by the shaded region. The plots on the right of Figure 3.16 are the mean and standard deviations of the responses of all 71 neurons to the Ramped and the Damped envelope shapes. These normalised responses indicate that the Ramped envelope shape elicits a more gradual change in spike rate across envelope ITDs when compared to the Damped envelope shape which stimulates a more rapid change in spike rate across a narrower range of envelope ITDs. The rapid change in firing rate across a narrow range of ITDs close to 0 ITD for the Damped envelope shape is followed by a more constant spike rate for the large positive envelope ITDs compared to a steady rise and fall pattern for the Ramped envelope shape.

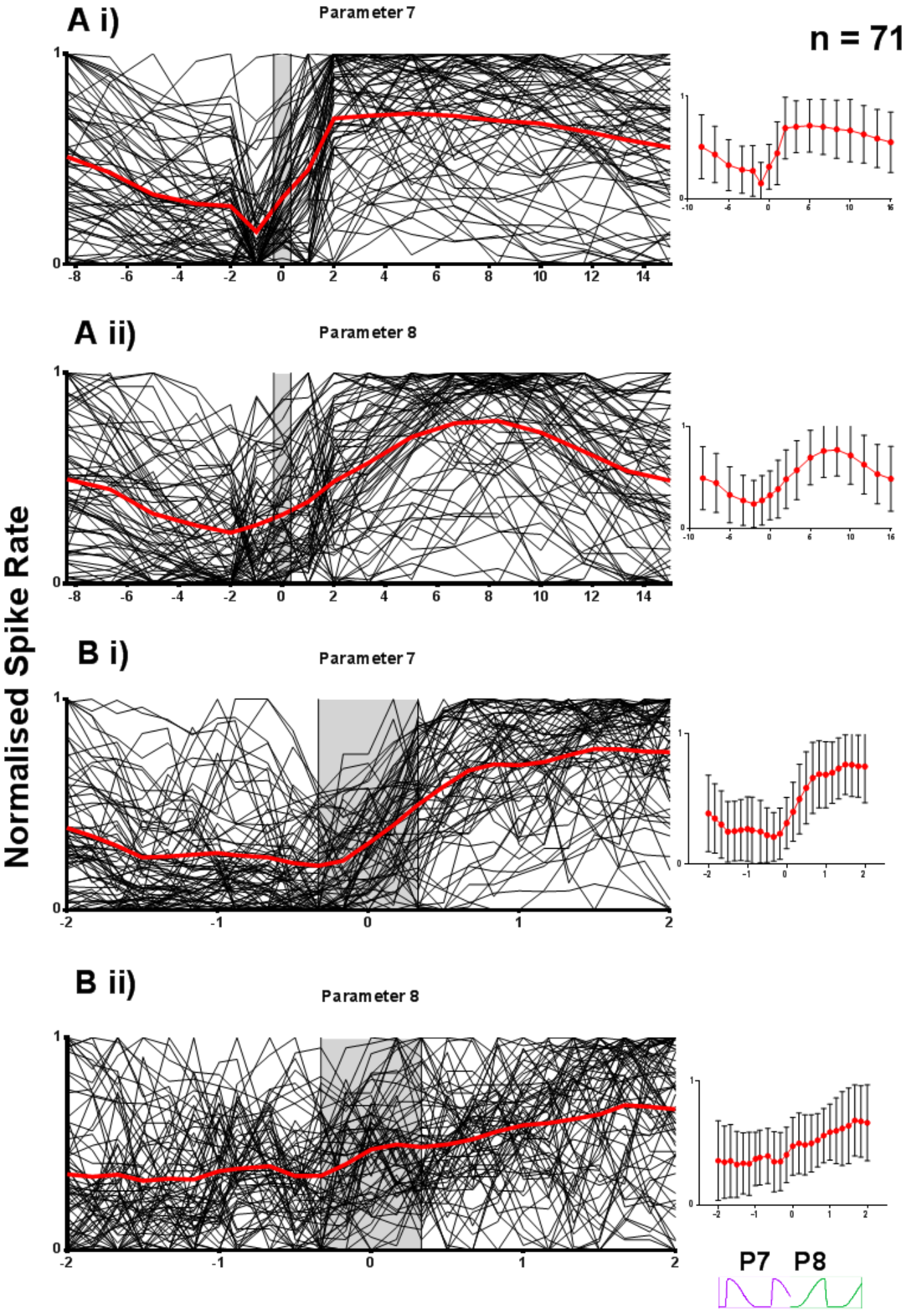


Figure 3.16 A-B) see next page for legend.

Figure 3.16 Normalised rate ITD functions for all 71 neurons for A) Damped and B) Ramped for the i) fine and ii) coarse recordings. Red line = mean normalised response for all 71 neurons. Shaded region = ± 330 µs physiological range of guinea pig ITDs. Right side panels = mean and standard deviations of the normalised responses of all 71 neurons.

Standard Separation (D)

Figure 3.17 depicts the normalised mean firing rates for the Damped and Ramped envelope shape from 71 neurons in the ± 2ms ITD range. The grey shaded region depicts the ± 330 µs guinea pig physiological range for ITDs. The responses used to calculate the value of standard separation (D) (Sakitt 1973) are denoted by black crosses. D = 0.6155 for the Ramped envelope shape and D = 1.3006 for the Damped envelope shape, over twice as large, representative of the rapid change in firing rate across a narrow range of envelope ITDs for the Damped envelope shape compared to the Ramped envelope shape.

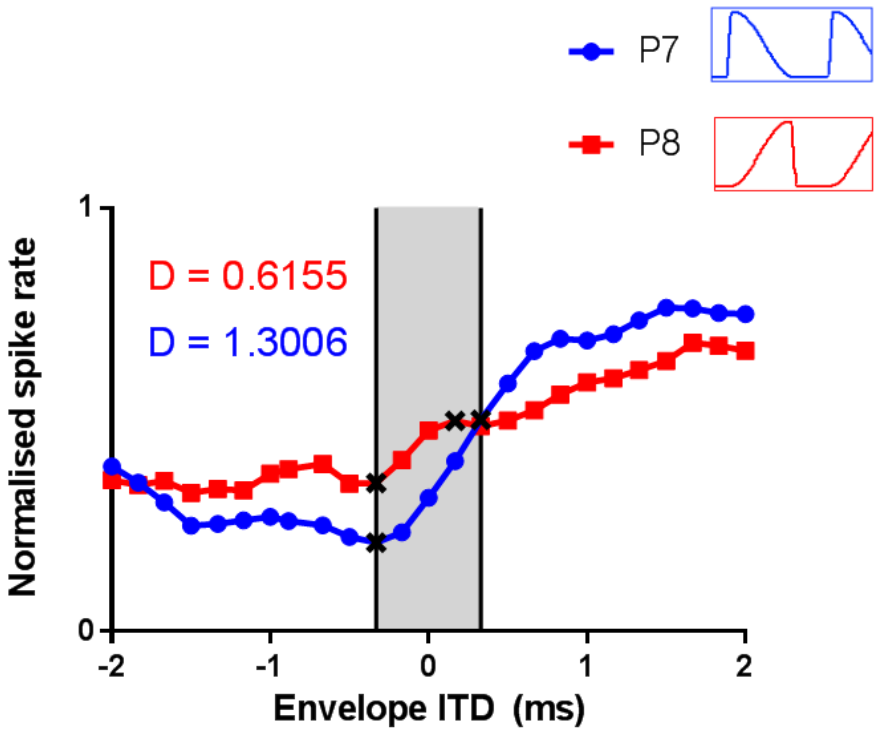


Figure 3.17

Figure 3.17 Normalised mean firing rates for Damped (P7) and Ramped (P8) from 71 neurons in the ± 2ms fine ITD range. The grey shaded region depicts the ± 330 μs guinea pig physiological range of ITDs. The responses used to calculate the value of standard separation (D) are denoted by black crosses.

JND for ITD

Each neuron's JND for ITD was found by applying a standard separation value (D) (Sakitt 1973) as described in Section 3.2.6 - *Data Analysis*. The criteria selected for accepting a threshold were D = 2, 1.5 and 1, in order to provide a comprehensive evaluation of ITD JNDs as defined by different degrees of change in the firing rate of a neuron. Figure 3.18 plots JNDs for ITD for all neurons with a JND that met the specific criterion within the range between ± 2ms ITD for the Damped and Ramped envelope shape and 8.

Figure 3.18A illustrates that for D = 2, 30 IC neurons (43%) attain a JND threshold that occurs between ±2ms ITD, with 5 (17%) of these occurring within the guinea-pig physiological range of ±330 μs ITD. Reducing the value of D to 1.5 equates to a more easily discriminable change in spike rate being set as the criterion required to reach an ITD JND. This is reflected in the increasing number of neurons that reach threshold as the D is set at 1.5 and 1 with 42 (59%) and 55 (77%) reaching threshold, respectively. Alongside the increasing numbers of neurons showing a threshold ITD for the 'Damped' stimulus, a greater proportion of JNDs are observed within the physiological range; 15/42 (36%) for D = 1.5 and 26/55 (47%) for D = 1. In contrast far fewer neurons are sensitive to ITDs carried in the envelope of the Ramped envelope shape, for D =2, 1.5 and 1, the proportion of sensitive neurons was 13% (9/71), 27% (19/71) and 54% (38/71), respectively.

Figure 3.18B plots the absolute JNDs (i.e. pooled across positive and negative ITDs) for the neuron population in response to the Ramped and Damped parameters. For each D level, the geometric mean of the data is represented by the red horizontal line. For the three threshold criteria, D = 2, 1.5 and 1, the mean ITD JNDs for the Damped envelope shape are 822 μs, 709 μs, and 535 μs, respectively. These mean values are somewhat larger than those ITD JNDs observed in human psychophysical studies (100 μs – Klein Hennig et al. (2011)). In response to the Ramped envelope shape 20/37 (54%) of the neurons, at D = 1, have ITD JNDs within the range of the behavioural thresholds and the mean response of 642 μs is close to the range of responses from the psychophysical study.

Figure 3.19 illustrates the frequency distribution of ITD JNDs for the present neuron population. The bin widths are 166 μs in order to distinguish the ± 330 μs physiological range of guinea pig ITD JNDs. Again the larger number of neurons reaching threshold in the presence of the Damped envelope shape and for smaller D levels is apparent.

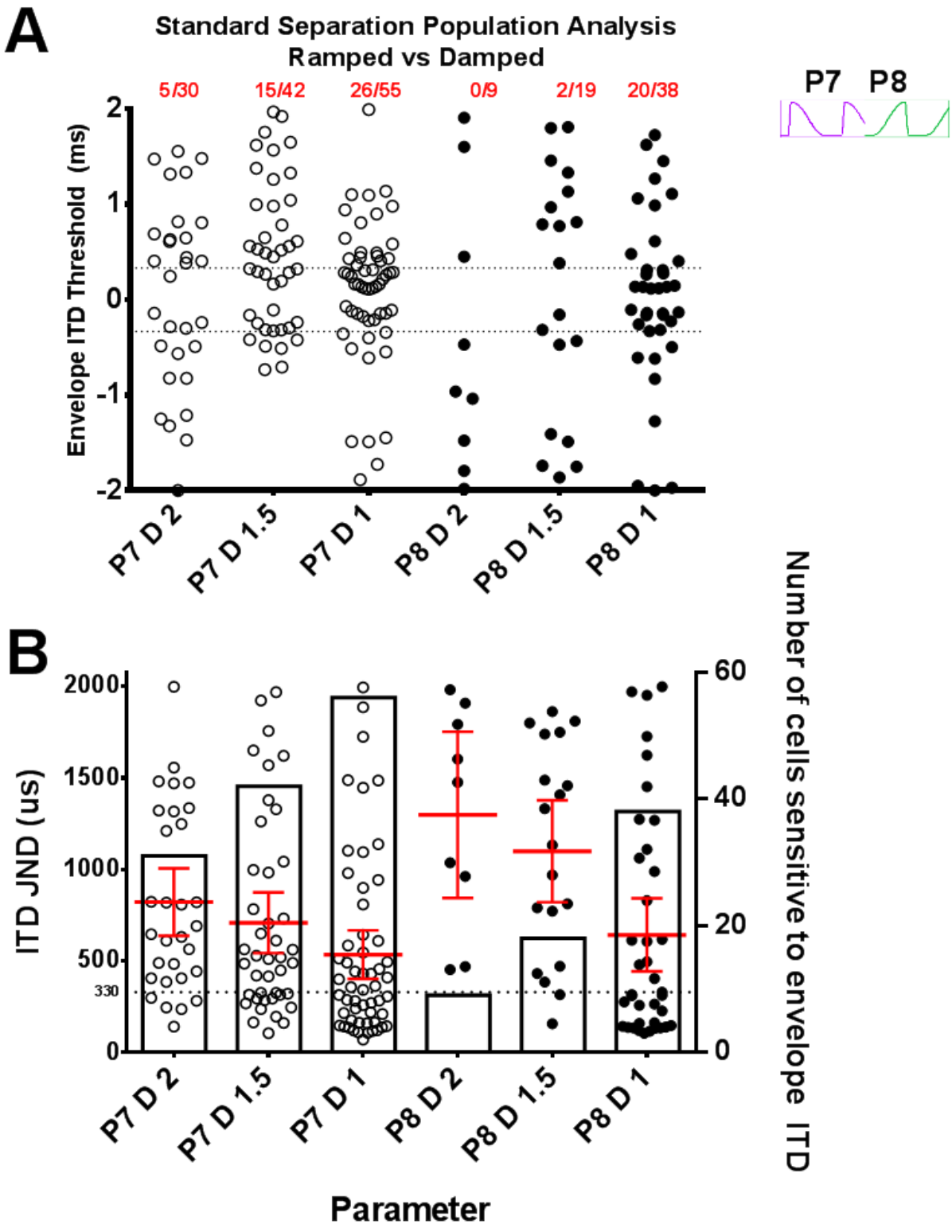


Figure 3.18 A) JNDs for ITD for D = 2, 1.5 and 1 between ± 2ms ITD for Damped and Ramped. Dashed lines = guinea pig physiological range of ±330 µs ITD. Red number = neurons with ITD JND within physiological range/neurons with ITD JND with ± 2ms ITD. B) Absolute JNDs for ITD (i.e. pooled across positive and negative ITDs) for D = 2, 1.5 and 1 between ± 2ms for Ramped and Damped parameters. For each D level, the mean and 95% confidence interval is in red.

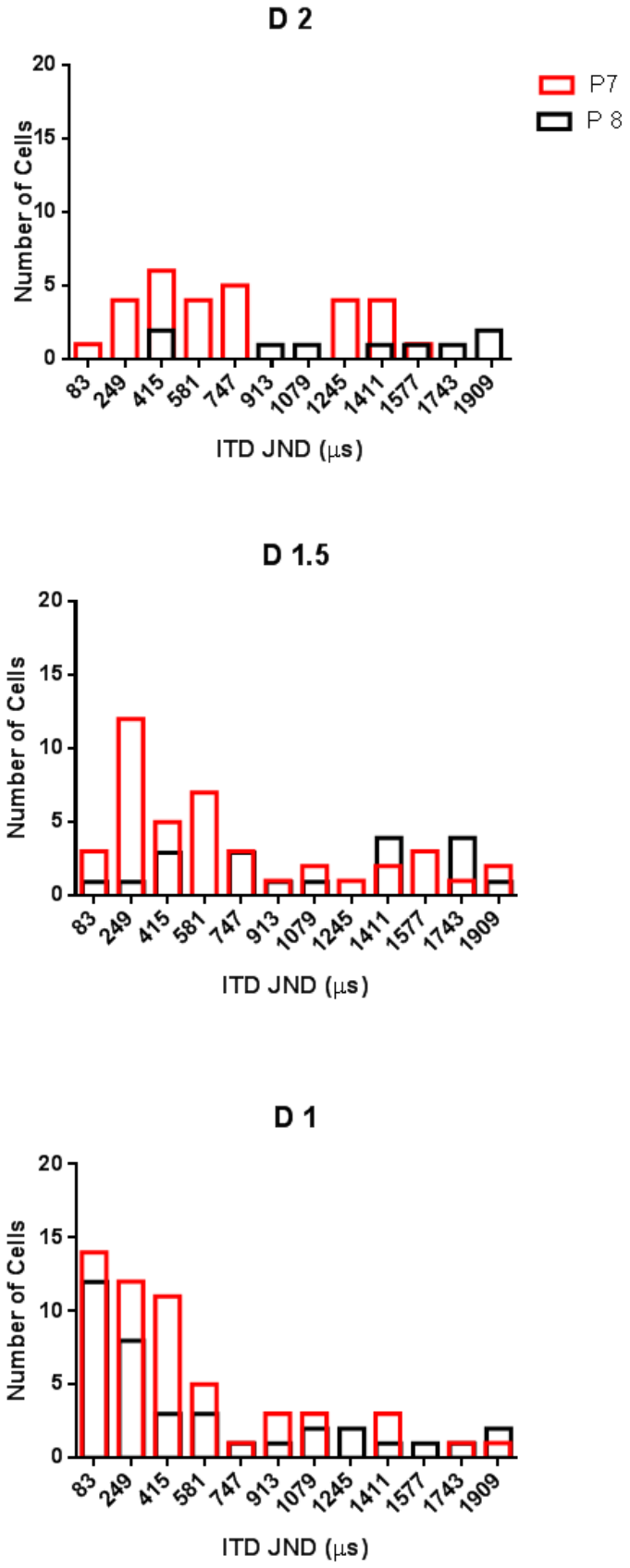


Figure 3.19

Figure 3.19 A-C) frequency distribution of ITD JNDs for the neuron population at D = 2, 1.5 and 1. Bin widths are 166 µs in order to distinguish the ± 330 µs physiological range of guinea pig ITD JNDs. Red = Damped (P7) Black = Ramped (P8)

3.3.3 Neuron Population Analysis

The 18 envelope shapes are divided into the same 8 comparison groups as in the single neuron analysis. Rather than focussing on the specific rate ITD functions for individual neurons, the rate ITD functions have been used to produce ITD JNDs for each neuron in response to each envelope shape. The criterion selected for accepting a threshold was a standard separation of D = 2, being the highest degree of difference in the spike responses between two ITDs that has been utilised thus far. This will provide robust categorisation of neurons that are ITD sensitive. For SAM modulation frequency and pure tone duration i), ITD JNDs for a criteria of D = 1 are also presented. In addition, only ITD thresholds that are within the maximum physiological range for a guinea pig (±330 μs (Sterbing, Hartung et al. 2003)) are included. The absolute JNDs presented here are therefore pooled across positive and negative ITDs for the neuronal population. All of the data has been analysed using the responses of 71 neurons as highlighted previously. In each Figure the left y-axis is the ITD JND (μs) on a logarithmic scale with the individual points plotted on each Figure representing the ITD threshold of a single neuron. The grey horizontal line is the geometric mean threshold of all neurons that reached the threshold criterion, and is plotted for each envelope shape. As well as the ITD thresholds, of interest is the percentage of the neuron population that are sensitive to a given envelope shape. The total number of sensitive neuron is therefore plotted using the right y-axis and is represented by the vertical bars in each Figure. The two horizontal dashed lines mark the limits of 330 μs and 660 μs.

SAM Tone Modulation Frequency

Figure 3.20: The ITD JNDs for neurons that are sensitive to SAM tone modulation frequencies of 333, 200, 100 and 40 Hz. For D = 2 there we no neurons with ITD JNDs smaller than 660 μs for ITDs carried in the envelope of

SAM tones with modulation frequencies of 333 and 200 Hz, within this neuron population. In response to the 100 Hz SAM tone (parameter 11) 2 neurons had ITD JNDs within the set limits while there were 3 neurons with ITD JNDs for the 40 Hz SAM tone (parameter 1). The ITD JNDs of the 3 neurons within the 0 to 660 μs range for parameter 1 were widely spread across this range with the smallest being 99 μs and the largest at 650 μs. The 2 neurons responsive to parameter 11 had relatively large and similar ITD JNDs with a mean of 536 μs. For a threshold of criterion of D = 1 there are more ITD sensitive cells and the mean threshold is also lower.

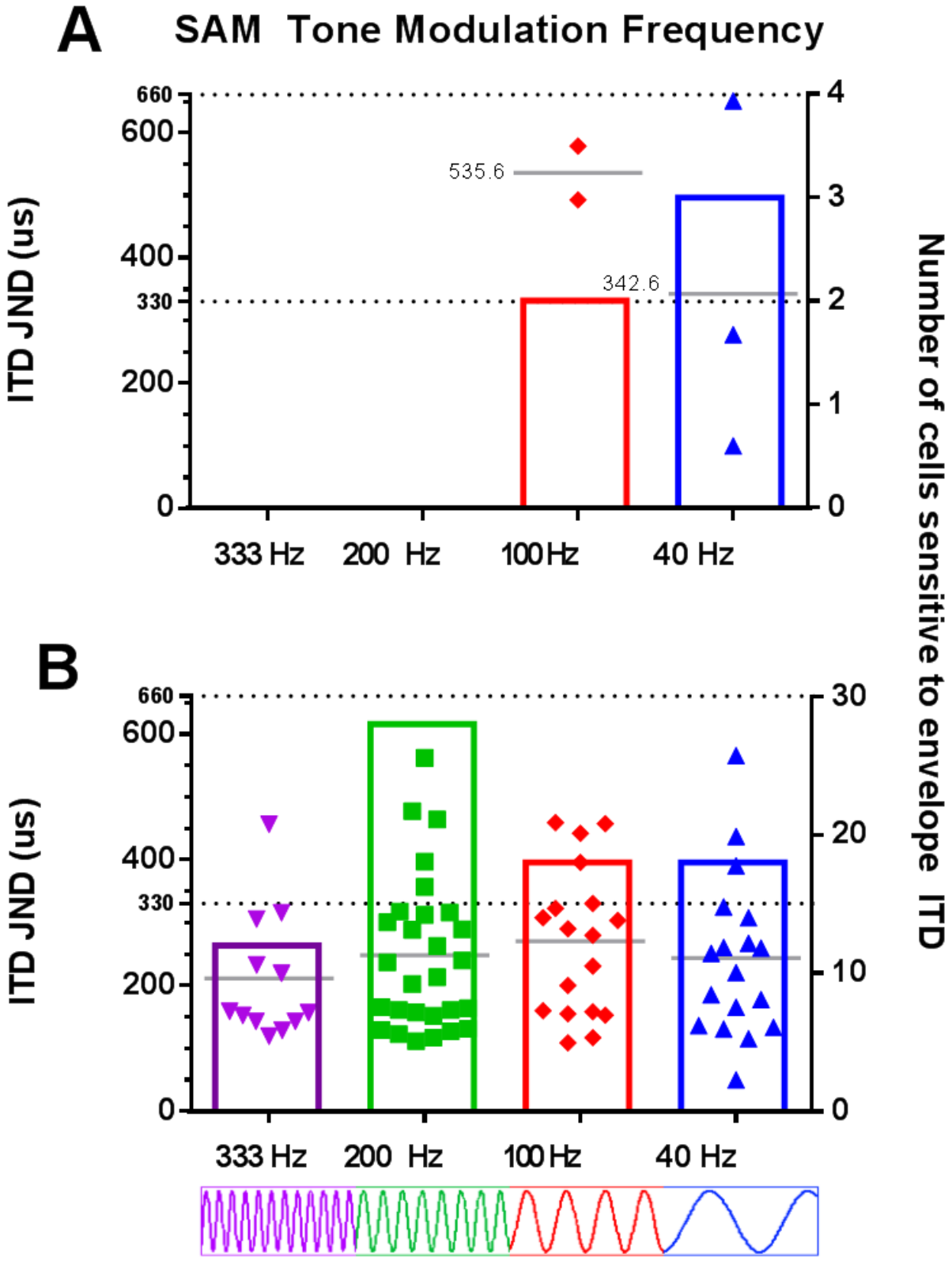


Figure 3.20 A) ITD JNDs for neurons that are sensitive to SAM tone modulation frequencies of 333, 200, 100 and 40 Hz when D = 2 and B) D = 1. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

PSW Modulation Frequency

Figure 3.21: The ITD JNDs for neurons that are sensitive to modulation frequencies of 333, 143, 91 and 40 Hz. The 333 Hz PSW has equal envelope segment durations to the SAM tone of the same frequency and is therefore also found to elicit no ITD JNDs from the current neuron population when D = 2. In response to the 143 Hz PSW (parameter 12) only 2 neurons had ITD JNDs within the set limits while a reduction in the modulation frequency to 91 Hz (parameter 13) and 40 Hz (parameter 2) increases the number of sensitive neurons to 8 (11%) and 11 (15%) respectively. The mean ITD JNDs of the current neuron population in response to parameter 2 and 13 are close to identical at 412.6 and 412.7 µs. The two neurons sensitive to parameter 12 both have thresholds at the lower end of the range of ITD JNDs elicited by parameter 2 and 13 resulting in a lower mean threshold of 252.5 µs.

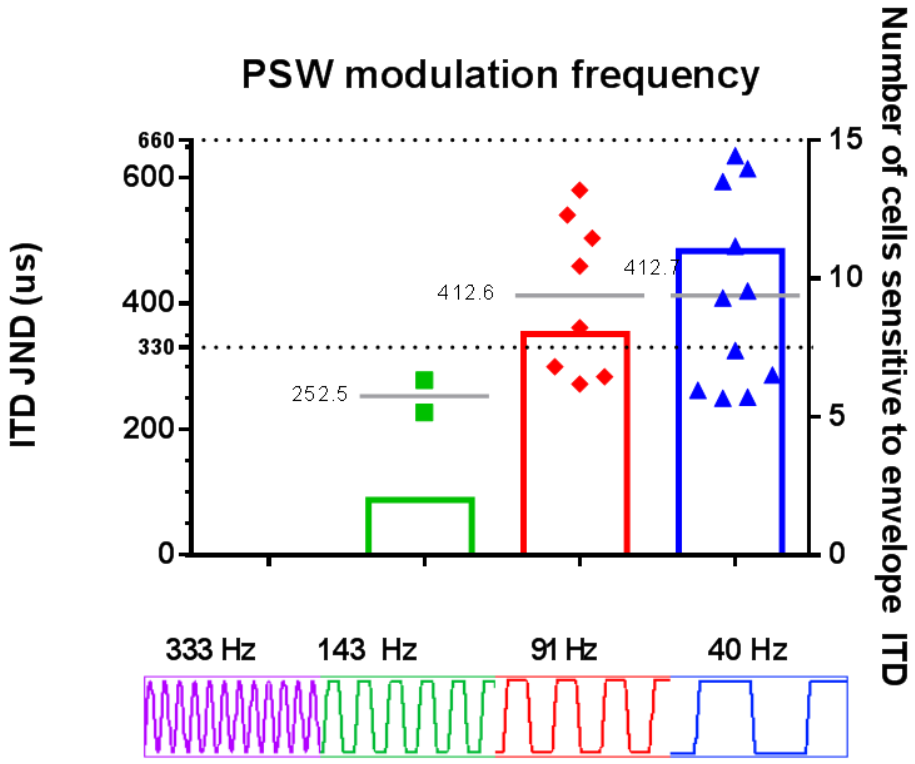


Figure 3.21 ITD JNDs for neurons that are sensitive to PSW modulation frequencies of 333, 143, 91 and 40 Hz when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Attack Duration

Figure 3.22: The ITD JNDs for neurons that are sensitive to envelope shapes with Attack durations of 1.5, 5 and 12.5 ms. In response to the 1.5 ms Attack duration (the Damped envelope shape) 14 (20%) neurons elicit an ITD JND. An increase in Attack duration to 5 ms (parameter 18) and 12.5 ms (parameter 17) results in a reduction to 5 and then 1 sensitive neuron respectively. 71% of the ITD JNDs of the current neuron population in response to the Damped envelope shape are within a relatively narrow range between 400 and is 550 µs. The 5 neurons sensitive to parameter 18 are spread across a wide range between 120 and 600 µs while the 1 neuron sensitive to parameter 17 has a relatively large ITD JND of 627 µs.

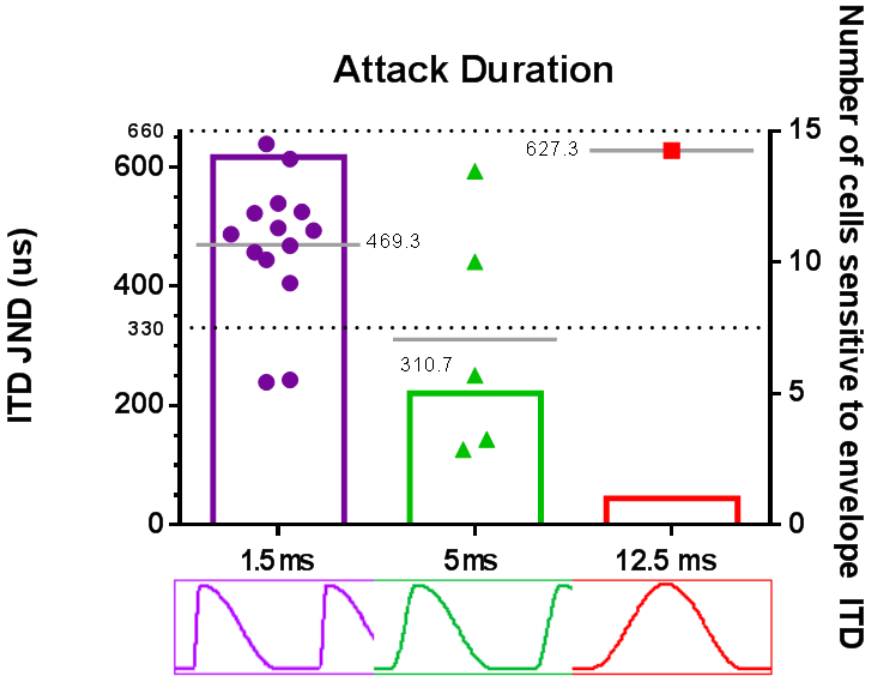


Figure 3.22 ITD JNDs for neurons that are sensitive to Attack duration of 1.5, 5 and 12.5ms when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Ramped vs Damped

Figure 3.23: The ITD JNDs for neurons that are sensitive to the temporally asymmetric envelope shapes of the Damped and Ramped envelope shape, the results of which have been described in detail in Section 3.3.2 – *Neuron population characteristics – Damped vs Ramped*. The Attack and Decay duration is either 1.5 ms or 15 ms for each parameter. In response to the Damped envelope shape that has a 1.5 ms Attack and 15 ms Decay (parameter 7) 14 (20%) neurons elicit an ITD JND with a mean of 469 µs. Switching the Attack and Decay durations (parameter 8) results in just one neuron having an ITD JND for the same threshold criteria. This single neuron has an ITD JND of 260 µs which is at the lowest end of the range of ITD JNDs elicited by the Damped envelope shape.

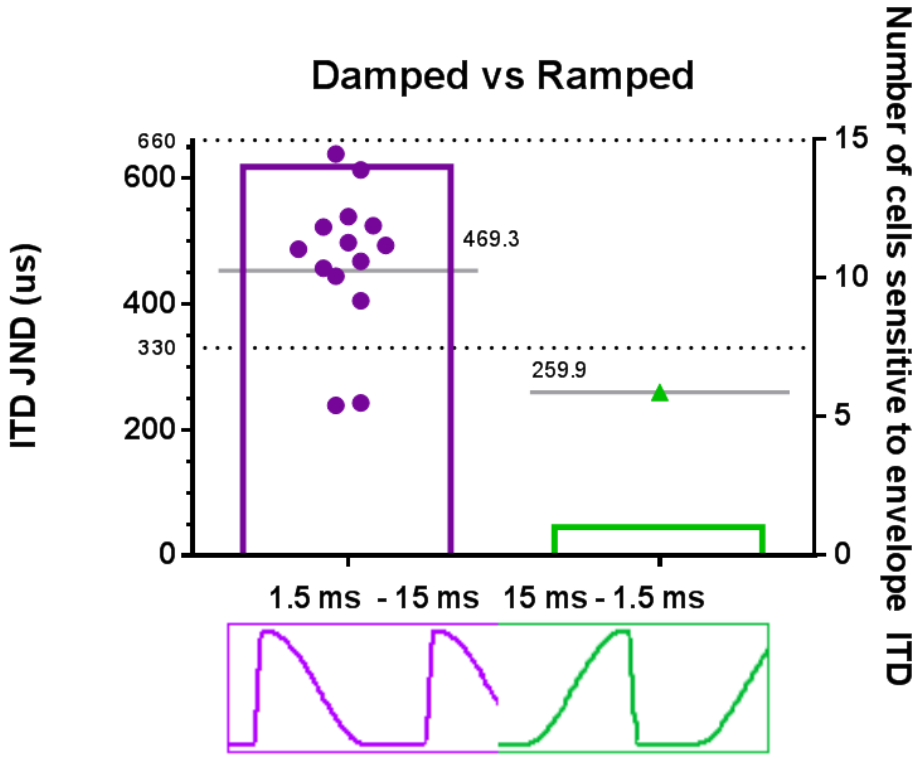


Figure 3.23 ITD JNDs for neurons that are sensitive to Damped and Ramped parameters when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Pause Duration i)

Figure 3.24: The ITD JNDs for neurons that are sensitive to Pause durations of 0, 4, 8 and 18 ms. The Pause duration separates PSW modulations that have a constant Sustain duration of 4 ms for all parameters. By varying the Pause duration and keeping the Sustain constant, the duty cycle is varied and results in a reduction in frequency modulation from 143 Hz to 40 Hz as the duty cycle decreases. Threshold criteria of both D = 2 (Figure 3.24A) and D = 1 (Figure 3.24B) were used to determine IPD JNDs. Figure 3.24A: In response to the PSW with 0 ms Pause (parameter 14) only one neuron had an ITD JND within the set limits. Increasing the Pause duration resulted in a monotonic increase in the number of sensitive neurons. A 4 ms duration (parameter 13) resulted in 8 (11%) sensitive neurons, an 8 ms duration (parameter 15) resulted in 11 (15%) sensitive neurons and an 18 ms duration (parameter 6) elicited the largest number of ITD JNDs from this neuron population with 21 (30%) producing a response within the set criteria. The range of ITD JNDs and the mean responses are fairly similar for both parameter 13 and parameter 15. All ITD JNDs occur between approximately 250 and 600 µs and the mean ITD JNDs are 413 µs and 452 µs respectively. As well resulting in the greatest number of sensitive neurons, parameter 6 also elicited the largest range of ITD JND values (between approximately 140 and 640 µs) and the lowest mean ITD JND at 355 µs. In Figure 3.24B the mean ITD JND is reduced for each parameter by between 100 µs 200 µs compared to Figure 3.24A although the mean ITD JND is relatively similar across parameters. The best ITD JNDs recorded for each parameter is improved with an increase in Pause duration.

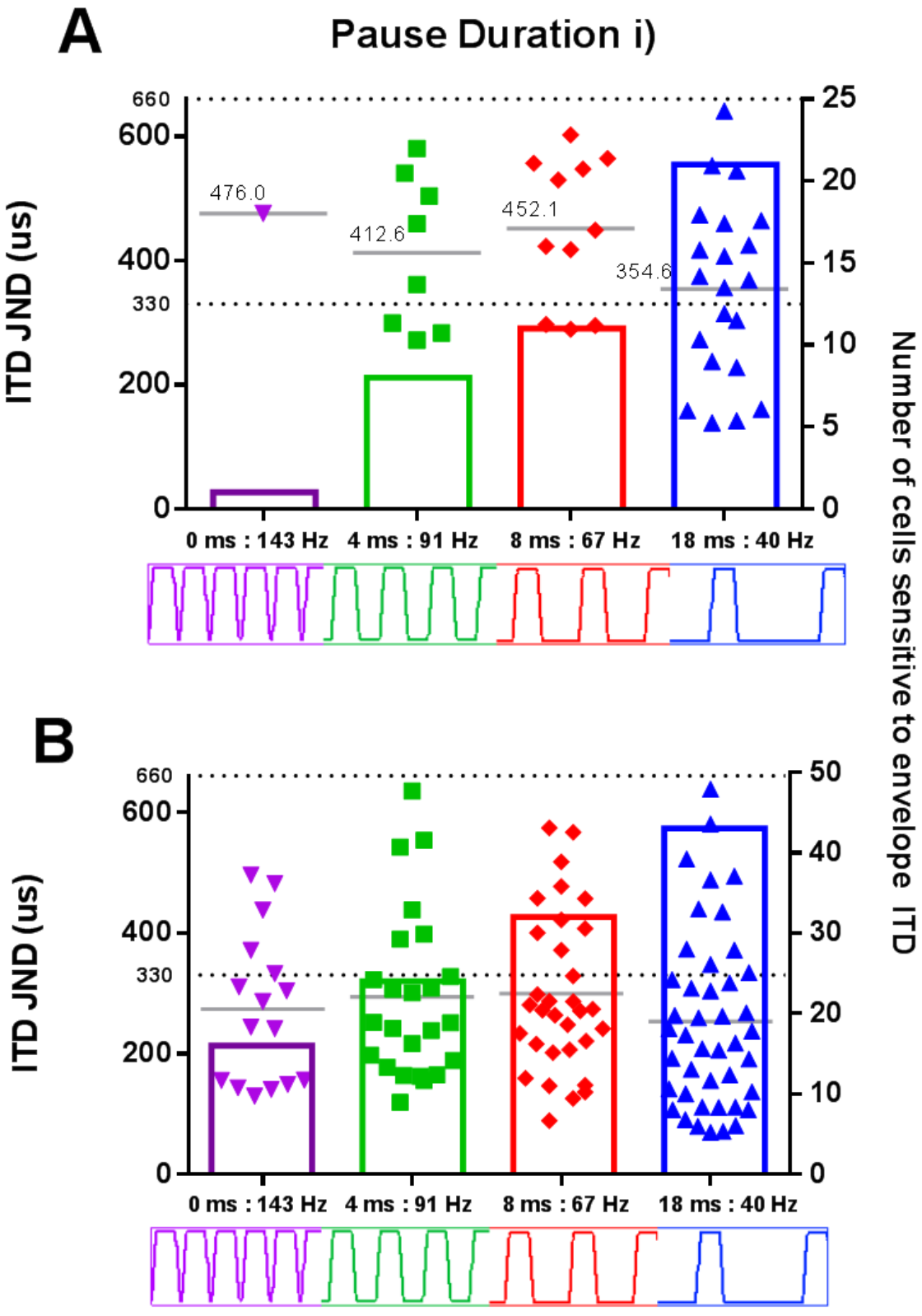


Figure 3.24 A) ITD JNDs for neurons that are sensitive to Pause durations of 0, 4, 8 and 18 ms when D = 2 and B) D = 1. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Pause Duration ii)

Figure 3.25: The ITD JNDs for neurons that are sensitive to Pause durations of 0, 4, 8, 11 and 18 ms. The duty cycle is kept constant within this comparison by varying the Sustain duration such that the frequency modulation is maintained at 40 Hz for all parameters within this comparison. In response to the PSW with 0 ms Pause (parameter 3) only one neuron had an ITD JND within the set limits. Increasing the Pause duration again resulted in a monotonic increase in the number of sensitive neurons. A 4 ms duration (parameter 4) resulted in 4 (6%) sensitive neurons, an 8 ms duration (parameter 5) resulted in 5 (7%) sensitive neurons, an 11 ms duration (parameter 2) resulted in 11 (15%) sensitive neurons and the 18 ms Pause duration elicited the largest number of ITD JNDs from this neuron population with 21 (30%) producing a response within the set criteria. The range of ITD JNDs in response to each parameter appears to increase as the Pause duration is increased as seen by the responses to parameter 4 (4 ms Pause) having a range of approximately 200 µs increasing to a range of approximately 500 µs for parameter 6 (18 ms Pause). The mean responses are similar for parameters 4, 5, 2 and 6 occurring within a relatively narrow range between 335.5 and 412.7 µs. The single ITD JND found in response to parameter 3 is equivalent to the smallest ITD JNDs elicited by the other parameters at 163 µs.

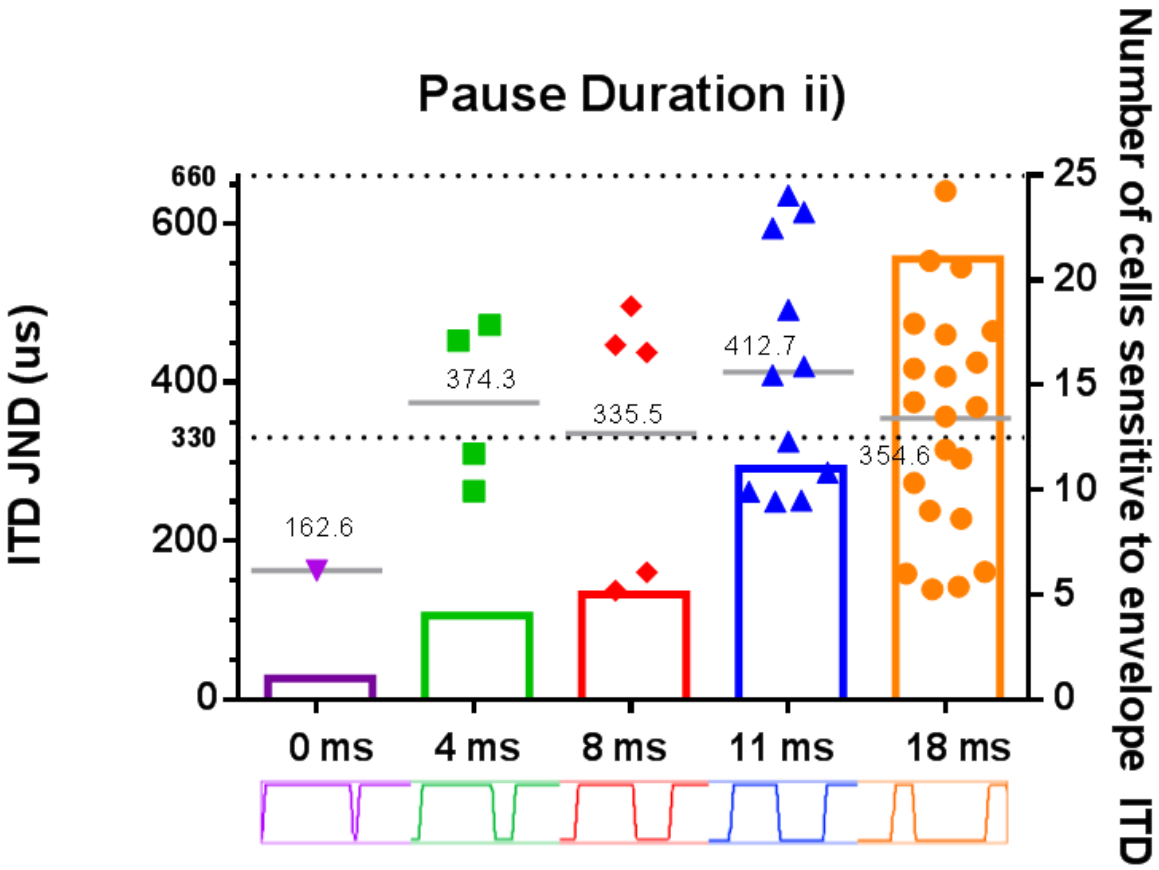


Figure 3.25 ITD JNDs for neurons that are sensitive to Pause durations of 0, 4, 8, 11 and 18 ms with a constant duty cycle when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Pause Duration iii)

Figure 3.26: The ITD JNDs for neurons that are sensitive to a SAM tone with a modulation frequency of 40 Hz and an equivalent envelope modulation shape with the addition of an 8 ms Pause duration that results in a 26 Hz modulation frequency. 3 neurons had ITD JNDs smaller than 660 µs for ITDs carried in the envelope of the SAM tone (parameter 1) while only 1 neuron had an ITD JND within the set criteria when the 8 ms Pause duration was included (parameter 17). The ITD JNDs of the 3 neurons within the 0 to 660 µs range for parameter 1

were spread across a relatively large range of 550 µs. The 1 neuron responsive to parameter 17 had an ITD JND equivalent to the largest elicited by parameter 1 at 627 µs.

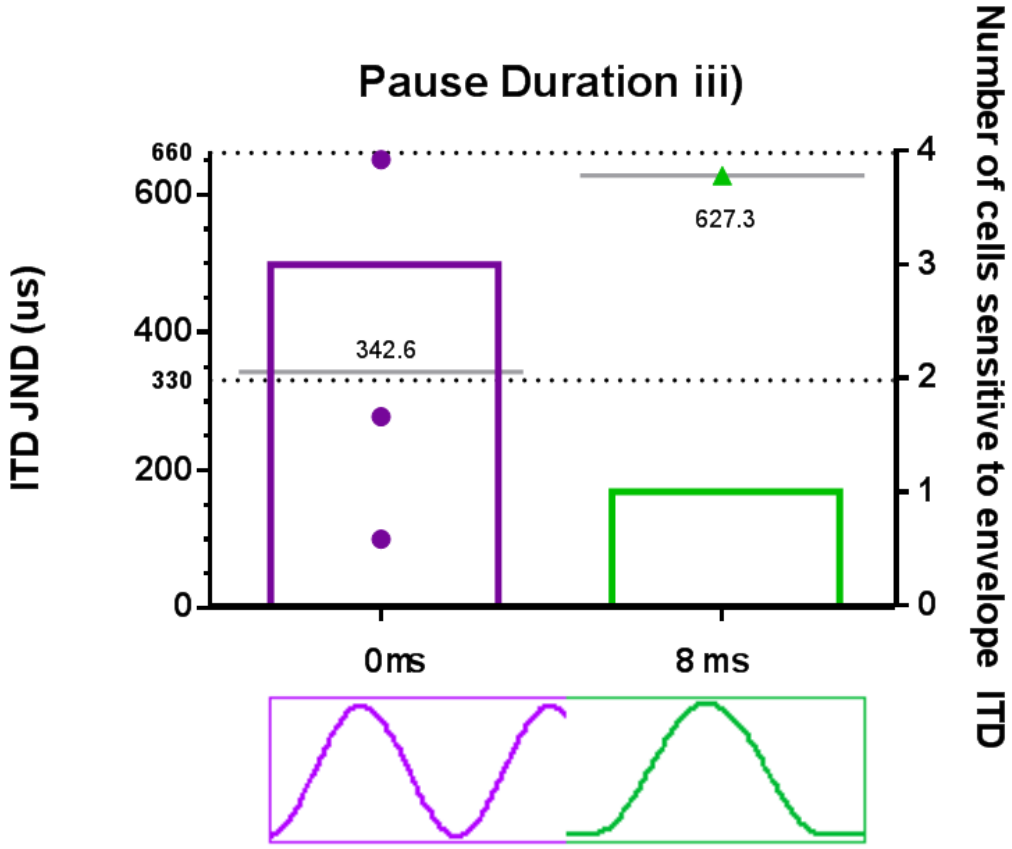


Figure 3.26 ITD JNDs for neurons that are sensitive to a 40 Hz SAM tone and an equivalent envelope modulation shape with the addition of an 8 ms Pause when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

Pause Duration iv)

Figure 3.27: The ITD JNDs for neurons that are sensitive to a SAM tone with a modulation frequency of 333 Hz and an equivalent envelope modulation shape

but with the addition of an 8 ms Pause duration that results in an 87 Hz modulation frequency. No neurons had ITD JNDs smaller than 660 µs for ITDs carried in the envelope of the SAM tone (parameter 9) while 16 (23%) neurons had an ITD JND within the set criteria when the 8 ms Pause duration was included (parameter 16). The ITD JNDs of the responsive neurons for parameter 16 were spread across a range between approximately 150 and 580 µs with a mean of 341 µs.

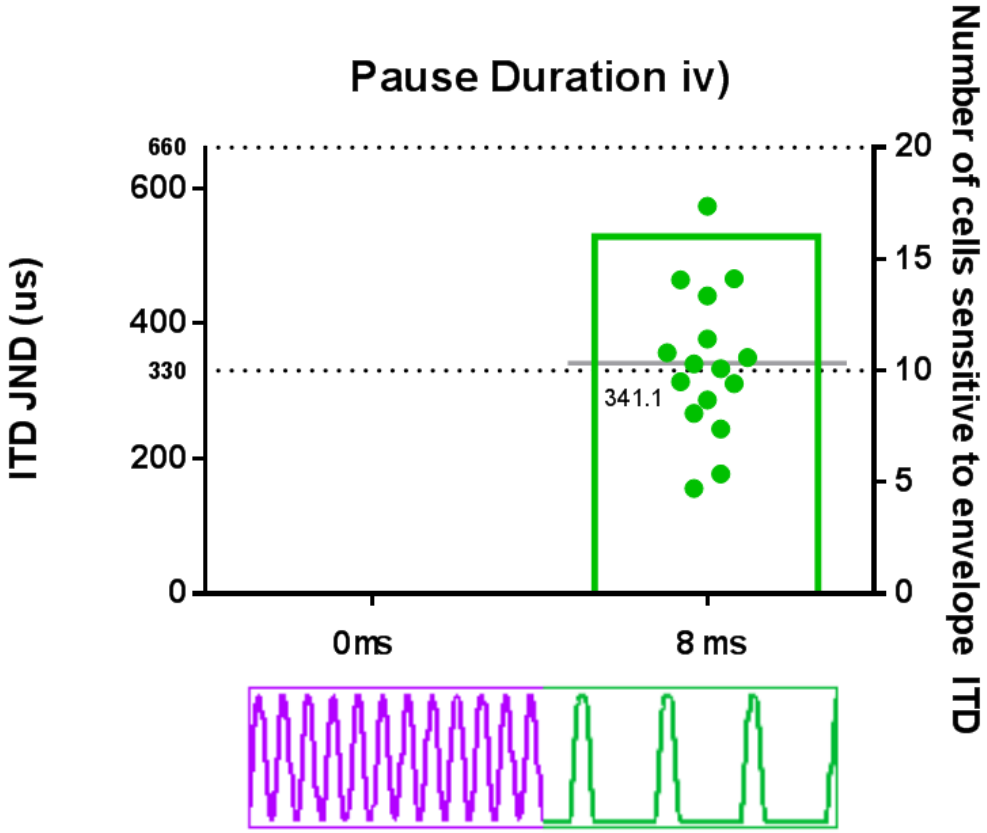


Figure 3.27 ITD JNDs for neurons that are sensitive to a 333 Hz SAM tone with and an equivalent envelope modulation shape with the addition of an 8 ms Pause when D = 2. Grey lines and number = mean ITD JND. Dashed line = 330 µs physiological range and 660 µs criteria for inclusion. Coloured bars plot number of sensitive neurons (right ordinate)

3.4 Discussion

The primary objective of the electrophysiology experiments presented in this thesis was to investigate the effect that manipulating the Attack, Decay, Pause and Sustain segments of the envelope of a high-frequency sound has on the response of inferior colliculus neurons. The responses of four example neurons have been presented in detail. Each of the neurons demonstrated the ability to phase-lock to the envelope of high-frequency sounds as seen in the raster plots and the spike distributions across the modulation cycle duration whereby the majority of spikes occur at periodic intervals. With regards the raster plots, phase-locking is expressed as a high number of tightly grouped spikes at periodic intervals, linked to the envelope modulation frequency. The degree to which phase-locking took place is observed as superior for the PSW when compared to the SAM tone as represented by a greater number of spikes occurring over a narrower range of the modulation cycle, in agreement with existing literature on neural responses phase-locking to the envelope of high frequency stimuli (Yin, Kuwada et al. 1984; Griffin, Bernstein et al. 2005; Dreyer and Delgutte 2006).

The ITD-rate-functions in response to the eighteen different envelope shapes were presented as 8 comparisons for each example neuron and produce a complex response configuration unique to each neuron. The individual neurons display sensitivity to changes in modulation frequency whereby at high modulation rates neurons no longer display envelope based ITD sensitivity. This is likely due in part to the reduction in the phase-locking to high modulation frequencies originating peripherally (Dreyer and Delgutte 2006). A further response trend observed across neurons is the greater ITD sensitivity to envelopes with longer Pause durations. This finding is in line with the assumption that the Pause prior to the Attack allows for a recovery of the

sensitivity of auditory nerve fibres and that the recovery has a short time constant (Klein Hennig, Dietz et al. 2011).

The neural responses in the PSTHs present a maximum at stimulus onset and although the focus of the data analysis in this thesis has been on the ongoing response, this immediately indicates the importance of the Attack segment of the stimulus and suggests that the Attack would produce the greatest ITD sensitivity due to its inherent connection to the stimulus onset. We have confirmed this by the comparison of Rate-ITD-functions for envelopes with very short and very long Attack segments whereby a short Attack segment results in far more ITD sensitive neurons as well as smaller ITD JNDs than does the longer Attack segment.

Utilising the envelope shapes of the present study it has been possible to assess how the output at the level of high-frequency ANFs, in terms of temporal firing pattern, affects ITD sensitivity. The enhanced ITD sensitivity observed for the PSW, characterised by relatively large changes in spike rate across ITDs, is proposed to be a result of the temporal pattern of action potentials firing in high-frequency ANFs that originate from the temporal signature of the stimulus envelope (Griffin, Bernstein et al. 2005). This effect can be seen most clearly in the raster plots (Figure 3.1 – 3.10 G-J) and the frequency distribution plots relating spike rate and modulation phase (Figure 3.1 - 3.10 D). The present data indicates that phase-locking is enhanced in response to the PSW envelope shape compared to the SAM tone, and for the Damped envelope shape compared to the Ramped , as the majority of spikes are observed to occur over a narrower range of the modulation cycle. In order to understand the factors that may be contributing to this effect we can consider a scenario whereby the activity of a given neuron matches precisely the shape of a PSW and SAM tone of equivalent modulation frequency. In this scenario the phase-locking would be found to be greatest in response to the PSW as the amplitude of the PSW reaches a

maximum over a shorter time period. This in turn results in an increased probability of a neuron firing during a specific, narrower time period compared to a response occurring during the slow Attack phase of the SAM tone. In this case a neuron will reach its internal threshold criteria for firing at a time that may be less well defined. This effect on phase-locking may underlie the results found in the psychophysical studies comparing SAM tones to transposed-tones (Bernstein and Trahiotis 2002). The Pause duration effect as observed in the neuron population analysis (Figure 3.24) extends the parameters of the SAM tone, PSW or transposed-tones and was conducted in response to the human psychophysical findings of Klein Henning et al. (2011). The responses to envelope shapes with long Pause durations between waveform modulations presented here demonstrate a greater change in spike response rate across different ITDs when compared to those envelopes containing short Pause durations. This observation is in agreement with the assumption that a Pause segment that precedes an Attack segment allows for a recovery of sensitivity of neurons within the auditory pathway and that increasing the pause duration from 0 to 18 ms can result in an improvement in ITD sensitivity

As well as phase-locking and sensitivity to ITDs being greater in response to envelopes with short Attack and long Pause durations, the individual neuron data shows a reduction in ITD sensitivity with increasing modulation rate, consistent with previous electrophysiology studies (Yin, Kuwada et al. 1984; Griffin, Bernstein et al. 2005). The suggestion that reduced ITD sensitivity with increasing modulation frequency is a result of the attenuation of spectral components by the limited width of the corresponding peripheral filters was brought in to question by the use of the Ramped and Damped envelope shapes. These two parameters have equal power spectra but temporally asymmetric envelope shapes and result in ITD JNDs that differ by a factor of about 4 (Klein Hennig, Dietz et al. 2011). The response of auditory neurons is known to be greatest at stimulus onset and to be followed by variable levels of adaptation

(Smith 1977; Smith 1979) and so the effect of the Attack segment is not unexpected. The shorter first spike latency found for the Damped compared to the Ramped stimulus is also in line with earlier studies that found shorter first-spike latencies in response to shorter rise times in the cat primary auditory cortex (Heil 2001).

The population characteristics of the 71 neurons that were assessed reveal a stimulus specific first spike latency that is on average 10 ms shorter for the Damped stimulus compared to the Ramped. The normalised ITD-rate-functions reveal that the Damped stimulus elicits the greatest change in spike rate across the physiological range and the assessment of neuronal ITD JNDs reveals that the mean ITD JND is lower in response to the Damped stimulus, again consistent with the psychophysical findings of Klein Hennig et al. (2011). From the total population of neurons from which ITD sensitivity data was collected, the number of neurons sensitive to envelope ITDs increased when the Pause duration segment of the envelope was increased. The data presented are also consistent with the enhanced ITD sensitivity generated by transposed-tones observed in psychophysical data with human listeners (Bernstein and Trahiotis 2002; Bernstein and Trahiotis 2009). More neurons were sensitive to ITDs within the envelope of high frequency sounds that had shorter, and therefore steeper, Attack segments. Firing rates were more modulated as a function of ITD, and ITD JNDs were consistently lower in response to the PSW than to SAM tones.

The data presented in this thesis contributes to the ongoing discussion of what the neural correlates of sound localisation are. By assessing the individual neural responses to envelope ITDs it has been possible to assess the envelope segments of modulated signals that affect an individual neurons response including its ITD sensitivity as well as the response across a sample population of neurons within the inferior colliculus. By comparing the ITD JNDs of neurons to both the

Ramped and Damped stimulus it has been possible to provide evidence for three factors in particular that may be directly connected to the behavioural findings of psychophysical studies that have previously explored responses to very similar stimuli such as in the work of Klein-Hennig et al. (2011). These three factors can be most clearly observed in Figure 3.18B and Figure 3.23 that present the ITD JNDs from all 71 ITD sensitive neurons assessed in this thesis in response to the Ramped and Damped stimulus. Figure 3.18B evaluates ITD JNDs that were obtained from the 'fine' recordings covering the ±2 ms envelope ITD range while Figure 3.23 displays the ITD JNDs of neurons that were ITD sensitive within the range of ±330 µs. From the data presented the possible contributors to the underlying neural correlate of sound localisation are proposed to be i) the mean population ITD JND ii) the size of the ITD sensitive neuron population and iii) the ITD JND of the best performing neurons. The known behavioural ITD JNDs in response to similar envelope shapes, whereby ITD JNDs are worse by a factor of about 4 for the Ramped stimulus compared to the Damped stimulus (Klein Hennig, Dietz et al. 2011) , makes it is possible to explore these potentially important neural correlates for sound localisation. These correlates are found in the data as follows; The mean population ITD JND in Figure 3.18B is a found to be smaller in response to the Damped stimulus than the Ramped stimulus by a factor of approximately 1.5 when D = 2. The mean population ITD JND does however decrease for both the Ramped and the Damped envelope shape when D = 1 to mean ITD JNDs that a very similar between the two although the ITD JND does remain smallest in response to the Damped stimulus. These relatively small differences between population mean ITD JNDs compared to the behavioural thresholds suggests that this may not be the driving factor underlying ITD sensitivity but could be either a contributing factor as the trend is consistently in same direction as that observed behaviourally or could be present only as a result of the true neural activity that drives sound localisation. The next factor proposed to contribute to the behavioural findings is the size of as observed in Figure 3.18B. The ITD sensitive neuron population is larger for

the Damped stimulus than the Ramped stimulus by a factor of approximately 3.3 when D = 2 whereby 30 neurons are sensitive to the Damped stimulus while only 9 are sensitive to the Ramped stimulus. When D = 1 the difference factor decreases to 1.4. When considering only those neurons sensitive within the narrower range of ITDs of ±330 μs, which may more be ecologically relevant, when D = 2, 14 neurons are found to be sensitive to the Damped stimulus with only 1 sensitive to the Ramped stimulus. This suggests that the total number of neurons that are ITD sensitive to a particular envelope shape does contribute to the behavioural sensitivity thresholds. Lastly, considering the ITD JNDs of the most sensitive neurons we find neural thresholds equivalent to those obtained behaviourally. In Figure 3.18B for example, when D = 2, the best neuron in response to the Damped stimulus has an ITD JND of approximately 150 μs while in response to the Ramped stimulus the best neuron has an ITD JND of approximately 450 μs. As a first approximation this is remarkably similar to the findings of the behavioural study using the equivalent envelope shapes and suggests the possibility of a significant role for a relatively small number of the most sensitive neurons for a given stimulus in sound localisation.

4 General Discussion and Conclusion

In order to contribute to the current understanding of ITD sensitivity two complimentary approaches have been utilised in order to provide a foundation of behavioural and neurophysiological data that provides evidence for the possible mechanisms involved in sound localisation. The first part of this project was a psychophysical lateralisation study in which were measured the just-noticeable-differences (JNDs) of ITDs as a function of a reference ITD for human subjects in a manner similar to that of Bernstein et al. (2002). The aim was to establish the smallest difference in interaural phase that a subject could identify between tones presented to the left and right ear for ten reference IPDs at five different frequencies and how the human sensitivity varied when the IPD was conveyed in either the fine-structure or the envelope structure of an acoustic signal. These experiments resulted in data that suggest fundamental differences in the processing of ITDs conveyed in the fine structure and envelope structure of acoustic signals and provided a sound justification for a deeper exploration of how envelope ITDs are processed at the neural level.

The second part of this thesis consisted of a series of neurophysiological experiments in which single cell recordings from the guinea pig inferior colliculus were made in response to eighteen envelope shapes that varied in their duration of their Pause, Attack, Sustain, and Decay as previously undertaken psychophysically by Klein Hennig et al. (2011) in order to establish their importance in driving ITD sensitivity as well as establishing what the potential neural correlates for sound localisation are in order to provide targets for further research and analysis.

The importance of understanding the underlying neural mechanisms of binaural processing and human performance in spatial listening tasks was the most significant driving factor in undertaking the work contained within this thesis.

Although extensive physiological data exists, most models of human ITD processing are based upon early models of cross correlation and the interpretation of behavioural data is largely based on the assumptions of these models. Models of binaural hearing that are built upon physiological studies are critical to the interpretation of new psychophysical data that are not fully described by existing models including the rapid decline in sensitivity to envelope based ITDs for modulation frequencies above 512 Hz. In working to understand the link between single neuron recordings and human perception of ITDs there is the possibility of being able to apply this knowledge to the development of bilateral cochlear implantation, particularly with regard to the role of the temporal fine structure and the envelope. Current cochlear implant digital sound processing strategies will often provide envelope cues to high-frequency auditory channels but it is possible that the manner in which this is performed can be optimised based upon the present findings of the ideal envelope modulation shapes that have been found to enhance ITD sensitivity.

The ability to detect sounds in noisy environments is greatly enhanced by being able to assess the similarity of the sound at the two ears and normal human listeners are exquisitely sensitive to these interaural differences, relying on them to hear and understand speech at low signal-to-noise ratios. Understanding brain mechanisms contributing to the detection of these interaural differences can have a positive effect on the functionality of bilateral CIs that currently provide independent stimulation patterns at the two ears, likely reducing the ability to exploit interaural differences required for spatial listening.

The present psychophysical and physiological experiments indicate sensitivity to ITDs in the envelopes of high-frequency modulated sounds. It is somewhat remarkable, given the prior knowledge of binaural cues for sound localisation that under certain parameter conditions temporal information in the envelopes of high-frequency modulated tones can be equivalent to temporal information

normally present in the output of low-frequency channels. The first example of this that we explored was the tranposed-tone but in exploring the possible neural mechanisms involved, and with effort applied to optimising envelope shapes, the Damped envelope shape has been found to also offer very good ITD sensitivity in both human subjects, and now in single neurons within the IC. Low modulation frequency transposed-tones result in ITD discrimination thresholds equivalent to low-frequency tones for locations on the mid-line (Bernstein and Trahiotis 2002) and in the present experiment we were able to show that it is even possible for envelope based ITD discrimination to go beyond this at locations off-midline. The enhanced ITD sensitivity has also been shown in the responses of inferior colliculus neurons of the guinea pig (Griffin, Bernstein et al. 2005). Given the importance of envelope based ITDs to cochlear implant users as, in many cases, they are the only form of ITD information available, a result of the stimulus fine-structure being discarded under standard cochlear implant sound processing strategies, the importance of optimising the available cues by adjusting the Pause and Attack segments of the modulation envelope may be a worthwhile route for exploration. Further understanding of the brain mechanisms that contribute to performance in envelope-ITD coding may therefore be vital to future bilateral cochlear implant processing strategies. This is supported by recent work by Laback et al. (2011) who compared the abilities of cochlear implant users and normal hearing subjects in discriminating envelope based ITDs. As expected they found large differences in the ability to discriminate ITDs between the two groups but encouragingly, the improvements resulting from envelope shape adjustments were found in both groups. Combining the existing psychophysical and electrophysiological data may therefore improve the processing of envelope ITDs to the point that cochlear implant users can begin to see improved outcomes in more challenging listening environments.

There is no definitive description of the exact neural mechanisms that underlie the difference in ITD JND as a result of the different envelope shapes of the Ramped and Damped parameters as was observed in both the human psychophysical and electrophysiological studies. The data presented can however offer some insight in to the components that are likely to be involved and, ultimately, may lead to improvements in communication for those who are most in need.

The data presented in this thesis provides a foundation for future research in the field of cochlear implantation processing strategies. Combining the findings of existing behavioural studies that use normal hearing subjects as well as our current understanding of sound localisation cues and the neural correlates that may underlie them, an approach can be developed that combines both behavioural thresholds from cochlear implant users with optimised electrical-pulse modulations with a focus on the signal onset, Attack and Pause segments of the envelope modulations in order to develop a clinical application for use in cochlear implants. This approach could be combined with a parallel programme measuring neural responses to direct electrical stimulation of the auditory nerve recorded in a similar manner to the present electrophysiology experiments in order to gain a further understanding of the digital processing needs for improved hearing with cochlear implants. This approach to further research provides the possibility of optimisation of the electrical stimulation that human cochlear implant users have access to for the purpose of improving sound localisation abilities and speech discrimination in noise for cochlear implant users, currently two of the greatest challenges that cochlear implant users experience. The neural recordings using animal models for cochlear implantation would provide further insight in to the difference that acoustic and electrical stimulation has on neural activity while a further application could assess how the deafened auditory brain processes envelope-optimised stimulation as those discussed in this thesis. The manner in which a signal

envelope can be optimised as discussed presently would be well applied to the processing of acoustic signals such as speech, given that speech is in effect an envelope modulated waveform. It is possible therefore that further research into applying a digital processing strategy that provides Attack and Pause segment optimisation to the extracted envelope of an acoustic signal may provide the opportunity to discover if this approach has the potential to alleviate some of the greatest challenges currently experienced by cochlear implants users.

Appendix

I. *Envelope Shapes and corresponding frequency spectrums*

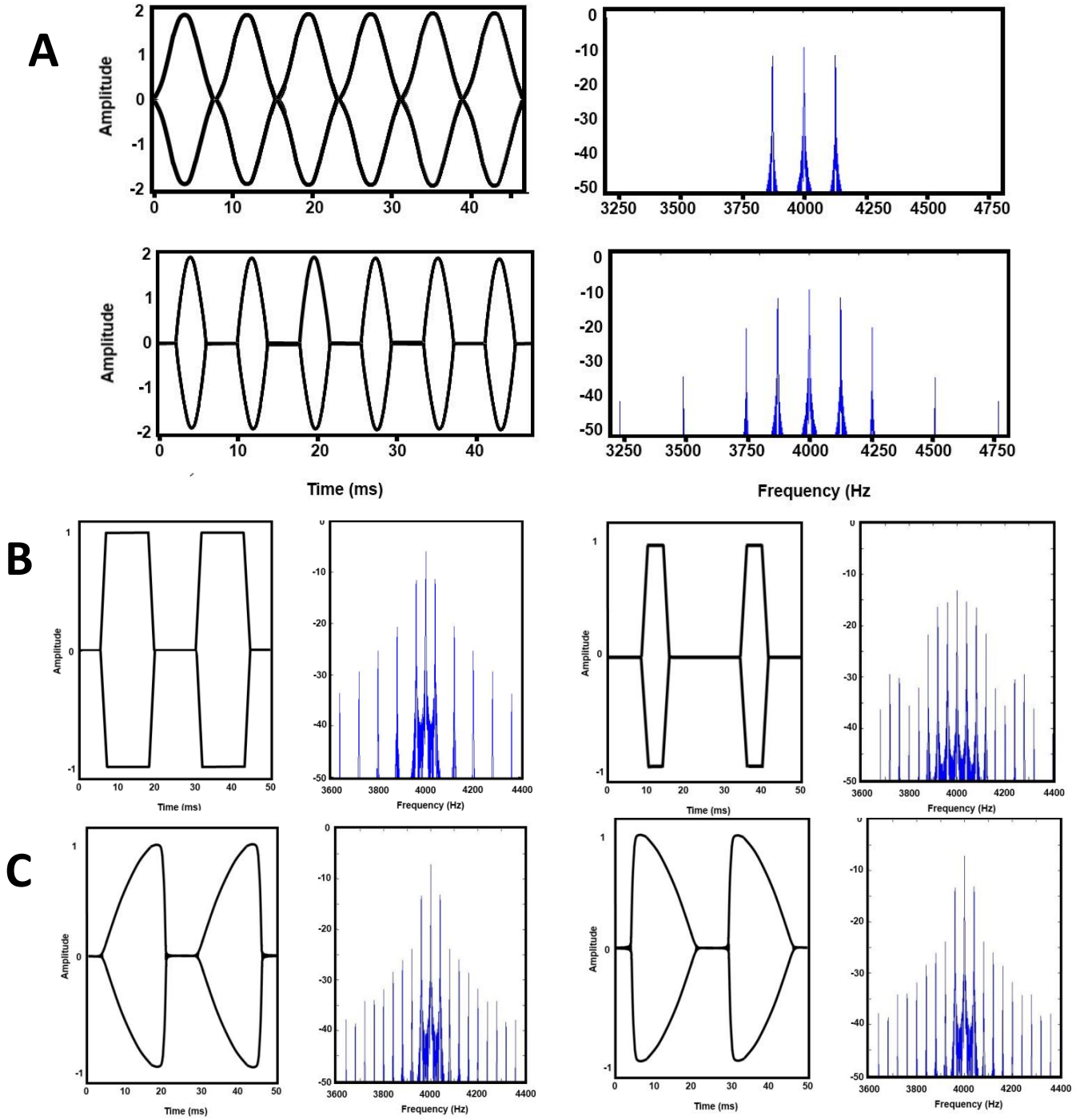


Figure 3.28 A) Top row: 50 ms segment illustration of 128 Hz SAM tone envelope (left) and frequency spectrum (right). Bottom Row: 50 ms segment illustration of 128 Hz transposed-tone envelope (left) and frequency spectrum (right). B) Envelope illustration of 40 Hz PSW with 11 ms Sustain duration and 11 ms Pause duration (left) and 4 ms Sustain duration and 18 ms Pause duration (right) with accompanying frequency spectrums. C) Envelope illustration of 40 Hz Ramped parameter with 15 ms Attack duration and 1.5 ms Decay duration (left) and Damped parameter with 1.5 ms Attack duration and 15 ms Decay (right) with accompanying frequency spectrums.

II. Equal Loudness

Due to the known equal loudness curves for normal human hearing, a sound presented at 75 dB SPL will have a subjective loudness dependent on the frequency at which it is presented. Loudness levels are measured in phons such that pure tones of different frequencies perceived by a normal-hearing individual as having the same loudness have the same phon level. In order to ensure that changes to IPD threshold are due to the stimulus IPD and not ILDs, all sounds presented were at a constant value of 75 dB HL, ensuring that the perceived loudness of all stimuli, regardless of frequency, would be equal.

A 64 Hz and a 512 Hz tone, both presented at 75 dB SPL, will be perceived as being approximately 35 dB different (42 and 77 phons respectively). The headphone output for each test frequency was adjusted using a measuring amplifier to ensure that all sound levels would be perceived at 75dB HL.

An increase in 40 phons at 512 Hz has little effect on IPD JNDs at both 0° and 60° reference IPD. At 64Hz the effect that an increase of 29 phons has is again small, relative to the threshold change between 0° and 60° reference IPD. Zwislocki (1956) observed that at 500Hz, an increase from 50 dB SL to 110 dB SL had very little effect on threshold IPD.

III. Masking noise

Amongst studies that demonstrate a subjects ability to detect ITDs in the envelope of complex high frequency stimuli, Henning (1974) was one of the first to show that temporal information in the envelope of complex stimuli can be used for the lateralisation of amplitude modulated stimuli. Even when the carrier frequencies are higher than those thought to be linked to ITD based localisation (Rayleigh 1907). It was important therefore to first establish that the

lateralisation abilities observed in these experiments was indeed based on temporal information in the high frequency channels and not due to non-linear distortions producing energy in lower frequency channels. These distortions would occur at frequencies equal to, and multiples of, the modulation rate (Nuetzel and Hafter 1976) and could be caused by nonlinearities in the equipment, or in middle ear or basilar membrane mechanics. Henning (1974) initially addressed this using a high intensity, low-frequency masking noise to mask any aural harmonic components below 600 Hz. The study found that the low-frequency masking noise did not disrupt the lateralisation of high frequency complex waveforms suggesting that low-frequency information in the waveform was not necessary for lateralisation.

Nuetzal and Hafter (1976) masked the information carried in the first five harmonics of the modulator and all of their subjects could still complete their lateralisation tasks. In doing so they verified the finding that the high frequency channels are utilised and ruled out any essential role for distortions at higher multiples of the modulation frequency. Later studies, including Bernstein and Trahiotis (2002), employed a continuous diotic noise low pass filtered at 1300 Hz for the high frequency stimuli in order to prohibit a subjects' use of information at low spectral frequencies.

These earlier studies used SAM tones with varying carrier and modulation frequencies. The present psychophysical tests utilise a 4 kHz carrier frequency with modulations up to 512 Hz. The effect of a diotic noise, low pass filtered at 1300 Hz was tested on the IPD discrimination of a 512 Hz modulated transposed-tone. This is where the available psychophysics literature points to a departure between pure tone and transposed-tone ITD thresholds.

IV. Four AFC

Leek (2001) discussed how the time saving with a 2AFC protocol converging at the 71% correct level is less efficient than a 3AFC protocol converging at the 79% correct level. I have chosen a three-interval, two AFC protocol combined with a three-down one-up procedure converging at the 79% correct level for the present study. Studies observing similar psychophysical effects have used two AFC protocols with two cues (Domnitz and Colburn 1977; Bernstein and Trahiotis 2002). In order to establish which protocol was best suited to the present task the effect on IPD thresholds when using a four-interval two AFC protocol and a three-interval, two AFC protocol was measured.

The four-interval task produced less sensitive thresholds for both the 256Hz and 512Hz stimuli and for reference IPDs of both 0° and 60°. These results combined with the time saved using the three-interval protocol meant that this was adopted for the present psychophysics experiments.

V. Psychophysics Pilot Study

The ranges of pure tone frequencies, modulation frequencies and reference IPDs used in the psychophysics experiments were selected by means of a series of pilot studies. The pilot studies revealed the range of reference IPDs and test frequencies that would allow for detailed analysis of the test subject's IPD sensitivity without assessing every possible reference IPD and test frequency. The pilot studies enabled test parameters to be selected for use in the final test protocol that maximized the efficiency and accuracy of data collection. 32 reference IPDs were explored in the pilot studies; ± 0, 22.5, 30, 45, 60, 67.5, 90, 110, 112.5, 120, 135, 140, 157.5, 160, 170, 180. Ten reference IPDs were selected for the final test protocol ± 0, 30, 60, 90, 120. Above 120° IPD, for some subjects, there is a reversal of the sound lateralisation and so to maintain the focus of the study on the range of unambiguous cues for sound localisation, the upper limit for the reference IPD was kept to 120°. The reference IPDs selected from the pilot studies were equally spaced 30° IPD apart, with five IPD

values 120° IPD is covered, the range of reliable human IPD detection. JNDs were initially calculated in terms of IPD in order to facilitate the observation of changes in discrimination threshold across all of our test frequencies on a single linear axis as thresholds would remain within a limited range of 0-150° IPD. When converting IPD to ITD, a log scale is desirable in order to observe the thresholds across frequency with clearly resolved data points (Thresholds exist within the range of 0-10000μs). In testing left and right leading reference IPDs we can exclude any sidedness bias that may exist.

Existing data (Bernstein and Trahiotis 2002) show human abilities in detecting IPDs for TTs to be limited to modulation frequencies below 512 Hz, and to be most sensitive below 1300Hz for pure tones (Mills 1958). The pilot studies tested 14 pure-tone frequencies; 64, 128, 100, 200, 256, 400, 467, 512, 600, 733, 800, 866, 1000, 1024 Hz. Five pure tone frequencies were selected for the final test protocol; 64,128,256,512 and 1024 Hz. Five TT modulation frequencies were selected for the final test protocol; 32, 64, 128, 256 and 512 Hz.

The adaptive tracking procedure was refined in order to maximise accuracy and minimise time. The initial step size was large, allowing subjects to more rapidly reach an IPD close to their discriminability threshold. After this the step size was reduced and the average threshold at a particular number of step reversals taken as the threshold. Beginning with the average of the last 14 reversals, after the pilot studies the average of the last 8 reversals was found to be accurate, quicker and avoided raised threshold due to loss of a subject's attention. The final test protocol covered the range of IPDs and frequencies appropriate to human abilities as well as that of the existing literature (Mills 1958; Bernstein and Trahiotis 2002) thus facilitating data comparisons.

Bibliography

Arsenault, M. D. and J. L. Punch (1999). "Nonsense-syllable recognition in noise using monaural and binaural listening strategies." The Journal of the Acoustic Society of America **105**(3): 1821-1830.

Batra, R., S. Kuwada, et al. (1997). "Sensitivity to Interaural Temporal Disparities of Low- and High-Frequency Neurons in the Superior Olivary Complex. I. Heterogeneity of Responses." Journal of Neurophysiology **78**(3): 1222-1236.

Beckius, G. E., R. Batra, et al. (1999). "Axons from Anteroventral Cochlear Nucleus that Terminate in Medial Superior Olive of Cat : Observations Related to Delay Lines." The Journal of neuroscience **19**(8): 3146-3161.

Bernstein, L. R. (2001). "Auditory processing of interaural timing information: New insights." Journal of Neuroscience Research **66**(6): 1035-1046.

Bernstein, L. R. and C. Trahiotis (1994). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise." The Journal of the Acoustic Society of America **95**(6): 3561-3567.

Bernstein, L. R. and C. Trahiotis (1996). "On the use of the normalized correlation as an index of interaural envelope correlation." The Journal of the Acoustic Society of America **100**(3): 1754-1763.

Bernstein, L. R. and C. Trahiotis (2002). "Enhancing sensitivity to interaural delays at high frequencies by using transposed stimuli." The Journal of the Acoustic Society of America **112**(3): 1026.

Bernstein, L. R. and C. Trahiotis (2003). "Enhancing interaural-delay-based extents of laterality at high frequencies by using ``transposed stimuli"." The Journal of the Acoustical Society of America **113**(6): 3335-3347.

Bernstein, L. R. and C. Trahiotis (2009). "How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli." The Journal of the Acoustic Society of America **125**(5): 3234-3242.

Blauert, J. (1982). "Binaural localization." Scandinavian audiology. Supplementum **15**: 7-26.

Blodgett, H. C., W. A. Wilbanks, et al. (1956). Effect of Large Interaural Time Differences upon the Judgment of Sidedness. The Journal of the Acoustical Society of America. **28:** 639-643.

Bronkhorst, A. W. and R. Plomp (1988). "The effect of head-induced interaural time and level differences on speech intelligibility in noise." The Journal of the Acoustical Society of America **83**(4): 1508-1516.

Bronkhorst, A. W. and R. Plomp (1992). "Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing." The Journal of the Acoustical Society of America **92**(6): 3132-3139.

Butler, R. A. and K. Belendiuk (1977). "Spectral cues utilized in the localization of sound in the median sagittal plane." The Journal of the Acoustic Society of America **61**(5): 1264-1269.
Cant, N. B. and R. L. Hyson (1992). "Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil." Hear Research **58**(1): 26-34.
Carlile, S. and D. Pralong (1994). "The location-dependent nature of perceptually salient features of the human head-related transfer functions." The Journal of the Acoustic Society of America **95**(6): 3445-3459.
Carr, C. E. and M. Konishi (1990). "A circuit for detection of interaural time differences in the brain stem of the barn owl." The Journal of neuroscience **10**(10): 3227.
Cherry, E. C. (1953). "Some Experiments on the Recognition of Speech, with One and with Two Ears." The Journal of the Acoustical Society of America **25**(5): 975-979.
Clark, G. M. (1969). "The ultrastructure of nerve endings in the medial superior olive of the cat." Brain Research **14**(2): 293-305.
Colburn, H. S. (1973). "Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination." The Journal of the Acoustical Society of America **54**(6): 1458-1470.
Colburn, H. S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise." The Journal of the Acoustical Society of America **61**(2): 525-533.
Colburn, H. S. and N. I. Durlach (1978). "Models of binaural interaction." Handbook of perception **4**: 467-518.
Colburn, H. S. and P. Esquissaud (1976). "An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli." The Journal of the Acoustical Society of America **59**(S1): S23.
Deatherage, B. H. and I. J. Hirsh (1959). "Auditory Localization of Clicks." The Journal of the Acoustical Society of America **31**(4): 486-492.
Dietz, M., S. D. Ewert, et al. (2012). "Lateralization based on interaural differences in the second-order amplitude modulator." The Journal of the Acoustical Society of America **131**(1): 398-408.
Domnitz, R. (1973). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude." The Journal of the Acoustical Society of America **53**(6): 1549-1552.
Domnitz, R. H. and H. S. Colburn (1977). "Lateral position and interaural discrimination." The Journal of the Acoustical Society of America **61**: 1586-1598.
Dreyer, A. and B. Delgutte (2006). "Phase Locking of Auditory-Nerve Fibers to the Envelopes of High-Frequency Sounds: Implications for Sound Localization." Journal of Neurophysiology **96**(5): 2327-2341.

Durlach, N. I. (1963). "Equalization and Cancellation Theory of Binaural Masking-Level Differences." The Journal of the Acoustical Society of America **35**(8): 1206-1218.
Feddersen, W. E., T. T. Sandel, et al. (1957). Localization of High-Frequency Tones, The Journal of the Acoustic Society of America.
Friesen, L. M., R. V. Shannon, et al. (2001). "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants." The Journal of the Acoustical Society of America **110**: 1150.
Gardner, M. B. and R. S. Gardner (1973). "Problem of localization in the median plane: effect of pinnae cavity occlusion." The Journal of the Acoustic Society of America **53**(2): 400-408.
Goldberg, J. M. and P. B. Brown (1968). "Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study." Journal of Neurophysiology **31**(4): 639-656.
Grantham, D. W. (1984). "Interaural intensity discrimination: Insensitivity at 1000 Hz." The Journal of the Acoustical Society of America **75**(4): 1191-1194.
Griffin, S. J., L. R. Bernstein, et al. (2005). "Neural Sensitivity to Interaural Envelope Delays in the Inferior Colliculus of the Guinea Pig." Journal of Neurophysiology **93**(6): 3463-3478.
Grothe, B., M. Pecka, et al. (2010). "Mechanisms of Sound Localization in Mammals." Physiological Reviews **90**(3): 983-1012.
Grothe, B. and D. H. Sanes (1993). "Bilateral inhibition by glycinergic afferents in the medial superior olive." Journal of Neurophysiology **69**(4): 1192-1196.
Grothe, B. and D. H. Sanes (1994). "Synaptic inhibition influences the temporal coding properties of medial superior olivary neurons: an in vitro study." The Journal of neuroscience **14**(3 Pt 2): 1701-1709.
Guinan, J. J., B. E. Norris, et al. (1972). "Single Auditory Units in the Superior Olivary Complex: II: Locations of Unit Categories and Tonotopic Organization." International Journal of Neuroscience **4**(4): 147-166.
Hafter, E. R. and S. C. Carrier (1972). "Binaural Interaction in Low-Frequency Stimuli: The Inability to Trade Time and Intensity Completely." The Journal of the Acoustical Society of America **51**(6B): 1852-1862.
Hafter, E. R. and J. D. Maio (1975). "Difference thresholds for interaural delay." The Journal of the Acoustical Society of America **57**(1): 181-187.
Hancock, K. E. and B. Delgutte (2004). "A Physiologically Based Model of Interaural Time Difference Discrimination." The Journal of neuroscience **24**(32): 7110-7117.
Harper, N. S. and D. McAlpine (2004). "Optimal neural population coding of an auditory spatial cue." Nature **430**(7000): 682-686.
Harris, G. G. (1960). "Binaural Interactions of Impulsive Stimuli and Pure Tones." The Journal of the Acoustical Society of America **32**(6): 685-692.

Hawley, M. L., R. Y. Litovsky, et al. (1999). "Speech intelligibility and localization in a multi-source environment." The Journal of the Acoustical Society of America **105**(6): 3436-3448.
Hawley, M. L., R. Y. Litovsky, et al. (2004). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer." The Journal of the Acoustical Society of America **115**(2): 833-843.
Heil, P. (2001). "Representation of sound onsets in the auditory system." Audiology and Neurotology **6**(4): 167-172.
Helmholtz, H. v. (1885). "On the sensations of tone (1863)." English translation by AJ Ellis.
Henning, B. (1980). "Some observations on the lateralization of complex waveforms." The Journal of the Acoustical Society of America **68**(2): 446-454.
Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms." The Journal of the Acoustical Society of America **55**(1): 84-90.
Hershkowitz, R. M. and N. I. Durlach (1969). "Interaural Time and Amplitude jnds for a 500-Hz Tone." The Journal of the Acoustical Society of America **46**(6B): 1464-1467.
Hind, J. E., J. M. Goldberg, et al. (1963). "Some Discharge Characterstics of Single Neurons in the Inferior Colliculus of the Cat. II. Timing of the Discharges and Observations on Binaural Stimulation." Journal of Neurophysiology **26**(2): 321-341.
Hirsh, I. J. (1948). "The Influence of Interaural Phase on Interaural Summation and Inhibition." The Journal of the Acoustical Society of America **20**(4): 536-544.
Javel, E. (1980). "Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones." The Journal of the Acoustic Society of America **68**(1): 133-146.
Jeffress, L. A. (1948). "A place theory of sound localization." The Journal of comparative and physiological psychology **41**(1): 35.
Jeffress, L. A., H. C. Blodgett, et al. (1952). "The Masking of Tones by White Noise as a Function of the Interaural Phases of Both Components. I. 500 Cycles." The Journal of the Acoustical Society of America **24**(5): 523-527.
Jeffress, L. A., H. C. Blodgett, et al. (1956). "Masking of Tonal Signals." The Journal of the Acoustical Society of America **28**(3): 416-426.
John, M. N. and R. H. Ervin (1976). Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration, ASA.
Johnson, D. H. (1980). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones." The Journal of the Acoustic Society of America **68**(4): 1115-1122.
Joris, P. X. (2003). "Interaural Time Sensitivity Dominated by Cochlea-Induced Envelope Patterns." The Journal of neuroscience **23**(15): 6345-6350.

Joris, P. X. and T. C. Yin (1992). "Responses to amplitude-modulated tones in the auditory nerve of the cat." The Journal of the Acoustic Society of America **91**(1): 215-232.
Joris, P. X. and T. C. Yin (1995). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences." Journal of Neurophysiology **73**(3): 1043-1062.
Klein Hennig, M., M. Dietz, et al. (2011). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays." The Journal of the Acoustical Society of America **129**(6): 3856.
Klumpp, R. G. and H. R. Eady (1956). "Some Measurements of Interaural Time Difference Thresholds." The Journal of the Acoustical Society of America **28**(5): 859-860.
Kuhn, G. F. (1977). "Model for the interaural time differences in the azimuthal plane." The Journal of the Acoustical Society of America **62**(1): 157-167.
Kuwada, S., T. R. Stanford, et al. (1987). "Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit: effects of changing frequency." Journal of Neurophysiology **57**(5): 1338.
Laback, B., I. Zimmermann, et al. (2011). "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing." The Journal of the Acoustical Society of America **130**(3): 1515-1529.
Leakey, D. M., B. M. Sayers, et al. (1958). Binaural Fusion of Low and High Frequency Sounds, ASA.
Leek, M. R. (2001). "Adaptive procedures in psychophysical research." Perception and Psychophysics **63**(8): 1279.
Leslie, R. B. and T. Constantine (1994). Detection of interaural delay in high frequency sinusoidally amplitude modulated tones, two tone complexes, and bands of noise, ASA.
Levitt, H. (1971). "Transformed Up Down Methods in Psychoacoustics." The Journal of the Acoustical Society of America **49**(2B): 467.
Licklider, J. C. R. (1948). "The Influence of Interaural Phase Relations upon the Masking of Speech by White Noise." The Journal of the Acoustical Society of America **20**(2): 150-159.
Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals." The Journal of the Acoustical Society of America **80**(6): 1608-1622.
Loftus, W. C., D. C. Bishop, et al. (2004). "Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive." Journal of Comparative Neurology **472**(3): 330-344.
McAlpine, D., D. Jiang, et al. (2001). "A neural code for low-frequency sound localization in mammals." Nature Neuroscience **4**(4): 396-401.
Mills, A. W. (1958). On the Minimum Audible Angle.
Mills, A. W. (1960). "Lateralization of High-Frequency Tones." The Journal of the Acoustical Society of America **32**(1): 132-134.

Müller, J., F. Schon, et al. (2002). "Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system." Ear and Hearing **23**(3): 198-206.
Nuetzel, J. M. and E. R. Hafter (1976). "Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration." The Journal of the Acoustical Society of America **60**(6): 1339-1346.
Oxenham, A. J., J. G. W. Bernstein, et al. (2004). "Correct tonotopic representation is necessary for complex pitch perception." Proceedings of the National Academy of Sciences of the United States of America **101**(5): 1421-1425.
Palmer, A. R. (1982). "Encoding of rapid amplitude fluctuations by Cochlear-nerve fibres in the guinea-pig." Arch Otorhinolaryngol **236**(2): 197-202.
Palmer, A. R. and I. J. Russell (1986). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells." Hearing Research **24**(1): 1-15.
Pecka, M., A. Brand, et al. (2008). "Interaural Time Difference Processing in the Mammalian Medial Superior Olive: The Role of Glycinergic Inhibition." The Journal of neuroscience **28**(27): 6914-6925.
Pollack, I. and J. M. Pickett (1958). "Stereophonic Listening and Speech Intelligibility against Voice Babble." The Journal of the Acoustical Society of America **30**(2): 131-133.
Polyak, S. (1941). "The retina."
Pressnitzer, D. and R. Patterson (2001). "Distortion products and the perceived pitch of harmonic complex tones." Physiological and psychophysical bases of auditory function. Maastricht: Shaker Publishing: 97-104.
Rayleigh, J. W. S. (1877). The theory of sound, Macmillan and co.(London).
Rayleigh, L. (1907). "XII. On our perception of sound direction." Philosophical Magazine Series 6 **13**(74): 214-232.
Rose, J. E., J. F. Brugge, et al. (1967). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey." Journal of Neurophysiology **30**(4): 769-793.
Rose, J. E., D. D. Greenwood, et al. (1963). "Some Discharge Characterstics of Single Neurons in the Inferior Colliculus of the Cat. I. Tonotopical Organization, Relation of Spike-Counts to Tone Intensity, and Firing Patters of Single Elements." Journal of Neurophysiology **26**(2): 294-320.
Rose, J. E., N. B. Gross, et al. (1966). "Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source." Journal of Neurophysiology **29**(2): 288-314.
Roth, G. L., L. M. Aitkin, et al. (1978). "Some features of the spatial organization of the central nucleus of the inferior colliculus of the cat." The Journal of Comparative Neurology **182**(4): 661-680.
Rupert, A., G. Moushegian, et al. (1963). "Unit Responses to Sound from Auditory Nerve of the Cat." Journal of Neurophysiology **26**(3): 449-465.

Ryugo, D. K. (1992). "The auditory nerve: peripheral innervation, cell body morphology, and central projections." The mammalian auditory pathway: Neuroanatomy **1**: 23-65.
Sakitt, B. (1973). "Indices of Discriminability." Nature **241**(5385): 133-134.
Sayers, B. M. (1964). "Acoustic-Image Lateralization Judgments with Binaural Tones." The Journal of the Acoustical Society of America **36**(5): 923-926.
Sayers, B. M. and E. C. Cherry (1957). "Mechanism of Binaural Fusion in the Hearing of Speech." The Journal of the Acoustical Society of America **29**(9): 973-987.
Sayles, M. and I. M. Winter (2008). "Ambiguous Pitch and the Temporal Representation of Inharmonic Iterated Rippled Noise in the Ventral Cochlear Nucleus." The Journal of neuroscience **28**(46): 11925-11938.
Shackleton, T. M., R. Meddis, et al. (1992). "Across frequency integration in a model of lateralization." The Journal of the Acoustical Society of America **91**(4): 2276-2279.
Skottun, B. C., T. M. Shackleton, et al. (2001). "The ability of inferior colliculus neurons to signal differences in interaural delay." Proceedings of the National Academy of Sciences **98**(24): 14050-14054.
Smith, P. H. (1995). "Structural and functional differences distinguish principal from nonprincipal cells in the guinea pig MSO slice." Journal of Neurophysiology **73**(4): 1653-1667.
Smith, P. H., P. X. Joris, et al. (1998). "Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat." Journal of Neurophysiology **79**(6): 3127-3142.
Smith, P. H., P. X. Joris, et al. (1993). "Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: Evidence for delay lines to the medial superior olive." The Journal of Comparative Neurology **331**(2): 245-260.
Smith, R. L. (1977). "Short-term adaptation in single auditory nerve fibers: some poststimulatory effects." Journal of Neurophysiology **40**(5): 1098-1111.
Smith, R. L. (1979). "Adaptation, saturation, and physiological masking in single auditory-nerve fibers." The Journal of the Acoustical Society of America **65**(1): 166-178.
Spitzer, M. W. and M. N. Semple (1998). "Transformation of binaural response properties in the ascending auditory pathway: influence of time-varying interaural phase disparity." Journal of Neurophysiology **80**(6): 3062.
Sterbing, S. J., K. Hartung, et al. (2003). "Spatial Tuning to Virtual Sounds in the Inferior Colliculus of the Guinea Pig." Journal of Neurophysiology **90**(4): 2648-2659.
Stern, R. M., A. S. Zeiberg, et al. (1988). "Lateralization of complex binaural stimuli: A weighted-image model." The Journal of the Acoustical Society of America **84**(1): 156-165.

Steven van de, P. and K. Armin (1997). "A new approach to comparing binaural masking level differences at low and high frequencies." The Journal of the Acoustical Society of America **101**: 1671-1680.
Stevens, S. S. and E. B. Newman (1936). "The localization of actual sources of sound." The American Journal of Psychology **48**: 297-306.
Stiebler, I. and G. Ehret (1985). "Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution." The Journal of Comparative Neurology **238**(1): 65-76.
Strelcyk, O. and T. Dau (2009). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing." The Journal of the Acoustical Society of America **125**(5): 3328-3345.
Sullivan, W. and M. Konishi (1984). "Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl." The Journal of neuroscience **4**(7): 1787-1799.
Tan, X., X. Wang, et al. (2008). "First spike latency and spike count as functions of tone amplitude and frequency in the inferior colliculus of mice." Hearing Research **235**(1–2): 90-104.
Tasaki, I. (1954). "Nerve impulses in individual auditory nerve fibers of guinea pig." Journal of Neurophysiology **17**(2): 97-122.
Tollin, D. J. (2003). "The lateral superior olive: a functional role in sound source localization." Neuroscientist **9**(2): 127-143.
Tsuchitani, C. (1977). "Functional organization of lateral cell groups of cat superior olivary complex." Journal of Neurophysiology **40**(2): 296-318.
Tsuchitani, C. and J. C. Boudreau (1967). "Encoding of Stimulus Frequency and Intensity by Cat Superior Olive S-Segment Cells." The Journal of the Acoustical Society of America **42**(4): 794-805.
Webster, F. A. (1951). "The Influence of Interaural Phase on Masked Thresholds I. The Role of Interaural Time-Deviation." The Journal of the Acoustical Society of America **23**(4): 452-462.
Woodworth, R. (1938). "Experimental psychology."
Yin, T. C. and J. C. Chan (1990). "Interaural time sensitivity in medial superior olive of cat." Journal of Neurophysiology **64**(2): 465-488.
Yin, T. C., J. C. Chan, et al. (1986). "Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise." Journal of Neurophysiology **55**(2): 280-300.
Yin, T. C. and S. Kuwada (1983). "Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency." Journal of Neurophysiology **50**(4): 1020-1042.
Yin, T. C. T., S. Kuwada, et al. (1984). "Interaural time sensitivity of high-frequency neurons in the inferior colliculus." The Journal of the Acoustical Society of America **76**(5): 1401-1410.

Yost, W. A. (1974). "Discriminations of interaural phase differences." The Journal of the Acoustical Society of America **55**(6): 1299-1303.
Yost, W. A. and R. H. Dye (1988). "Discrimination of interaural differences of level as a function of frequency." The Journal of the Acoustical Society of America **83**(5): 1846-1851.
Yost, W. A., F. L. Wightman, et al. (1971). "Lateralization of Filtered Clicks." The Journal of the Acoustical Society of America **50**(6B): 1526-1531.
Young, L. L., Jr. (1976). "Time-Intensity Trading Functions for Selected Pure Tones." Journal of Speech and Hearing Research **19**(1): 55-67.
Zohar, O., T. M. Shackleton, et al. (2011). "First Spike Latency Code for Interaural Phase Difference Discrimination in the Guinea Pig Inferior Colliculus." The Journal of neuroscience **31**(25): 9192-9204.
Zwislocki, J. and R. S. Feldman (1956). "Just Noticeable Differences in Dichotic Phase." The Journal of the Acoustical Society of America **28**: 860-864.