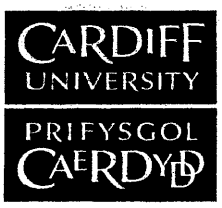

BINDING SERVICES
Tel +44 (0)29 2087 4949
Fax +44 (0)29 20371921
e-mail bindery@cardiff.ac.uk

Binaural resolution

by

Andrew Joseph Kolarik

September 2006
Cardiff University

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary of thesis

The aim of the experiments described within this thesis was to measure binaural temporal and spectral resolution. Previous investigations that have studied temporal resolution (e.g. Bernstein et al., 2001) assumed that interfering noise dilutes delayed noise within the temporal window. The first two experiments described in this thesis have validated the dilution concept for correlated interfering noise, but not for uncorrelated interfering noise, the presence of which has a more detrimental effect than interfering correlated noise. The study by Bernstein et al. (2001) suggested that the equivalent rectangular bandwidth (ERD) of the binaural temporal window is considerably smaller than estimates made in previous studies (e.g. Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998). The results from the experiments in this thesis disagree with those of Bernstein et al., and suggest that several factors led to their findings, including lack of control over the coherence of the stimulus due to the use of a detection task, the short duration of their stimuli, and the use of diotic interfering noise. The ERD of the binaural temporal window was found to range from 110-349 ms across listeners, a finding consistent with binaural sluggishness.

In the frequency domain, a study by Sondhi and Guttman (1966) that investigated the frequency selectivity of the binaural system found evidence suggesting that binaural auditory filters are substantially wider than monaural auditory filters. Conversely, Kohlrausch (1988) measured auditory filters that were comparable to monaural filters. The results from the experiment conducted in this thesis found that binaural auditory filters are substantially wider than monaural auditory filters. Best fits were found to be 2-parameter asymmetric Gaussian filters with an ERB that ranged from 99-198 Hz at a centre frequency (CF) of 250 Hz, 138-215 Hz at a CF of 500 Hz, and 229-285 Hz at a CF of 750 Hz.

Acknowledgements

This thesis was undertaken at the School of Psychology, Cardiff University, with the supervision of Dr. John F. Culling. Thank you for your help and guidance.

Thanks are due to my family and friends for their support.

The research was supported by a studentship from the EPSRC.

CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

1.1 TIME AND FREQUENCY PERCEPTION IN HEARING.

One of the more striking features of the auditory system is the ability it demonstrates in resolving stimuli in terms of time and frequency. Temporal and spectral resolution describes the limits to our capabilities when discriminating changes in stimuli over time and frequency respectively, and our ability to track moving sound sources, to perceive speech in noise, and to 'hear out' a melodic line in the presence of other melodies when listening to music are all governed to some degree by these limits.

Temporal and frequency resolution can be conceived in similar ways. Frequency selectivity has been envisaged as a set of bandpass filters with continuously overlapping passbands (Fletcher, 1940) now known as auditory filters, and temporal processing has been described using a temporal window model conceptually similar to that of an auditory filter. The temporal window applies progressively lower weight to events that precede or lag the centre of the window, thus attenuating information which occurred ahead or behind a given point in time. The auditory filter works on the same principles, but in the frequency domain, thus the auditory filter passes information at a given centre frequency without attenuation, and applies progressively less weight to spectral information occurring above or below the centre frequency of the filter. Both the temporal window and auditory filter can be characterised in terms of shape and extent about their central time/frequency.

The majority of research that has been conducted into frequency and temporal resolution in hearing has been performed within the monaural domain, which is concerned with listeners' ability to resolve stimuli by listening through a single ear. However, listening in daily life generally involves using both ears (binaural listening)

which assists in the localisation of sound sources, and helps detect and identify sounds in noise. This thesis is concerned with measuring the temporal resolution and frequency selectivity of the binaural system, and comparing it to existing research describing monaural resolution.

The auditory system integrates information in a variety of ways. Resolution refers to the limits of the ability of the system to detect changes in stimuli over time or frequency. This limit can be characterised by a minimum compulsory integration, described by an auditory filter in the spectral domain, and by a temporal window in the temporal domain. In the spectral domain, auditory filters can be described in terms of the bandwidth of the filter over which information is integrated, and it is the bandwidth of the filter that determines the resolution of the system. The narrower the filter bandwidth, the better the resolution of the system. In the temporal domain, temporal windows can be described in terms of time constants, the duration of which determine the resolution of the system, and the duration over which information is integrated. The shorter the duration of the temporal window, the better the resolution of the system. For stimulus durations less than the duration of the temporal window, thresholds drop by approximately 3 dB for each doubling of the duration of the signal, as doubling the duration of the signal doubles the energy of the signal present within the temporal window. For stimulus durations within and beyond the duration of the window, an additional reduction of 1.5 dB per doubling which is due to a decrease in 'decision' noise; an advantage arising from having the signal available for a longer period of time (e.g. Breebart, van de Par and Kohlrausch, 2001). On the other hand, Viemeister and Wakefield (1991) have contended that the decrease in thresholds beyond the time constant of the temporal window is due to the listener using a 'multiple looks' strategy, whereby listeners combine information from the outputs of

a range of temporal windows that constitute a number of 'looks' at the input, in order to improve their performance.

The binaural system plays a major role in assisting in the localisation of sound sources and identifying signals in noise. These abilities form the basis of the measurement techniques that allow us to measure spectral and temporal resolution, and are described below.

1.1.1 The binaural system and sound localization.

The direction of a sound relative to the listener can be defined by two angles: elevation and azimuth. Sound source elevation refers to the direction of a sound source relative to the listener's head in terms of angle in the vertical plane; sounds directly in front of the listener have 0° elevation and sounds coming from directly above the listener have 90° elevation. Sound source azimuth refers to the direction of a sound source relative to the listener's head in terms of angle in the horizontal plane. If the source is directly ahead of the listener, the azimuth is 0°, if directly to the left of the listener, 90°, if behind, 180°, and if directly to the right, 270° (equivalent to -90°). Listeners can use a range of cues to determine sound source azimuth, including interaural time difference (ITD) and interaural level difference (ILD).

ITD is the dominant cue for sound localization (Wightman and Kistler, 1992). When a sound source is lateralized to one side of the head, the sound that reaches one ear will be delayed relative to the other (see Fig. 1.1).

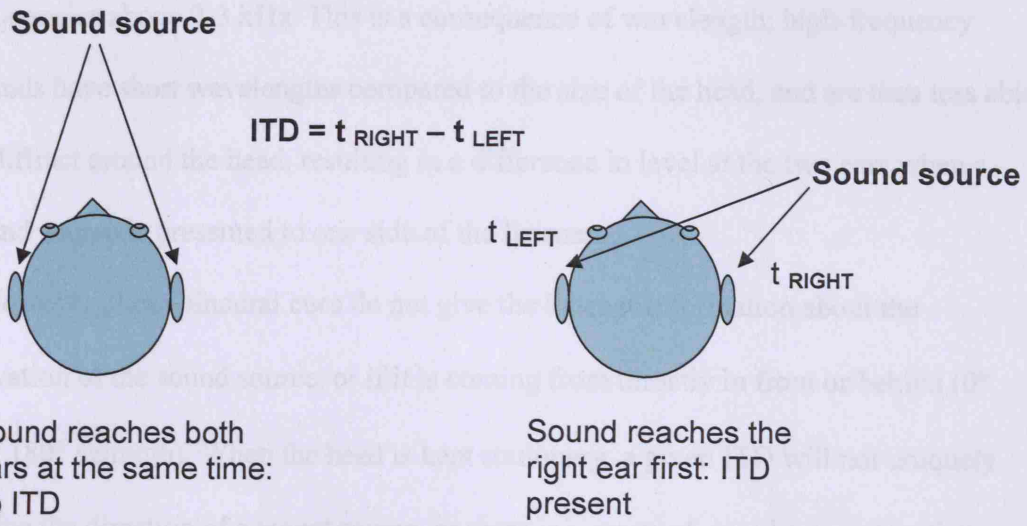

FIG. 1.1. Sound localisation and Interaural Time Delays (ITDs).

The size of the human head limits the ecological range of the ITDs that we experience in everyday life to approximately ±600 μs. When compared to a reference of zero ITD, the smallest ITD that the auditory system can discriminate is as low as 10 μs, for both pure tone stimuli (Domnitz, 1973; Domnitz and Colburn, 1977), and broadband noise (Klumpp and Eady, 1956; Mossop and Culling, 1998). When broadband noise stimuli are presented to listeners, the sound image can be described as fused and centred within the head, and as the ITD of the noise increases, the intracranial position of the sound moves towards the leading ear (for a review, see Durlach and Colburn, 1978), an experience known as lateralization. As ITDs extend beyond the physiological range, a noise image remains lateralized to the leading ear but becomes more and more diffuse, until the image fills the head and becomes impossible to lateralize (Blodgett, Wilbanks, and Jeffress, 1956).

At higher frequencies, listeners tend to rely on interaural level differences (ILDs) to localise sounds. ILD depends both on the difference in level due to the disparity in distance between the two ears, and the occluding effect of the head (Kuhn, 1977). The magnitude of the effect of the occlusion or 'head-shadow' is much higher at

frequencies above 2-3 kHz. This is a consequence of wavelength; high-frequency sounds have short wavelengths compared to the size of the head, and are thus less able to diffract around the head, resulting in a difference in level at the two ears when a sound source is presented to one side of the listener.

However, these binaural cues do not give the listener information about the elevation of the sound source, or if it is coming from directly in front or behind (0° and 180° azimuth). When the head is kept stationary, a given ITD will not uniquely define the direction of a sound source, as there is a range of sound source locations that can produce that ITD. This range of source locations is known as the 'cone of confusion'; the surface of the cone describes all sound source locations that would produce the ITD in question (Moore, 1989). Elevation and disambiguation between the frontal and rear hemifields are provided by spectral changes in the sound caused by reflections along the corrugations of the pinna (Batteau, 1967; Musicant and Butler, 1985; Butler, Humanski, and Musicant, 1990; Lopez-Poveda and Meddis, 1996), head movements that change the binaural cues, (Wallach, 1940; Perret and Noble, 1997), and movements of the sound source that the listener controls (Wightman and Kistler, 1999).

1.1.2 Interaural correlation, coherence, and the binaural masking level difference.

Interaural correlation can be defined as the point-by-point correlation coefficient between the waveform at the listener's left ear and the waveform presented to the right ear. It is distinct from interaural coherence, which is defined as the point-by-point correlation coefficient between the waveform at the listener's left ear and the waveform presented to the right ear *after* one waveform has been time shifted to

maximize the correlation; the maximum of the cross-correlation function. The term 'correlation' is sometimes used in this sense (e.g. Grantham, 1995), but not in this thesis. Traditionally, if the two waveforms are identical, then the noise is denoted as No; the suffix 'o' indicates zero difference in interaural phase. No noise has an interaural correlation of unity. If the two waveforms are statistically independent, then the noise is denoted Nu (suffix denoting 'uncorrelated') and the correlation has a mean of zero (Akeroyd and Summerfield, 1999). The perception of the noise varies according to its interaural correlation: No noise is perceived as narrow, compact and centralised in the head, whereas Nu noise is broad and diffuse across the head (Gabriel and Colburn, 1981; Grantham, 1995). An alternative percept was reported by Blauert and Lindemann (1986), who perceived a compact auditory event for No broadband noise which split into two events, one at each ear, for Nu noise. $N\tau$ noise denotes coherent noise that has an interaural delay ('τ' represents time delay) and is perceived to be fused and lateralised to one side of the head. The magnitude of lateralization depends on the magnitude of the delay. Nm noise is the notation used to describe noise presented to one ear only ('m' for monaural), and $N\pi$ noise denotes noise that is phase-inverted at one ear, resulting in a π radian interaural phase difference across the ears.

The threshold of a signal in noise can often be considerably lower when the listener is listening with two ears (binaurally) rather than just one (monaurally). Detectability of a pure tone in noise can be measured by obtaining the masked threshold: the level of the signal when it is just masked by the noise. Signals are more easily detected in noise in binaural conditions if there is an interaural phase difference between the interfering noise and the signal (Hirsh, 1948). For example, when the noise is in-phase across the ears and the phase of the signal is shifted by 180° or π radians ($NoS\pi$), or

vice versa (NπSo), the signal can be detected at a lower signal-to-noise ratio than if the noise and signal are both in-phase (NoSo), or both out-of-phase (NπSπ). The difference between masked thresholds in these cases is called the binaural masking level difference (BMLD), and is measured in dB (see Fig. 1.2).

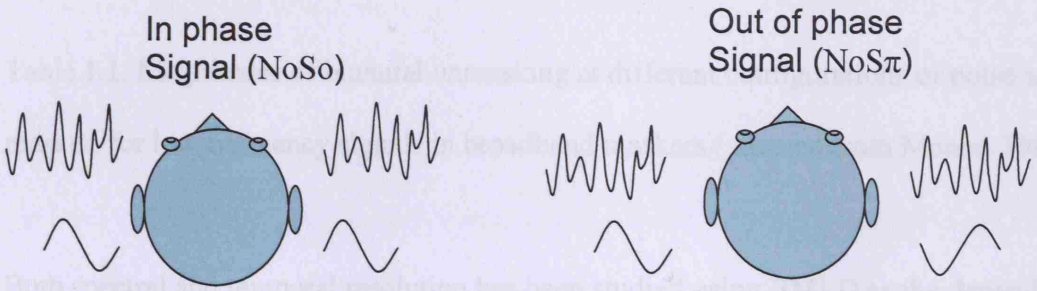

FIG. 1.2. Two stimulus configurations in a tone detection task. In NoSo, both the noise and the signal are identical at each ear. In NoSπ, the noise is identical at each ear but the signal is phase-inverted at the right ear. The BMLD between these conditions is approximately 15 dB at low signal frequencies and for broadband noise.

The largest BMLDs occur for lower frequencies, and the BMLD falls to 2-3 dB for signal frequencies above 1500 Hz in broadband noise (Durlach, 1978). The NoSo, NmSm and NπSπ conditions do not give the listener a binaural advantage and all give roughly the same reference threshold. Different configurations of noise and masker give the listener a BMLD of a magnitude comparable to those listed below (see Table 1.1), and are expressed relative to the reference threshold.

Interaural configuration	BMLD (dB)
NuSm	1.5
NuSπ	3
NuSo	4
NπSm	6
NoSm	9
NπSo	13
NoSπ	15

Table 1.1. Magnitudes of binaural unmasking at different configurations of noise and masker, for low frequency signals in broadband maskers (adapted from Moore, 1989).

Both spectral and temporal resolution has been studied using BMLD as the dependent variable (e.g. Sondhi and Guttman, 1966; Grantham and Wightman, 1979; Culling and Summerfield, 1998), as BMLD is an exclusively binaural phenomenon which makes it an ideal dependent measure for measuring binaural resolution.

1.2 SPECTRAL RESOLUTION.

1.2.1 Frequency selectivity and masking.

One of the fundamental characteristics of the auditory system is the ability to analyse frequency, to resolve the components in a complex sound. One of the earliest attempts to measure this ability was made by Fletcher (1940), who measured the monaural threshold of a sinusoid in a band of noise. The threshold of the signal initially increased as the bandwidth of the masker was increased, but beyond a certain point thresholds flattened off and further increases in the bandwidth of the noise did not change the signal threshold. Fletcher envisaged the auditory system as a bank of bandpass filters with overlapping passbands, and suggested that the listener performed the task by attending to the output of the auditory filter centred on the signal

frequency. As the bandwidth of the noise increases, more noise gets into the filter and thresholds increase. When the bandwidth of the noise is equal to the bandwidth of the auditory filter, further increases in the bandwidth of the noise fail to increase detection thresholds as the noise falls outside the passband of the filter. Fletcher labelled the bandwidth beyond which thresholds cease to increase the critical bandwidth (CB), and assumed it was close to the bandwidth of the critical band at the same centre frequency of the noise. Fletcher made the working assumption that the shape of the band was rectangular, with flat top and vertical edges, so that all components within the band are passed equally and all components outside the band are completely attenuated. The value of the CB was estimated by fitting the data with a horizontal line at larger bandwidths where the thresholds are broadly consistent, and fitting a second line along the slope at smaller bandwidths where thresholds are seen to increase as bandwidth increases. The intercept of the two lines defines the critical bandwidth.

However, measurement of the auditory critical band (now described as an auditory filter) can be influenced by the choice of masker. When a bandpass noise masker is employed, as it was in the study of Fletcher (1940), the highest signal-to-noise ratio may not be obtained by centring the auditory filter at the signal frequency. A better signal-to-noise ratio may be obtained by listening to a filter that is higher or lower in frequency than the signal. This process is known as off-frequency listening (Patterson, 1976). To prevent the listener using an off-frequency listening strategy, a masker must be employed that limits the advantage in terms of signal-to-noise ratio that a shift in filter centre frequency might produce. There are further problems associated with the use of a narrowband masker, as the listener may be able to combine the information from several auditory filters to lower the internal noise (van de Par and Kohlrausch,

1999), instead of listening through a single filter. A masker that is able to address these issues consists of broadband noise with a spectral notch around the signal frequency (see Fig. 1.3).

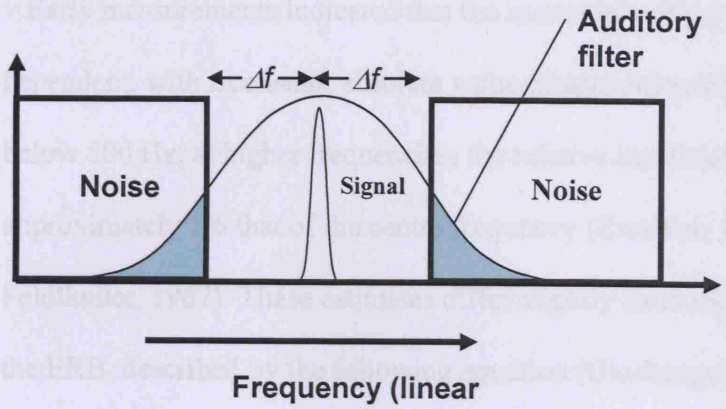

FIG. 1.3. Schematic illustration of the notched-noise method used by Patterson (1976) to measure auditory filter bandwidth.

The measure of bandwidth associated with the notched-noise technique is called the equivalent rectangular bandwidth (ERB). The ERB of a given filter is equal to the bandwidth of a perfect rectangular filter, which has a transmission in its passband equal to the maximum transmission of the specified filter, and transmits the same power of white noise as the specified filter (Moore, 1989), thus allowing the ERB and CB to be compared. The ERB has the same centre frequency and spectral density of the given filter and is calculated by integrating the spectral information within the filter.

In order to measure the filter, the threshold of the sinusoidal signal is measured as a function of the width of the spectral notch in the noise masker, rather than as a function of noise bandwidth when the masker is a bandpass noise centred over the signal. As the spectral notch width increases, less noise gets into the filter (indicated by the shaded areas in Fig. 1.3), and thresholds decrease. Different filter shapes have

been assumed in order to obtain best fits to the psychophysical data. Gaussian, rounded-exponential (or roex for short), and exponential shapes are among those most often used.

Early measurements indicated that the monaural critical bandwidth was frequency dependent, with a constant absolute value of approximately 100 Hz at frequencies below 500 Hz; at higher frequencies, the relative bandwidth was found to be approximately 1/6 that of the centre frequency (Zwicker, 1961; Zwicker and Feldtkeller, 1967). These estimates differ slightly from more recent measurements of the ERB, described by the following equation (Glasberg and Moore, 1990):

$$\mathrm{ERB} = 24.7(4.37F+1)$$

where F is the centre frequency in kHz. According to this method, the ERB is approximately 78 Hz at a centre frequency of 500 Hz, and broadens at higher frequencies (e.g. 564 Hz at 5 kHz). The value of the ERB, especially at low frequencies, tends to be smaller than critical bandwidth estimates (Moore and Glasberg, 1983).

Studies that have measured the bandwidth of the binaural auditory filter have demonstrated conflicting results; bandwidths comparable to monaural measurements have been found (e.g. Kohlrausch, 1988), which contrast with studies indicating much larger binaural bandwidths (e.g. Sondhi and Guttman, 1966). The measurement of binaural auditory filters is discussed in more detail in chapter 5.

1.3 TEMPORAL RESOLUTION – A 'SLUGGISH' SYSTEM?

Moore, Glasberg, Plack, and Biswas (1988) speculated that the temporal resolution of the auditory system could be modelled using a temporal window (an intensity weighting function) that was the temporal analogue of the auditory filter measured in

the frequency domain by Fletcher (1940). It was proposed that the resolution of the monaural auditory system could be measured by the effective duration of the window. A sinusoidal signal presented monaurally at 500 or 2000 Hz was presented between two 200 ms noise bursts. The duration of the gap was systematically varied, and the signal was presented symmetrically and asymmetrically within the gap in order to measure any asymmetry in the window. It was assumed that the window was centred at the position that provided the best signal-to-noise ratio. The data were fitted with temporal windows with a variety of shapes, and the best-fit to the data was found when each side of the window was modelled to be the sum of two roex functions. The equivalent rectangular duration (ERD, the temporal analogue of equivalent rectangular bandwidth in the frequency domain) of the window was found to be approximately 8 ms.

Plack and Moore (1990) investigated the monaural temporal window in more detail, measuring temporal windows at a variety of frequencies and levels. The ERD of the window was found to decrease slightly from 13 to 7 ms as the centre frequency increased from 300 to 8100 Hz, and decrease slightly as the level increased e.g. from 10 ms at a centre frequency of 2700 Hz with a 20-dB masker to 7 ms at a 40-dB masker level. All window shapes were found to be asymmetric, with a longer lower lobe than upper lobe.

Early studies investigating temporal acuity found that the ability to temporally resolve binaural disparities was far lower than for monaural amplitude modulation (Grantham and Wightman, 1978; Grantham, 1982). There are a range of studies that have indicated that the binaural system is 'sluggish', meaning that in comparison to the monaural system, it cannot resolve fast changes in interaural stimulation (e.g. Perrot and Musicant, 1977; Grantham and Wightman, 1979; Perrot and Pacheco,

1989; Kollmeier and Gilkey, 1990; Yama, 1992; Culling and Summerfield, 1998; Holube, Kinkel, and Kollmeier, 1998; Akeroyd and Summerfield, 1999). The slower response of the binaural system compared to that of the monaural system is assumed to be due to the action of a binaural temporal window that is longer than the monaural temporal window. Binaural windows are typically found to have integration times of approximately 100 ms (e.g. Grantham and Wightman, 1979; Culling and Summerfield, 1998), whilst the monaural window is much shorter, around 7-13 ms (Moore et al., 1988; Plack and Moore, 1990). Conversely, some studies investigating the binaural system have found no evidence of sluggishness (Wagner, 1991; Witton, Green, Rees and Henning, 2000; Bernstein, Trahiotis, Akeroyd and Hartung, 2001; Witton, Simpson, Henning, Rees, and Green, 2003).

A variety of studies that have attempted to measure the temporal resolution of the binaural system have found indications that the system is sluggish. Some experimenters found evidence of sluggishness when manipulating the time interval between two static sound sources activated sequentially (Perrott and Pacheco, 1989), and by discriminating between static and dynamic stimuli using localization tasks (e.g. Perrott and Musicant, 1977). Others have manipulated the rate of modulation of the ITD of a target stimulus (Grantham and Wightman, 1978), periodically varied the interaural correlation of the masker (Grantham and Wightman, 1979), or varied the correlation of the masker in a stepwise fashion (Kollmeier and Gilkey, 1990), or in a rectangular manner (Culling and Summerfield, 1998). When periodic variation is employed in a study, the interaural phase of the masker is modulated throughout stimulus presentation. In contrast, in studies that employ a stepwise design, the correlation of the masker is changed only at a single point in time during the stimulus, for example from in-phase to out-of-phase. Studies that have a rectangular change in

correlation in their design have two changes in the correlation in the masker during stimulus presentation, e.g. from out-of-phase to in-phase to out-of-phase. Different studies that have employed different methods have produced a wide range of time constants (Holube et al., 1998), and are described below.

1.3.1 Sound localization: measurements of the Minimum Angle and the Minimum Audibility Movement Angle.

The minimum audible angle (MAA) is the minimum angle of arc that a pair of sequentially presented stationary sound sources must be separated by in order for a listener to discriminate between them and a single stationary source emitting the same sounds (Mills, 1958). Perrott and Pacheco (1989) presented listeners with bursts of broadband pink noise from two sound sources activated sequentially. The task was to indicate whether the second sound came from the left or right of the first sound, and the interstimulus-onset-interval (ISOI) between the two sounds was varied between 1 and 900 ms. They found that MAAs were dependent on the ISOI and systematically decreased as the ISOI increased (from 4.65° at 1 ms to 0.93° at 150 ms). No improvement in MAA was observed beyond 150 ms, indicative of a minimum integration time (MIT) of around 150 ms which blurred together the internal representation of the two stimuli at shorter ISOIs.

Poor temporal resolution in the binaural system has also been noted in studies of the minimum auditory movement angle (MAMA); the angle of arc that a moving sound source must pass through in order to be discriminated from a stationary sound source. The MAMA is larger than the MAA (Harris and Sergeant, 1971; Harris, 1972; Perrott and Musicant, 1977; Chandler and Grantham, 1992), and increases with the sound source's angular velocity. In a study by Perrott and Musicant (1977), at the slowest

angular velocity that was tested (90°/s), the MAMA was 8.3°, and increased to 21.2° at the fastest velocity tested (360°/s). Minimum integration times were approximately 92 ms at 90°/s, and 59 ms at 360°/s. A study of the MAMA by Grantham (1986) found that the MIT ranged between 150 and 300 ms, and a study by Chandler and Grantham (1992) found an average MIT of 336 ms.

The MITs apparent in studies of the MAA and the MAMA are of the same order of magnitude. Both indicate that when different source locations for static or moving stimuli are presented within 100 ms of one another, there is a decrement in listeners' ability to detect the change in position.

1.3.2 Studies of binaural temporal resolution involving periodic changes in stimulus parameters.

A study by Grantham and Wightman (1978) provided one of the earliest demonstrations of binaural sluggishness. Three experiments were conducted that employed wide-band stimuli with sinusoidally modulated ITDs. The static reference was either a dichotic noise stimulus (experiment 1), or a dichotic noise stimulus with an 'image width' that matched the distance of excursion of the moving stimulus (experiment 2). When the rate of modulation was higher than 5 Hz, listeners could no longer discern between a moving and a static stimulus. Above 20 Hz, thresholds decreased due to the use of cues other than change in lateral position. In the third experiment, the sinusoidally moving noise stimulus used in the previous two experiments was used to mask a binaural click. The rate of modulation of the masker was manipulated, and the position of the noise image in the head at the moment the click was delivered was controlled by manipulating the instantaneous ITD of the masker. At low modulations, thresholds for detecting the click when it was presented

on the opposite side of the head to the noise were lower than when the click was presented on the same side of the head. As the movement rate of the masker was increased, the binaural system was less able to follow the movement of the masker and thresholds became independent of masker position at a modulation rate of approximately 5 Hz.

In a second study (Grantham and Wightman, 1979), stimuli consisted of a short interaurally phase inverted probe tone (Sπ) and a narrow-band Gaussian noise masker centred at the signal frequency. The masker's interaural correlation was varied sinusoidally between 1 and -1 at rates varying from 0-4 Hz. The probe tone was presented at various points in the modulation cycle of the masking noise. BMLD was measured as a function of the rate of modulation of the masker, by comparing thresholds when the instantaneous stimulus was NoSπ to thresholds when the instantaneous stimulus was NπSπ. At masker modulations close to 0 Hz, signal threshold decreased monotonically as the masker correlation altered from -1 to unity. As the modulation rate increased, the function relating signal threshold to masker correlation was flattened at all signal frequencies. For a modulation frequency of 4 Hz, there was no effect of masker correlation on signal detectability, and the BMLD was abolished. The flattening of the functions is attributable to the binaural system smoothing the dynamic fluctuations of the stimulus; beyond the cut-off frequency of 4 Hz, the stimulus fluctuations become too fast to be resolved. As this cut-off frequency was considerably lower than comparable monaural experiments (e.g. Viemeister, 1977), Grantham and Wightman labelled the slow response of the system as 'binaural sluggishness'. A single-sided exponential temporal window was fitted to the data, and its time constant defined as a 'binaural minimum integration time.' Across all conditions and subjects, time constants ranged from 44-243 ms.

It has since been found that the application of continuous periodic changes to the masker consistently produces larger time constants than when the change is stepwise or rectangular (Holube et al., 1998). The reasons for this are not entirely understood, although a number of suggestions have been made to explain these observations. Constantly varying the interaural correlation in a periodic fashion causes the percept of the noise image to vary continuously, compared to stepwise or rectangular changes where the 'binaural image' is more well-defined throughout stimulus presentation. Thus, it has been argued that the linear sliding temporal window model should not be applied when the masker is varied in a periodic manner, as the binaural image changes are not taken into account in the modelling procedure (Holube et al., 1998). In addition, when applying continuous changes in correlation to the masker, only a low modulation rate is required to abolish any binaural advantage (Culling and Summerfield, 1998), as only a slight departure from a correlation of 1 results in a large reduction in BMLD (Robinson and Jeffress, 1963). The sharp increase in detection thresholds for small increases in decorrelation of the masker results in large estimates of binaural minimum integration times.

Kollmeier and Gilkey (1990) argued that there may be a variety of binaural transmission channels that correspond to different binaural abilities, such as localisation (e.g. Blauert, 1968), lateralization (e.g. Blauert, 1972; Grantham and Wightman, 1978), and detection (e.g. Grantham and Wightman, 1979). These channels may be described by individual time constants that are none the less within the same order of magnitude. Within this framework, the large values of time constants obtained from BMLD studies employing periodic variations in masker correlation have been interpreted using EC theory (Durlach, 1972). In the experiments of Grantham and Wightman (1979) and Holube et al. (1998), because the correlation

is changing continuously, the outputs of different binaural processing channels with various equalization mechanisms have to be compared, and the output with the highest signal-to-noise ratio selected. The comparison process may produce additional sluggishness, leading to higher time constants than those involving stepwise or rectangular changes in correlation. This form of sluggishness is distinct from conventional or 'strategy sluggishness' (Holube et al., 1998), and was termed 'binaural analyzer sluggishness' by Kollmeier and Gilkey (1990).

1.3.3 Studies of binaural temporal resolution involving stepwise or rectangular changes in the interaural correlation of the masker.

Kollmeier and Gilkey (1990) measured the temporal response of the binaural system using a binaural analogue of monaural forward and backward masking. The threshold level of a 20 ms, 500 Hz Sπ tone was obtained in the presence of a 750-ms noise masker, which switched from Nπ to No after 375 ms, or switched from an No configuration to an Nπ configuration. The Sπ tone was presented at points before and after the phase transition of the masker, so that the stimulus configuration was either NoSπ or NπSπ depending on the temporal position of the tone. As the duration between signal tone offset and noise phase transition increased, thresholds decreased gradually. Binaural thresholds were compared directly to a monaural situation, where the phase of the masker was held at Nπ, but the level was lowered by 15 dB, corresponding to the maximum BMLD resulting from the phase transition in the binaural condition. The experimenters modelled temporal window functions with five different shapes to the data: rectangular, triangular, Gaussian, exponential, and rounded exponential. The best fits were obtained by a double-sided exponential window, with an ERD ranging from 33.2 to 83.2 ms. These ERDs are of the same

order of magnitude as Grantham and Wightman's (1979) 'binaural minimum integration time' of 44-243 ms.

However, the configuration of the stimuli in Kollmeier and Gilkey's experiment allows the listener to exploit a phenomenon called 'off-time' listening (Moore, Glasberg, Plack, and Biswas, 1988; Plack and Moore, 1990) that is analogous to 'off-frequency listening' in the frequency domain (Patterson, 1976). The signal-to-noise ratio may be improved by moving the temporal window forward or backward in time in relation to the presentation of the signal. This is true for all window shapes except for exponential windows, where off-time listening makes no difference to thresholds as the integral of an exponential is an exponential with the same time constant. This is because the amount of signal energy entering the window is reduced by the same proportion as the amount of noise energy entering the window as the centroid of the exponential window is moved further from the noise, thus listening to an exponential temporal window that is not centred directly over the signal does not result in any advantage in terms of signal-to-noise-ratio (Culling and Summerfield, 1998). In their study, Kollmeier and Gilkey allowed for off-time listening by assuming that listeners centred their temporal window at the offset of the tone in the forward masking condition and at the onset of the tone in the backward masking condition (an offset in each case of 10 ms), thus giving them an advantage in terms of signal-to-noise ratio. However, off-time listening was not explicitly modelled, as it is possible that Kollmeier and Gilkey's listeners were off-time listening by more than 10 ms. Assuming that the listener can choose from a range of temporal windows centred at different points in time, the sharpness of the windows that were not modelled using an exponential shape may have been over-estimated if their listeners were using an off-time listening strategy. The apparent sharpening of the window may have resulted in

better fits being obtained for the sharper exponential rather than rounded-exponential window shapes that were modelled by Kollmeier and Gilkey (Culling and Summerfield, 1998). This advantage may have been reduced if off-time listening was prevented or explicitly modelled.

In order to reduce their listener's use of an off-time listening strategy, Culling and Summerfield (1998) conducted a study using a binaural analogue of the notched-noise techniques used in measurement of the auditory filter (Patterson, 1976), in which off-time listening was modelled. Stimuli consisted of a burst of diotic noise temporally flanked on each side by 400 ms of uncorrelated noise. A 10-ms Sπ tone was presented during the correlated portion of the noise in one interval of a 2I-2AFC task. The listener is able to detect the presence of the signal by centring a temporal window over the Sπ tone. The threshold for detecting an Sπ tone in No noise is approximately 12 dB lower than in an Nu noise (e.g. Moore, 1989). Any increase in threshold as the duration of the diotic noise is reduced can be taken as an indication of uncorrelated noise entering the optimally placed window, reducing the amount of binaural unmasking. Three independent variables were manipulated: the duration of the correlated portion of the stimulus (0-960 ms), the signal frequency (125, 250, 500 and 1000 Hz), and signal level (20, 30, 40, and 50 dB (SPL)/Hz). Three base functions were modelled to the data: Gaussian, rounded-exponential, and exponential, and three fit-types (simple, floor, and skirt). The best fit (a simple Gaussian) was found to be largely independent of frequency and level, as was the ERD of the windows which ranged from 55 to 188 ms. The goodness-of-fit improved monotonically (but not significantly), with the roundness of the peak of the fitting function, i.e., worst fits were obtained for exponential windows, moderate fits were found for rounded-exponential windows, and best fits were obtained for Gaussian windows.

1.3.4 Studies of binaural temporal resolution involving temporal separation between signal and masker, and binaural gap detection.

Yama (1992) compared masking level differences for different temporal separations of masker and signal. Thresholds were obtained for diotic (NoSo) and dichotic (NoSπ) conditions using a forward-masking paradigm. The level of the noise masker was varied between 32 and 80 dB, and the temporal separation between signal and masker varied from simultaneous to 100 ms. An additional condition at a level of 24 dB was added for the simultaneous masking condition. The slopes of the masking functions in both conditions were observed to decrease as temporal separation increased. After correction for NoSo masking effectiveness, the MLD decreased by only 1.4 dB as temporal separation increased from 5 to 100 ms, indicating a long time constant for the binaural system comparable to studies that found evidence of binaural sluggishness (Grantham and Wightman, 1978; Grantham and Wightman, 1979; Kollmeier and Gilkey, 1990).

Akeroyd and Summerfield (1999) used a binaural analogue of gap detection in order to measure the ERD of the binaural temporal window without incorporating a BMLD design. A binaural 'gap' in interaural correlation was created by presenting listeners with a burst of uncorrelated noise between two bursts of diotic noise. The temporal window smoothes the gap into a dip at the output of the window, and assuming that binaural gap detectability is limited by the detectability of the dip, the ERD of the temporal window can be calculated providing that the just-noticeable-difference (jnd) in correlation from unity is known. Measurements of binaural-gap thresholds and jnds in interaural correlation from unity were obtained and analyzed using a computational model of binaural processing. A temporal window with a mean ERD of 140 ms was obtained, a result consistent with binaural sluggishness.

1.3.5 Binaural resolution and interaural level differences.

Unlike ITD resolution, studies investigating the binaural system's ability to resolve fluctuating ILDs found that the system is not sluggish in comparison to the monaural system (Blauert, 1972; Burns and Colburn, 1977; Grantham, 1984; Bernstein and Trahiotis, 1992), providing support to the notion that ITDs and ILDs are processed independently (Grantham, 1984; Moore, 1989). As there is widespread agreement among previous studies that indicate that ILD processing is not sluggish, this thesis focuses on the binaural temporal resolution of ITDs only.

1.3.6 Studies that demonstrate no binaural sluggishness.

In comparison with the literature described in the preceding sections, a range of studies have produced results that demonstrate that the binaural system is able to resolve temporal fluctuations to a degree that is comparative with the monaural system (e.g. Pollack, 1978; Wagner, 1991; Witton et al., 2000; Bernstein et al., 2001; Witton et al., 2003). Pollack (1978) presented binaural random polarity-modulated pulse trains that switched between correlations of 1 and -1, and demonstrated that listeners could detect switching periods of 2-4 ms, a period comparable to monaural time constants. However, Grantham (1982) argued that Pollack's data could be explained in terms of a binaural temporal window with a time constant of 100 ms. In a previous study, Grantham and Wightman (1979) modelled the binaural system as a simple *RC* lowpass filter, and from their data obtained a time constant of 100 ms. The cutoff frequency for the system is approximately 2 Hz, but with an attenuation rate of 6 dB/octave, Grantham (1982) stated that the output of the system would be down by 48 dB at 500 Hz. A 500-Hz binaural switching rate could be detected within the tail of the skirt of the filter if the binaural input was large enough, and as Pollack's stimuli

switched in correlation between 1 and -1, the magnitude of the binaural input was at its maximum.

Witton et al. (2000) investigated temporal resolution by introducing interaural phase modulation (IPM) in their stimuli. In the first experiment, a 500-Hz tone was presented in the left ear only (monaural condition). Two intervals were presented, where the 500-Hz tone was modulated in one interval but not in the other interval. The second experiment was identical to the first, except that a static, unmodulated tone was also presented to the right ear in both intervals, thus constituting a dichotic condition. At a modulation rate of 1 Hz, listeners were found to have thresholds an order of magnitude lower in the dichotic condition than the monaural condition. This 'dichotic advantage' remained present for modulation depths up to approximately 40 Hz, thus suggesting that detection of internal phase modulation is not sluggish. These results suggest that the binaural system is not adversely affected by the speed of the phase changes up to 40 Hz and listeners are able to follow rapid interaural modulation of perceived location. However, it is possible that the target interval could be distinguished from the distracter interval in the dichotic condition on the basis of coherence changes. Introducing IPM reduces the interaural coherence of the stimulus, leading to a change in perceived image width (Gabriel and Colburn, 1981) or loudness (Culling and Edmonds, 2006). Gabriel and Colburn (1981) reported that correlated Gaussian noise is perceived as fused and/or compact, whereas uncorrelated noise is perceived as diffuse and filling the whole head. Culling and Edmonds (2006) gave listeners a loudness-matching task and presented 500 ms bursts of noise with various interaural coherences. Cues associated with image width changes were controlled for as listeners were required to match stimuli in loudness using an intensity offset, which has a negligible impact on image width. Their results indicated that uncorrelated noise

was perceived to be louder than diotic noise. Thus, thresholds in the dichotic condition of Witton et al. (2000) could have been improved through width or loudness cues that were not temporal in nature.

A second study (Witton et al., 2003) was designed to examine dichotic linear ramp modulations, which give listeners the percept of smooth unidirectional horizontal movement. Initially, a 500 Hz tone was held at a steady state ITD for 20 ms. Linear phase modulation (PM) was then applied to the tone for 960 ms, and during the final part of the stimuli the signal was held at a steady-state ITD for 20 ms. A detection condition was used to obtain thresholds for linear dichotic PM ramps, and for diotic PM ramps where no ITD was present. A discrimination condition was also used to establish sensitivity to direction rather than presence of modulation alone. In line with the earlier study investigating sinusoidally varying IPM (Witton et al., 2000), sensitivity to dichotic PM was found to be greater than diotic PM. However, once again thresholds may have been confounded by listeners using width or loudness cues arising from interaural coherence differences rather than laterality judgements alone. In the detection task, only one interval was decorrelated by IPM. Listeners were more sensitive in this condition compared to the discrimination condition in which both intervals were decorrelated by IPM, where laterality cues only were available.

Wagner's (1991) study of the binaural temporal resolution in the barn owl produced results that contrasted with the sluggish time constants found in human listeners. Stimuli consisted of an interaurally delayed probe noise temporally contiguous with uncorrelated masking noise, and kept the ITD fixed while varying the duration of the probe and masking noises. The owls' task was to discriminate a leftward or rightward target ITD and indicate the target with a head turn in the direction of the ITD. When a temporal window was fitted to the data, it was best described by a roex or exponential

shape (which produced similar best fits to the data) with two time constants; a peak of short duration (3-5 ms), and a longer skirt (30-50 ms), which are far shorter than the 100-ms time constants typically described in studies of human binaural temporal resolution (e.g. Culling and Summerfield, 1998).

Wagner's experiments formed the basis of a study by Bernstein et al. (2001), who used a similar experimental design to examine sensitivity to dynamic changes of ITD in humans, with similar results. They used an ITD detection task in which participants were asked to make lateralization judgements based on stimuli with a total duration of 100, 40, or 20 ms. In their first experiment, they employed a stimulus design in which an ITD was applied to a probe section of noise with a duration of 2, 4, 8, 16, 32 or 64 ms, temporally centred within No. Best fitting temporal windows were found to be described by exponential-skirt functions. These windows were composed of two time constants; one of a short duration (between 0.02 and 0.12 ms) that described the central peak of the window, and a second longer time constant (between 7.48 and 64.21 ms) that described the window skirts. The weighting parameter *w* for the skirt ranged from -13.98 to -16.99 dB. These temporal windows were far shorter than previous estimates such as Akeroyd and Summerfield (1999), who obtained a window with a mean ERD of 140 ms, and Culling and Summerfield (1998), who obtained a window 55 to 188 ms wide. Investigation of the monaural system provided evidence that the monaural temporal window has an ERD of approximately 13 ms (Plack and Moore, 1990), thus the data produced in the study of Bernstein et al. suggests that the binaural system has in a sense an even finer temporal resolution than the monaural system. The bulk of the thesis dedicated to temporal resolution attempts to explain why Bernstein et al's study produced results inconsistent with binaural sluggishness. Fundamental to Bernstein et al's analysis was the assumption that ITD is diluted by

interfering noise in a linear (1:1) fashion. The first two experiments investigated whether this assumption was valid.

1.4 RESEARCH QUESTION AND OUTLINE OF THESIS.

1.4.1 Outline of chapters.

The aim of the thesis was to measure the spectral and temporal resolution of the binaural system:

- Chapter 2 describes the general methods used throughout the thesis.

- Chapter 3 addresses temporal resolution, and the importance of the effect of different noises mixing within the temporal window. Experiment 1 examines how ITD discrimination thresholds for broadband noise change in the presence of interaurally correlated interfering noise. A regression analysis showed that thresholds doubled for every halving of the proportion of delayed noise power in the stimulus. Experiment 2 measures ITD discrimination thresholds for a range of interaural correlations obtained by mixing correlated and uncorrelated noise before applying a delay. Thresholds more than doubled for each halving of the correlation at all three stimulus durations tested.

- Chapter 4 examines whether stimulus coherence can act as a confounding cue in detection tasks designed to measure the binaural temporal window. Participants were asked to perform detection and discrimination tasks involving brief changes in ITD imposed on probe sections of noise temporally flanked with diotic noise. Experiments are described that use two different methods of obtaining thresholds; experiment 3 uses an adaptive track, and experiment 4 measures psychometric functions. When temporal windows were modelled to the detection data, windows with a narrow peak and wide skirt,

similar to those found by Wagner (1991) and by Bernstein et al. (2001), were obtained. However, windows fitted to the discrimination data were found to have large and indeterminate ERDs, suggesting that the narrow windows observed from the detection data were the result of listeners using a coherence cue that was absent in the discrimination task. The next two experiments attempted to obtain a measurable temporal window by repeating experiment 4, but using stimuli of greater duration that were long enough to encompass the whole window. Experiment 5 tested listeners with diotic masking noise surrounding the probe, and experiment 6 tested listeners with uncorrelated noise surrounding the probe. When temporal windows were modelled using the appropriate slope functions observed in experiments 1 and 2 to the data obtained in experiment 5, they were found to have narrow peaks and large and indeterminate skirts, a finding probably attributable to the distracting effect of auditory events occurring when the listener is presented with contiguous delayed noise and diotic noise. When temporal windows were modelled to the data of experiment 6 which ameliorated the effects of the auditory events, they were found to have an ERD of approximately 100 ms, supporting the notion of a 'sluggish' system.

- Finally, Chapter 5 examines spectral resolution using a tone detection task in a binaural analogue of the notched-noise technique in order to resolve the conflict between studies that have indicated that the binaural auditory filter is comparable to the monaural filter (e.g. Kohlrausch, 1988), which contrast with studies indicating much larger binaural bandwidths (e.g. Sondhi and Guttman, 1966). Consistent with the results reported by Sondhi and Guttman (1966), the

equivalent rectangular bandwidth of the binaural filter was found to increase with centre frequency, and was consistently larger than monaural bandwidths.

CHAPTER 2: GENERAL METHODS

2.1 INTRODUCTION

Methodology specific to a given experiment will be described in detail in the relevant chapter. Two procedures were used to gather thresholds throughout the experiments; the adaptive track, and the measurement of psychometric functions. These methods are described in this chapter, as is the window-fitting procedure that describes how auditory filters and temporal windows were fitted to the experimental data.

2.1.1 Participants.

Listeners that participated in the various studies consisted of the author of the thesis (AK), a member of the Hearing laboratory of Cardiff University (JC), and a variety of Cardiff University psychology undergraduates who were paid for their participation. Listeners were given at least 5 hours of preparatory training, and attended sessions that lasted no longer than 1 hour.

2.1.2 Stimuli and fitting algorithms.

The following methods of stimulus generation and presentation are common to the majority of the experiments described in the thesis. Any deviations from the methods described in this section will be described in detail in the methods section of the relevant experiment.

The stimuli were generated digitally with a sampling rate of 44.1 kHz and 16-bit sample depth using Matlab. Stimuli were bandlimited between 100 and 3000 Hz prior to presentation, with 10-ms onset and offset ramps. Stimuli were presented over Sennheiser HD590 headphones at an overall sound level of 75 dB through an MTR

HPA-2 Headphone Amplifier using a 24-bit Edirol UA-20 sound card in a single-walled IAC sound-attenuating booth within a sound-deadened room. The only exception to this was the use of Sennheiser HD650 headphones in experiment 7. In all experiments, trial-by-trial feedback was provided. Where applicable, interaural time delays were imposed on the stimuli by adding a ramp function to the phase spectrum at one ear. For the detection and discrimination stimuli that form the basis of the experiments described in chapter 4, independent noises were generated for the fringes of diotic noise preceding and following the probe, and for the probe itself. The filter/window fitting algorithms and sigmoid functions used to obtain thresholds from psychometric functions were programmed in C.

2.2 THE ADAPTIVE TRACK

An adaptive procedure was one of the methods used to obtain thresholds, in which an algorithm adjusted the stimulus parameters for each trial based on the participant's previous responses. The adaptive tracks employed in experiments 1, 3, and 7 are 2-down/1-up tracks that follow the Up-Down or Staircase method described by Levitt (1971), where the stimulus magnitude is increased after an incorrect response and decreased after two consecutive correct responses. 2-down/1-up tracks are employed in these experiments as they were used in the study of Bernstein et al. (2001), and a replication of the first experiment in that study was included as a condition in experiment 3. The step size is the magnitude by which the stimulus parameter was increased or decreased following a correct or incorrect response, and was decreased following several reversals after the track was initialised. Following a set number of initial reversals, the measurement phase of the track began. The measurement phase consisted of a set number of reversals, and following the final reversal the track was

terminated and the average of the reversals obtained within the measurement phase was taken as threshold. Figure 2.1 illustrates a typical adaptive track.

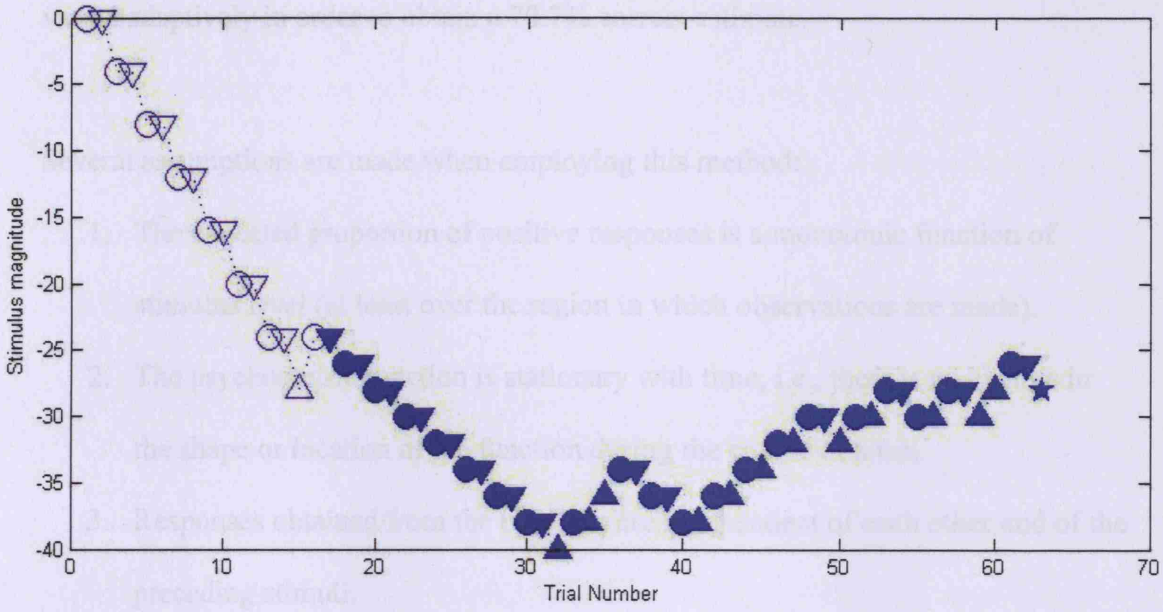

Fig. 2.1. An example adaptive track. The figure illustrates a 2-down1-up adaptive track that begins taking measurements after the second reversal, and terminates after ten reversals. The initial step size is 4 (these steps are signified by the open symbols), and is reduced to 2 (closed symbols) after the second reversal. The arrows indicate the direction of the step, circles indicate that a correct response has been elicited and the magnitude of the stimulus stays constant, and the star indicates the termination of the track. The figure is taken from an adaptive track collected in experiment 7, and in this example the threshold stimulus magnitude is -34.

In an adaptive track procedure, Levitt (1971) described how the stimulus magnitude is changed only after a sequence of responses that are classed as either UP or DOWN. The 2-down/1-up track converges at the stimulus magnitude at which the probability of a sequence of DOWN responses is equal to the probability of an UP response sequence (i.e. 0.5). The probability of obtaining a DOWN response sequence (two

consecutive correct responses) is $[P(X)]^2$. The stimulus magnitude converges on the value of X at which $[P(X)]^2=0.5$, or $P(X)=0.707$. Thus, the stimulus magnitude is varied adaptively in order to obtain a 70.7% correct estimate.

Several assumptions are made when employing this method:

1. The expected proportion of positive responses is a monotonic function of stimulus level (at least over the region in which observations are made).
2. The psychometric function is stationary with time, i.e., there is no change in the shape or location of the function during the course of a test.
3. Responses obtained from the observer are independent of each other and of the preceding stimuli.

When linear step sizes are used, the number of steps across the psychometric function may change across condition, resulting in different conditions having different lengths of adaptive track (Saberi, 1995). The use of logarithmic steps allows the same step size to be used for a range of stimulus conditions, thus preventing the length of the adaptive track changing.

Adaptive tracks were employed in experiments 1 and 3, in which the stimulus parameter that was adjusted was ITD, and experiment 7, in which the signal level was adjusted. In experiments 1 and 3, each adaptive track started with the ITD set to 500 μs, and the step size of the track was initially fixed to correspond to a factor of 0.2, which was reduced to a factor of 0.05 following two reversals. In experiment 7, a set arbitrary initial level was chosen, and the initial step size was 4 dB, which was reduced to 2 dB following two reversals. In all of the experiments that used adaptive tracks, following two initial reversals, the track was terminated after a measurement

phase that consisted of ten reversals. The parameters of the adaptive track were chosen following pilot testing. The initial magnitude of the stimulus parameter was chosen to be highly detectable. The large initial step size increased the rapidity with which the listener approached threshold, and the smaller step size in the measurement phase allowed threshold to be measured with greater sensitivity. Ten reversals were chosen to make up the measurement phase to allow the stimulus magnitude to converge around the final threshold value.

2.3 PSYCHOMETRIC FUNCTIONS

Measuring psychometric functions provides an alternate method to the adaptive track for measuring thresholds. Instead of adaptively varying the stimulus parameter to be measured, blocks of stimuli are generated with different magnitudes of stimulus parameter, and the percentage of correct responses at each magnitude is measured. A psychometric function is then constructed by plotting percentage correct against stimulus parameter magnitude. To construct the psychometric functions measured in the experiments in the thesis, thresholds were fitted with a logistic function (the logistic function was chosen as it has previously been demonstrated to provide a good overall fit to psychometric function data, e.g. Amitay et al., 2006) with two free parameters; slope and threshold, using the following formula:

$$y = \min + \frac{\max - \min}{(1 + e^{(-s(x-t))})}$$

s is the slope, t is the threshold, and min and max are the lowest and highest expected scores (50 and 100% respectively, for all psychometric functions in the thesis). The Simplex method (Nedler and Mead, 1965) was applied to obtain the parameters that produced the best fit to the data by minimising the sum-squared error between the empirical data and the predicted thresholds. In a 2A-FC task, if the participant

performs at chance, they will score 50% correct, and if they make no mistakes they will score 100%. The 75% point of the psychometric function is most often taken as threshold, as the slope of the psychometric function is steepest at this point (see Fig. 2.2).

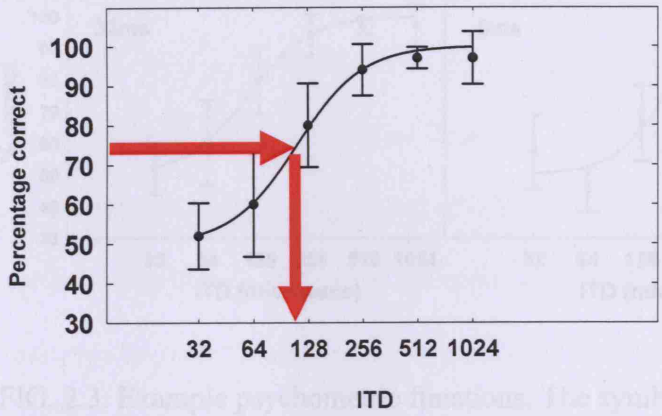

FIG. 2.2. An example of a psychometric function. In this example, the stimulus parameter is ITD. The 75% correct point is taken as threshold, which can be read off to be 110 μs.

The harder the task, the more errors the participant is likely to make. More errors tend to result in a rightward shift (or increase) in threshold. Figure 2.3 shows a range of psychometric functions for different levels of task difficulty.

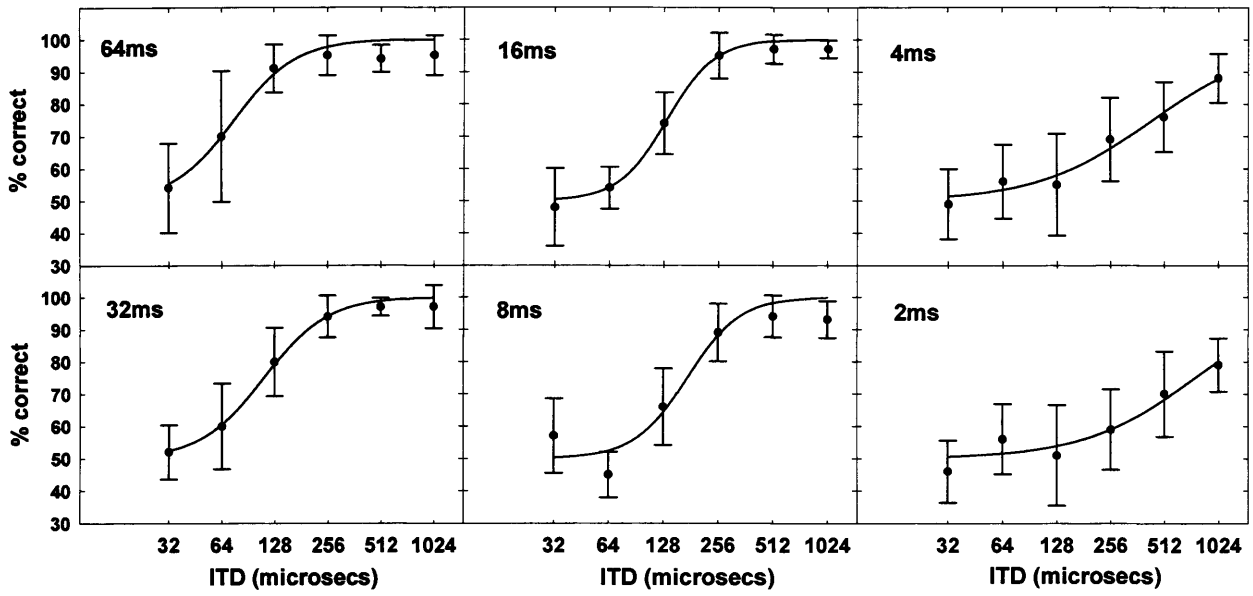

FIG. 2.3. Example psychometric functions. The symbols show observed threshold data. The lines show the closest match to this data achieved by scaling the predicted data to the empirically averaged data. Psychometric functions are illustrated for tasks that range from easiest (labelled as 64 ms), to hardest (labelled 2 ms). The data in this example is taken from a listener performing the detection task in experiment 4. The error bars represent ±1 standard error.

The 70.7% point of the function can also be taken in order to directly compare thresholds taken by measuring psychometric functions with thresholds taken using an adaptive track procedure with a 2I-2AFC task. Six-point psychometric functions were obtained in experiment 4, and eight-point psychometric functions were obtained in experiments 2, 5 and 6 (where the range of ITDs over which the psychometric function was measured was greater).

2.4 AUDITORY FILTER AND TEMPORAL WINDOW MODELLING

The temporal window/auditory filter fitting procedure employed throughout the thesis was similar to that used by Plack and Moore (1990) to model the monaural temporal window. In this procedure, several parameters that describe the window/filter to be modelled to a given dataset are given initial values which are then adjusted by an algorithm in order to achieve a fit with the lowest possible sum-squared error. Initially, a candidate function with a specified shape is chosen, and the values of the parameters associated with the function specified. A set of predicted thresholds are then produced on the basis of the specified parameters. Off-time/off-frequency listening is also explicitly modelled at this stage (except in the case where an exponential candidate function is assumed). The goodness of the fit is then evaluated by calculating the residual error between the calculated data and the empirical data. The parameters are then altered and fed back into the algorithm, in order to calculate whether the fit based on the modified parameters provides a better fit to the data than the fit based on the previous parameters. The algorithm continues to adjust the parameters in an iterative manner until it produces the best fit to the data using the specified candidate function. The Simplex method (Nedler and Mead, 1965) was implemented in the window-fitting program in order to obtain the window/filter parameters that produced the best fit to the data. The algorithm repeated the fitting procedure and adjusted the parameters in an iterative manner until it produced the best fit to the data using the specified candidate function. The steps involved in fitting a window/filter are described in detail below.

2.4.1 Shape selection.

The first step of window/filter fitting is the choice of candidate window/filter shape, or base function. Each window/filter can be modelled as being composed of either a peak (described as a simple fit), or a peak and skirt (described as a skirt fit), where the window/filter is made up of two functions which are both the same shape (Moore, 1989). In a skirt fit, the slope of the filter decreases after the attenuation reaches a certain value, after which the slope is described by the second function. The peak and skirt can be either symmetric or asymmetric. For each window/filter dataset, Gaussian, exponential, and roex shapes were modelled for subsequent comparison as to which base function fit the data best.

Each fit can be described by a number of parameters. When an asymmetric fit is assumed, the parameters *tpu* and *tpl* describe the upper and lower lobes of the window/filter respectively (throughout the thesis, *t* is used to denote a window/filter parameter, *p* denotes the peak, and *u* and *l* denote upper and lower lobes). When a symmetric fit is assumed, *tpu* and *tpl* are both given the same value. When a skirt fit is modelled, the parameter *tsl* (where *s* denotes skirt) defines the slope of the window/filter at times and frequencies far from the temporal midpoint of the probe or the centre frequency respectively, and *w* (which is used to denote the weighting of the skirt) defines the point where the skirt of the function begins. *tpu*, *tpl* and *tsl* are measured in ms (for temporal windows) or Hz (for auditory filters), and *w* (in this thesis) is reported in dB. Figure 2.4 illustrates simple and skirt fits, and the parameters associated with each fit.

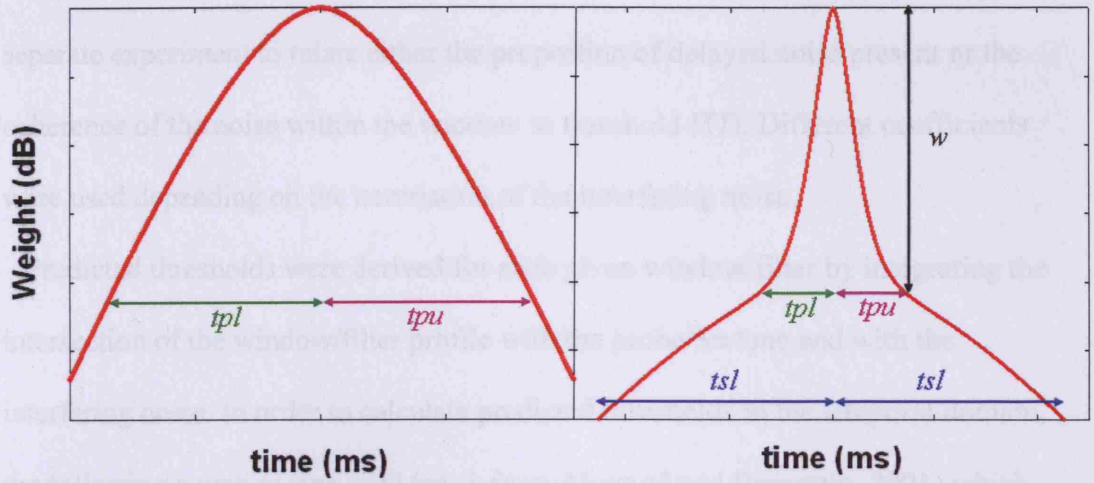

FIG. 2.4. Schematic simple and skirt window/filter shapes. The left panel shows a simple roex shape, and the right panel shows a skirt roex shape. Both shapes are labelled with their appropriate parameters.

When modelling the temporal window, a symmetric window was assumed, as (unlike a signal detection task such that used by Culling and Summerfield, 1998) the laterality task used to measure thresholds did not allow for the measurement of asymmetry in the window. Thus, a simple window consisted of a 1-parameter fit, and a skirt window consisted of a 3-parameter fit. When modelling the auditory filter, the filter was assumed to be asymmetric, thus simple filters consisted of 2-parameter fits, and skirt filters consisted of 4-parameter fits.

2.4.2 Threshold prediction.

When temporal windows were modelled, predicted thresholds were calculated by computing the integral of the area occupied by the probe, which was divided by the total integral of the window. This value represented the proportion of delayed noise present within the window when a correlated interferer was modelled. When an uncorrelated interferer was modelled, it also represented the coherence of the noise

within the window. Threshold ITD was determined using a coefficient derived from a separate experiment to relate either the proportion of delayed noise present or the coherence of the noise within the window to threshold ITD. Different coefficients were used depending on the correlation of the interfering noise.

Predicted thresholds were derived for each given window/filter by integrating the intersection of the window/filter profile with the probe/Sπ tone and with the interfering noise. In order to calculate predicted thresholds in the temporal domain, the following equation was used (eq. 1 from Akeroyd and Bernstein, 2001) which expresses the output of the temporal window (W) in the case of a diotic interferer. τ is the time relative to the peak of the window, and $w(\tau)$ is the function that defines the assumed shape of the temporal window. The overall duration of the probe is $|p_1|+|p_2|$and the overall duration of the stimulus is $|D_1|+|D_2|$.

$$W = ITD \frac{\int_{-p_1}^{p_2} w(\tau)d\tau}{\int_{-D_1}^{D_2} w(\tau)d\tau'}$$

The following equation (eq. 3 from Akeroyd and Bernstein, 2001) assumes that empirical threshold ITDs corresponded to a constant ITD, W_0 at the output of the window, and is obtained by rearranging eq. 1 of Bernstein and Akeroyd (2001) above.

$$thresholdI\,TD = W_0 \frac{\int_{-D_1}^{D_2} w(\tau)d\tau}{\int_{-p_1}^{p_2} w(\tau)d\tau'}.$$

The equations for the window functions and their integrals are listed below, as is the equation defining the 'erf' function (from Culling and Summerfield, 1998). Each function (in square brackets) represents the integral of one side of the filter. t represents time or frequency depending on whether a temporal window or auditory filter is being fitted, p represents one lobe of the peak of the function, and s represents

one lobe of the skirt of the function. In order to calculate the predicted thresholds, once a candidate function had been chosen, the relevant parameters were substituted into the appropriate equation. Asymmetry in the peak was modelled by substituting the parameters *tpl* and *tpu* for *p* for each lobe of the peak of the function. When the temporal window was modelled to be symmetric, *tpl* and *tpu* were given the same value. For both symmetric and asymmetric fits, the skirt was modelled to be symmetric, and *tsl* was substituted for *s*. Although a range of window shapes have been previously assumed and modelled to temporal window data (e.g. triangular or rectangular, Kollmeier and Gilkey, 1990), exponential, roex and Gaussian windows tend to provide better fits to empirical data (e.g. Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998), and are thus modelled to the temporal window data throughout this thesis. In a previous study, Bernstein et al. (2001) concluded that skirt fits provided better fits to temporal window data than simple fits. In order to examine this claim, simple and skirt functions were modelled to the temporal window data and compared in the thesis. Generally, roex filters have been assumed when measuring the monaural auditory filter (e.g. Glasberg and Moore, 2000). Simple and skirt Gaussian, exponential and roex filters were modelled to the auditory filter data in order to assess the best-fitting filter shape.

Simple exponential:

$$\int [e^{-t/p}]dt = pe^{-t/p}.$$

Simple Gaussian:

$$\int [e^{-(t/2p)^2}]dt = p\sqrt{\pi}erf(t/2p).$$

Simple rounded exponential:

$$\int [1+2t/p)e^{-2t/p}]dt = (p+t)e^{-2t/p}.$$

Weighted sum of two exponentials (peak and skirt):

$$\int [(1-w)e^{-t/p} + we^{-t/s}]dt = (w-1)pe^{-t/p} - wse^{-t/s}.$$

Weighted sum of two Gaussians (peak and skirt):

$$\int [(1-w)e^{-(t/2p)^2} + we^{-(t/2s)^2}]dt = (w-1)p\sqrt{\pi}erf(t/2p) + ws\sqrt{\pi}erf(t/2s).$$

Weighted sum of two rounded exponentials (peak and skirt):

$$\int [(1-w)(1+2t/p)e^{-2t/p} + w(1+2t/s)e^{-2t/s}]dt = (w-1)(p+t)e^{-2t/p} - w(s+t)e^{2t/s}.$$

The 'erf' function:

$$erf\ (x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$$

When the upper lobe of the window/filter was calculated, *t* was constrained to be greater than zero. When the lower lobe was calculated, *t* was constrained to be equal to or less than zero.

Using the final parameter values provided by the algorithm, the ERB (measured in Hz) or the ERD (measured in ms) could be calculated. The ERB of exponential and roex filters can be calculated using the following equations:

$$ERB = (1-w)(tpu+tpl)+w(tsl+tsu)$$

If a Gaussian shape was assumed:

$$ERB = \sqrt{\pi}((1-w)(tpu+tpl)+w(tsl+tsu))$$

When calculating the ERD, the following equations (distinct from those used to calculate the ERB) were used to account for the fact that the fits were symmetric. These follow from the equations used to calculate the ERB when *tpl* = *tpu*, and *tsl* = *tsu*:

$$ERD = 2((1-w)tp+wts)$$

If a Gaussian fit is modelled:

$$ERD = 2\sqrt{\pi}((1-w)tp+wts)$$

When auditory filters were modelled, the program was run with the parameters initially set to *tpl* = 100, *tpu* = 100, *tsl* = 100, and *w* = -100. As the best fits were observed to depend to an extent on the initial parameters the program was given, an additional fit was run with the weighting parameter set to -10 and the other parameters set to 100. Of the two sets of fits, the set with the lowest χ^2 value (N.B. χ^2 is the measure of goodness-of-fit, see section 2.4.3) was accepted as the best fit. When temporal windows were modelled, the program was run with the parameters initially set to *tp* = 10, *ts* = 100, and *w* = -10. As when auditory filters were fitted to the data, the best fits were observed to depend somewhat on the initial parameters the program was given. An additional fit was run with the weight parameter set to -100 and the other parameters set to 100.

Predicted thresholds for auditory filters were measured using the following procedure. For a given set of thresholds and No bandwidths, the effective correlation within the filter was calculated for each datapoint at the frequency in question by dividing the integral of the area occupied by the No noise by the total integral of the window (N.B. the tone was assumed to occupy an infinitesimally small area within the window, and was thus not assumed to be integrated with the rest of the noise within the window). The equation for calculating the BMLD is shown below, where T_{No} is the mean threshold in linear units observed at No bandwidths of 400 Hz, and T_{Nu} is the mean threshold in linear units for No bandwidths of zero, and ρ is the correlation:

$$BMLD = -10\log(T_{Nu} - \rho(T_{No} - T_{Nu}))$$

If the signal was not at the same frequency as the centre frequency of the filter, the predicted threshold was increased in proportion to the amount of attenuation caused by the filter.

Off-frequency listening was explicitly modelled for Gaussian and roex filter shapes by using a single parameter minimization method to position a candidate window/filter such that it predicts the lowest threshold. The centre frequency of the modelled filter was varied relative to the centre frequency of the No band in an iterative manner until the lowest disparity between predicted and empirical data was obtained.

2.4.3 Fit assessment.

The window/filter that provided the best fit produced the closest set of predicted thresholds to the empirical data. The χ^2 statistic was used to calculate the goodness of fit of each of the base functions. This statistic has been previously used to compare

the goodness of fit for a variety of temporal window functions (Culling and Summerfield, 1998). The statistic was generated using the following formula which took into account the number of free parameters in the filter/window shape:

$$\chi^2 = \frac{\text{SS error}}{\text{Degrees of freedom}} = \frac{\text{SS error}}{\text{No. datapoints} - \text{No. parameters}}$$

The lower the value of χ^2, the better the filter fit. Lower χ^2 values are generally obtained by having many datapoints and fewer parameters. There are cases where simple fits result in systematic variations between the fitted data and the empirical data, and a lower χ^2 value (i.e. a better fit) is obtained using a skirt fit despite increasing the number of parameters because the sum-of-squares error is reduced.

CHAPTER 3: DELAY MASKING

In order to model the shape and duration of the temporal window, the effect on perception of the presence of interfering noise within the window needs to be established. Interfering noise can take the form of diotic noise, which has an interaural correlation of 1 and an interaural time delay of zero, or uncorrelated noise, which has an interaural correlation of zero and no delay. When interfering noise enters the window, it is integrated with any delayed noise present within the window. In the following chapter, the effect of mixing delayed noise with interfering noise on lateralization is addressed.

When delayed noise and interfering diotic noise are present within the temporal window, the window may integrate the delayed and the diotic noise resulting in an internal or effective ITD that is smaller than the ITD conveyed by the delayed stimulus alone. As a result, the ITD of the delayed stimulus needs to be increased so that the internal ITD is of a magnitude that exceeds the threshold needed for detection. Bernstein et al. (2001) assumed that when broadband delayed target noise ($N\tau$) is heard in the presence of interfering correlated noise (No), the effective detectable ITD is a weighted average of the ITDs of the different noises present within the window. It was assumed that ITD threshold doubled for each halving of the proportion of delayed noise present within the window in the presence of diotic noise; that is, that the interfering diotic noise dilutes the delayed noise. The assumption that ITD is diluted in a 1:1 fashion by interfering noise is fundamental to the analysis of Bernstein et al. (2001). This assumption has not been tested empirically, and is the focus of the first experiment.

3.1 EXPERIMENT 1. DILUTION WITH CORRELATED NOISE.

3.1.1 Introduction.

In order to test the dilution hypothesis, participants were presented with stimuli designed to reflect the integrating properties of the temporal window. Mixing delayed noise with simultaneous diotic noise is equivalent to the action of the temporal window when it integrates temporally contiguous delayed and interfering noise. The temporal window is expected to average the ITD of the noises present within the window, thus the more interfering diotic noise that is presented within the window, the higher the threshold ITD becomes. By presenting listeners with stimuli made up of mixtures of noises with various proportions of delayed noise present, the change in threshold ITD as more correlated interfering noise is introduced into the temporal window can be examined.

Participants were given a 2I-2AFC discrimination task, where interaurally correlated noise was delayed and mixed with diotic noise, so that each interval was composed of delayed noise and diotic noise mixed at a pre-determined proportion (see Figure 3.1).

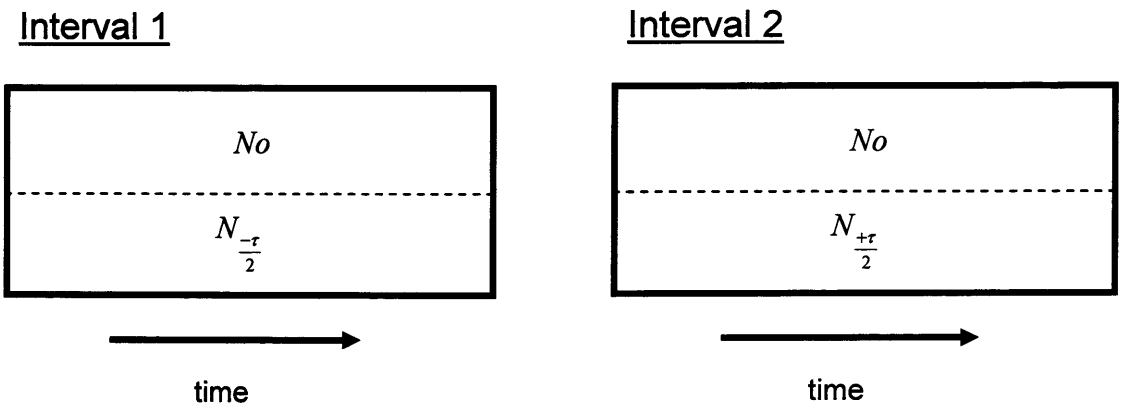

FIG. 3.1. Stimuli presented in experiment 1 (Dilution with correlated noise).

Half the delay was imposed on interval 1, and half the delay, in the opposite direction to interval 1, was imposed on interval 2. It was hypothesised that for each halving of the proportion of delayed noise present, the threshold ITD would double. Thus, on a log-log plot, as the proportion of delayed noise decreased and more diotic noise was mixed into the intervals, thresholds would increase in a linear manner on a 1:1 slope. The slope describing ITD threshold change is defined as the Correlated Masking Coefficient, or CMC.

3.1.2 Method.

Participants

Six participants took part in the experiment. One was male and five were female, aged between 18 and 25. All participants except for ET and RH had previous experience with psychophysical experiments. Untrained participants received at least 5 hours of training before data collection. They were paid upon completion.

Apparatus/Materials

Listeners' performed a 2-interval discrimination task. To create the first interval, three noises were generated: two independent Gaussian noises (N_1 and N_3), and a delayed copy of N_1 (N_2). The magnitude of the delay imposed on N_1 was half that of the total delay. To keep the power constant, the noises were scaled in amplitude in the following proportions (p), so that p is the proportion of the total power of the stimulus that is made up of delayed noise:

Channel one: $\sqrt{p}N_1 + \sqrt{1-p}N_3$

Channel two: $\sqrt{p}N_2 + \sqrt{1-p}N_3$

The second interval was created using the same procedure, except that the delay was lateralized to the opposite side of the head to interval 1. Each interval had a duration of 100 ms, with an inter-stimulus interval of 500 ms.

Design

The experiment had a repeated-measures design. The independent variable was proportion of delayed noise and was composed of seven conditions (1, 0.75, 0.563, 0.422, 0.316, 0.237, and 0.178). The dependent variable was threshold ITD.

Procedure

The listener's task was to identify in which direction the sound image moved from interval to interval (left to right or right to left), corresponding to the ITDs embedded within each interval. Participants pressed 1 if the noise moved from right to left, and 2 if it moved left to right. The direction of sound movement was randomised and trial-by-trial feedback was provided. The ITD was varied adaptively in order to obtain a 70.7% correct estimate (Levitt, 1971). Each adaptive track started with the ITD set to 500 μs. Initially, the step size of the adaptive track corresponded to a factor of 0.2, and was reduced to a factor of 0.05 following two reversals. Threshold was obtained after 10 reversals. The last 10 reversals comprised a measurement phase, and the average of the reversals within the measurement phase was taken as threshold. Nine experimental runs were performed, and measurement began after performance had stabilised; experimental thresholds were taken by averaging the remaining runs after the listener achieved thresholds below 500 μs at each proportion of delayed noise. Thus, experimental thresholds for AK, CH, RH and SW were composed of the mean of nine runs. HM and ET's thresholds took longer to stabilise below 500 μs, thus

HM's experimental thresholds were composed of the mean of eight runs, and ET's experimental thresholds were composed of the mean of four runs.

3.1.3 Results.

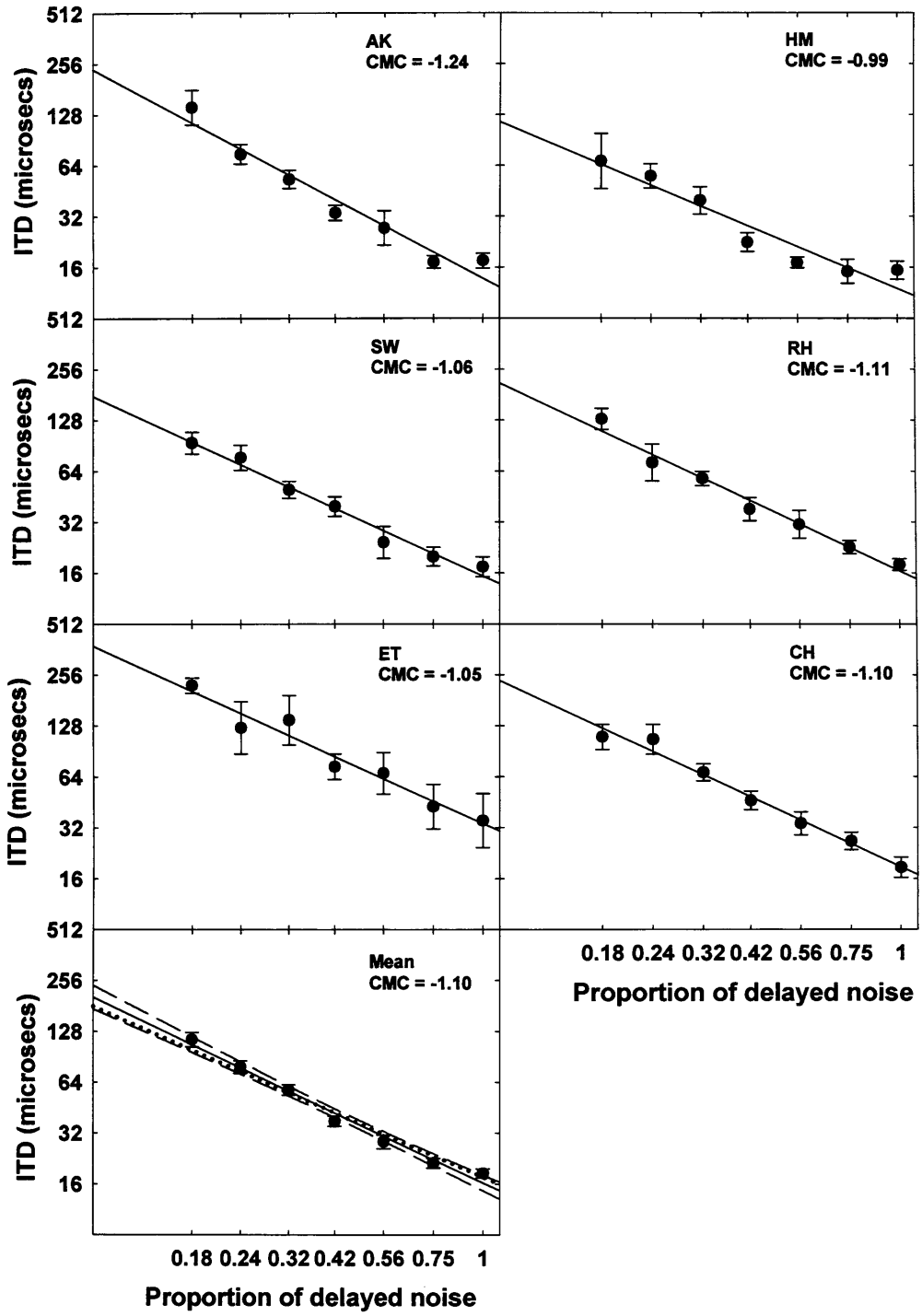

FIG. 3.2. Results for the experiment 1 (Dilution with correlated noise). Solid regression lines are plotted for individual listeners and the mean. The dotted line represents a slope of -1. Dashed lines represent 95% confidence intervals, calculated using Sigmaplot. Both axes are plotted logarithmically.

Figure 3.2 shows threshold ITDs measured at different proportions of delayed noise mixed with diotic noise. When $p = 1$, all the noise within each interval was delayed, producing the lowest thresholds for the three listeners (a geometric mean of 18.5 μs). Threshold ITD was found to approximately double for each halving of delayed noise present. The slope of the regression line of the mean observed data was -1.10, and accounted for 98% of the variance. The dotted line represents predicted values based on the assumption of dilution with a CMC of -1, and falls within the 95% confidence intervals of the mean data.

3.1.4 Discussion.

The results validate the dilution assumption. As the proportion of delayed noise decreased and more interfering diotic noise was mixed into the intervals, thresholds increased in a linear manner on a 1:1 slope, so that for each halving of the delayed noise present, the threshold ITD doubled. The 95% confidence intervals of a regression to the data encompassed the CMC of -1 that represented predicted values based on the assumption of dilution (dotted line in fig 3.2), demonstrating that the observed data was not significantly different from prediction based on the assumption of dilution.

It was observed that measuring the higher thresholds was problematic, because accuracy of lateralization judgements does not increase indefinitely with increasing ITD. As discussed in Chapter 1, the ecological range of naturally occurring ITDs is approximately 600 μs (Kuhn, 1977). ITDs of broadband noise within this range and up to 1 ms are described as fused and lateralized to one side of the head. Increasing ITDs beyond the physiological range causes the sound image to stay lateralized to one side, but the image becomes broader and more diffuse until it fills the head. Increasing

the magnitude of the ITD only makes the task easier for the listener up to approximately 700 μs, beyond which the task becomes rather more difficult, as there is a change in image cue (from lateralized and fused to lateralized and diffuse) when ITD magnitude rises above 700 μs, which could confuse the listener and cause them to make errors (Mossop and Culling, 1998). The use of ITD as the dependent variable in an adaptive track can thus violate assumption #1 of Levitt (1971):

'The expected proportion of positive responses is a monotonic function of stimulus level (at least over the region in which observations are made)' (p. 468).

The change in cue associated with large ITDs may affect the listener in two possible ways. First, if the track begins with a low value of ITD and the listener makes errors that bring the ITD above 700 μs, the listener is unlikely to be able to bring the ITD back down again, as the cue changes and increasing the magnitude of the ITD is no longer more likely to elicit a correct response. Second, if the track begins with the ITD magnitude set above 700 μs, the participant is more likely to make an error early on in the track, causing the step size to decrease and causing the listener difficulty in bringing the ITD down to a low threshold. It is likely that the former of these cue changes disrupted listeners at low proportions of delayed noise, as the starting point of the track was set at 500 μs.

In addition, the adaptive track was set to decrease the step size from a factor of 0.2 to a factor of 0.05 after the first two reversals, and obtain a threshold by averaging the ITD over the last 10 reversals. An error that is made early in the adaptive track can have a disproportionate affect on threshold ITD because it causes a reduction in step size, and the commencement of the measurement phase before the ITD approaches threshold. ITD can be reduced following a number of correct responses, but any further errors will result in an elevated final threshold. Following these observations,

experimental thresholds were taken by averaging the remaining runs after the listener achieved thresholds below 500 μs at each proportion of delayed noise. On the basis of these criteria, only two of the six participants had runs excluded from the final analysis; HM had one run removed, and ET had 5 runs removed.

Gathering psychometric functions constitutes an improved and more reliable method of obtaining thresholds compared to the adaptive track, as this method does not have the ITD cue change problem associated with it. Using this method, stimuli are produced with a range of fixed ITDs, and the percentage of correct responses at each ITD is recorded. From this data, the 71% correct point of the psychometric function can be taken as threshold, which can be directly compared to previous thresholds obtained using the adaptive track method.

Despite these problems, the validation of the dilution hypothesis enables the dilution concept to be utilised when temporal windows are fit to data that incorporates diotic interfering noise in the presence of delayed noise. The following experiment examines the effect of interaural coherence on ITD threshold, and examines whether thresholds change in the presence of an uncorrelated interferer in the same manner as they do in the presence of correlated interfering noise. Thresholds were taken by measuring psychometric functions.

3.2 EXPERIMENT 2. DILUTION WITH UNCORRELATED NOISE.

3.2.1 Introduction.

In their study of temporal resolution, Bernstein et al. (2001) assumed that when broadband delayed target noise (Nτ) is heard in the presence of interfering uncorrelated noise (Nu), the effective detectable ITD is a weighted average of the correlations of the different noises present within the temporal window. Thus,

Bernstein et al. assumed that the dilution concept that applied to diotic interfering noise was also valid for uncorrelated interfering noise. However, the effect of interaural correlation on the ability of participants to centre a noise was investigated by Jeffress, Blodgett and Deatherage (1962), who found that ITD threshold did not change in a manner consistent with the dilution concept.

In that study, correlated and uncorrelated noises were mixed in each ear to obtain interaural correlations of 1, 0.75, 0.5, 0.25, 0.2, 0.15, 0.1 and 0. Participants adjusted a delay line in the presence of continuous noise in order to centre the noise in their heads. The standard deviation of each participant's centring judgement was measured as a function of interaural correlation. The slope relating standard deviation of ITD to correlation was -0.48, which was considerably shallower than the slope of -1 that would have been observed if the 1:1 linear dilution concept applied.

In the current study, the effect of correlation on ITD threshold was assessed using stimuli and a procedure comparable to experiment 1. Listeners were again given a 2I-2AFC discrimination task, where the interaural coherence of each interval was manipulated by mixing correlated and uncorrelated noise before applying the delay (see Fig. 3.3). When delayed and uncorrelated noises are integrated, the coherence is equal to the proportion of delayed noise in the stimulus, so the proportions of delayed noise are numerically equivalent to the coherences in this case.

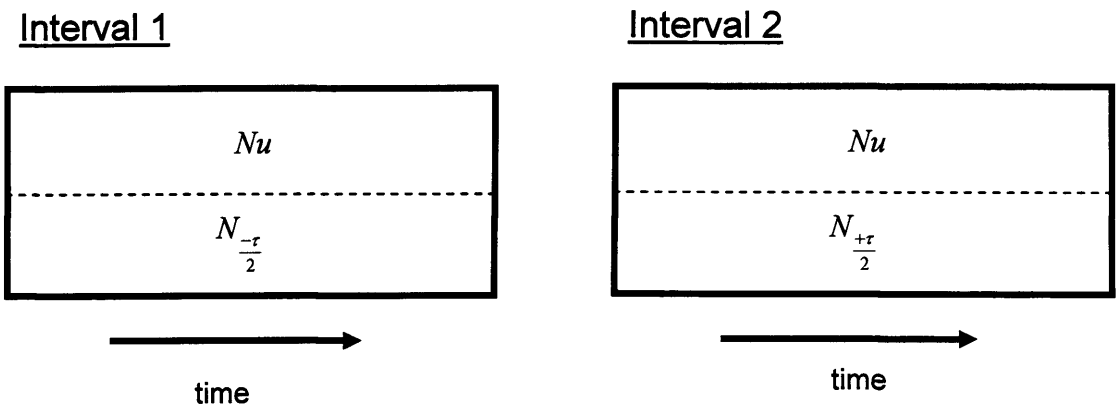

FIG. 3.3. Stimuli presented in experiment 2 (Dilution with uncorrelated noise).

Thresholds were gathered by taking psychometric functions. To examine any effects of stimulus duration and compare the results of the current experiment with that of experiment 1 (100 ms stimuli duration) and Jeffress et al. (1963, continuous stimuli), participants were tested at three stimulus durations (100, 500, and 1 s). In contrast to the previous experiment, the slope describing ITD threshold change is defined as the Uncorrelated Masking Coefficient, or UMC.

3.2.2 Method.

Participants

Three participants took part in the experiment. One was male and two were female, aged between 18 and 25. All participants had participated in experiment 1. They were paid upon completion.

Apparatus/Materials

Listeners' performed a 2-interval discrimination task. Two independent Gaussian noises (N_1 and N_2) were generated and mixed using the 'two-noise' method: one

channel was presented containing the common noise N_1, and the other channel contained a mixture of two noises N_1 and N_2 mixed in the proportion/coherence (ρ):

$$\rho N_1 + \sqrt{1-\rho^2} N_2$$

An interaural delay was added to one channel after noise mixing. Intervals were presented so that the first interval was lateralized to one side of the head, and the second lateralized to the other side of the head. The inter-stimulus interval was 500 ms.

Design

The experiment had a repeated measures design. The independent variables were coherence and duration. Participants were tested at coherences of 1, 0.75, 0.5, 0.25, 0.2, 0.15, 0.1, and 0.05, at stimuli durations of 100, 500, and 1 s. The dependent variable was threshold ITD, measured in μs.

Procedure

The listener's task was to identify in which direction the sound image moved between the two intervals (left to right or right to left), corresponding to the ITDs embedded within each interval. Participants pressed 1 if the noise moved from right to left, and 2 if it moved left to right. The direction of sound movement was randomised, and trial-by-trial feedback was provided. Initially, a block of trials was presented with a coherence of 1, 0.5, 0.2 and 0.1, and the order in which the eight ITDs were presented within the block was randomised. A block of trials with a coherence of 0.75, 0.25, 0.15 and 0.05 followed. The procedure was repeated so that 20 trials were repeated for each ITD and probe duration. Each block consisted of 8 ITDs x 4 coherences x 20 repetitions = 640 trials in all. Stimuli were presented with ITDs of

1024, 512, 256, 128, 64, 32, 16 and 8 μs. Participants were given at least six training runs at each coherence for each stimulus duration. They then performed three experimental runs. A psychometric function was fitted to the average of the three runs, and the 71% point of the function was taken as threshold in order to compare thresholds with those taken in experiment 1 with the adaptive track procedure.

3.2.3 Results.

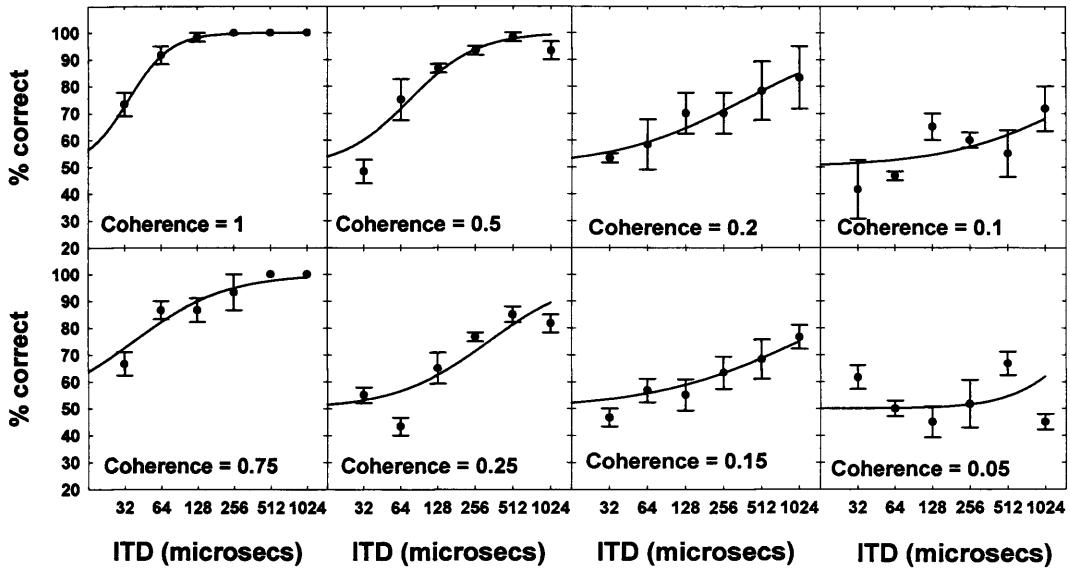

FIG.3.4. Psychometric functions taken from listener AK performing the task at a stimulus duration of 1 sec. The different panels indicate results at different coherences. The symbols show the observed threshold data. The lines show the closest match to this data achieved by scaling the predicted data to the empirically averaged data. The x axis is plotted logarithmically.

Fig. 3.4 shows psychometric functions plotted at each coherence for listener AK at a stimulus duration of 1 s. The psychometric function becomes visibly shallower as the coherence decreases.

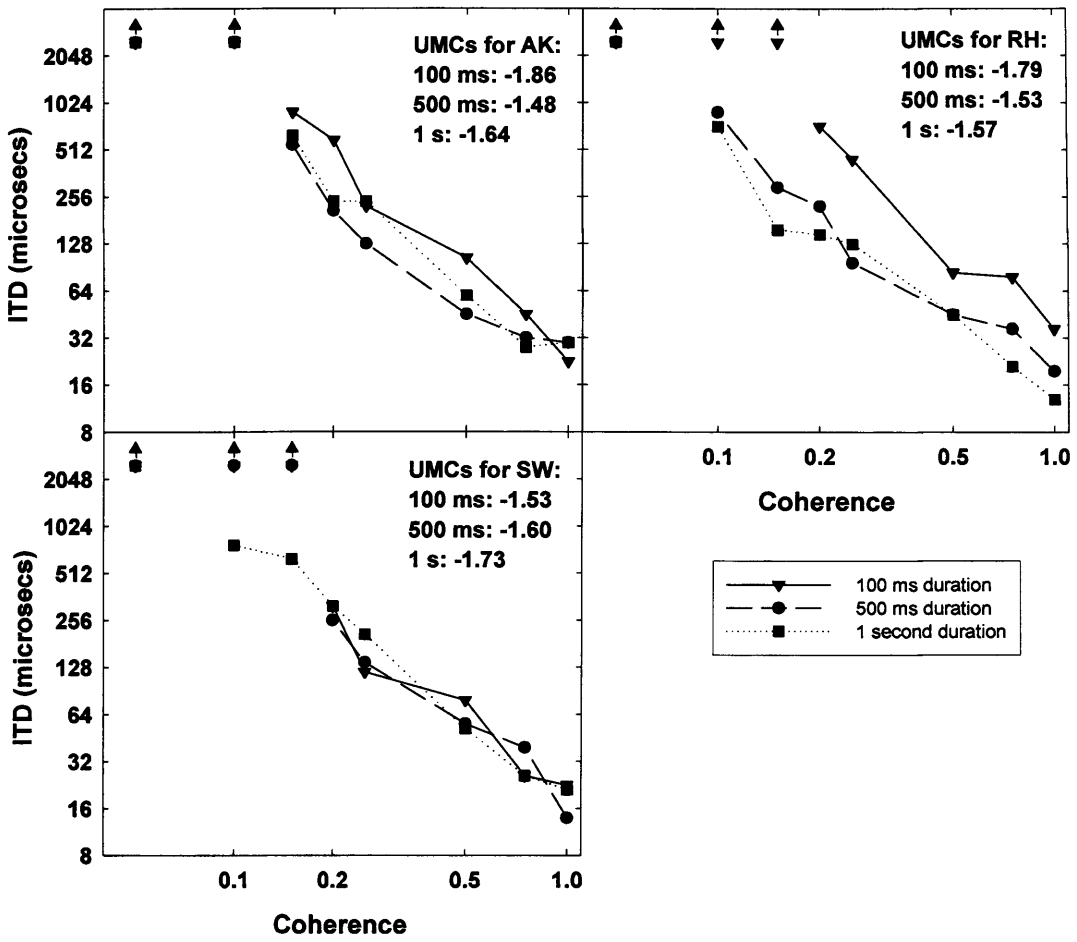

FIG. 3.5. Individual data for experiment 2 (Dilution with uncorrelated noise). Thresholds represent 71% discriminability. Upward pointing arrows indicate that a measure of threshold could not be obtained for those coherences as the 71% point of the psychometric function fell outside the measured range. Both axes are plotted logarithmically.

Fig. 3.5 describes individual data for the three stimulus durations. UMCs range from -1.48 to -1.86 (N.B. for clarity, regression lines are not plotted). When the coherence was unity, all the noise within each interval was delayed, producing the lowest thresholds for the three listeners. Threshold ITD was found to more than double for each halving of coherence (which is also the proportion of coherent noise power).

Thresholds at the lowest coherences could not be measured as the 71% point of the psychometric function fell outside the measured range.

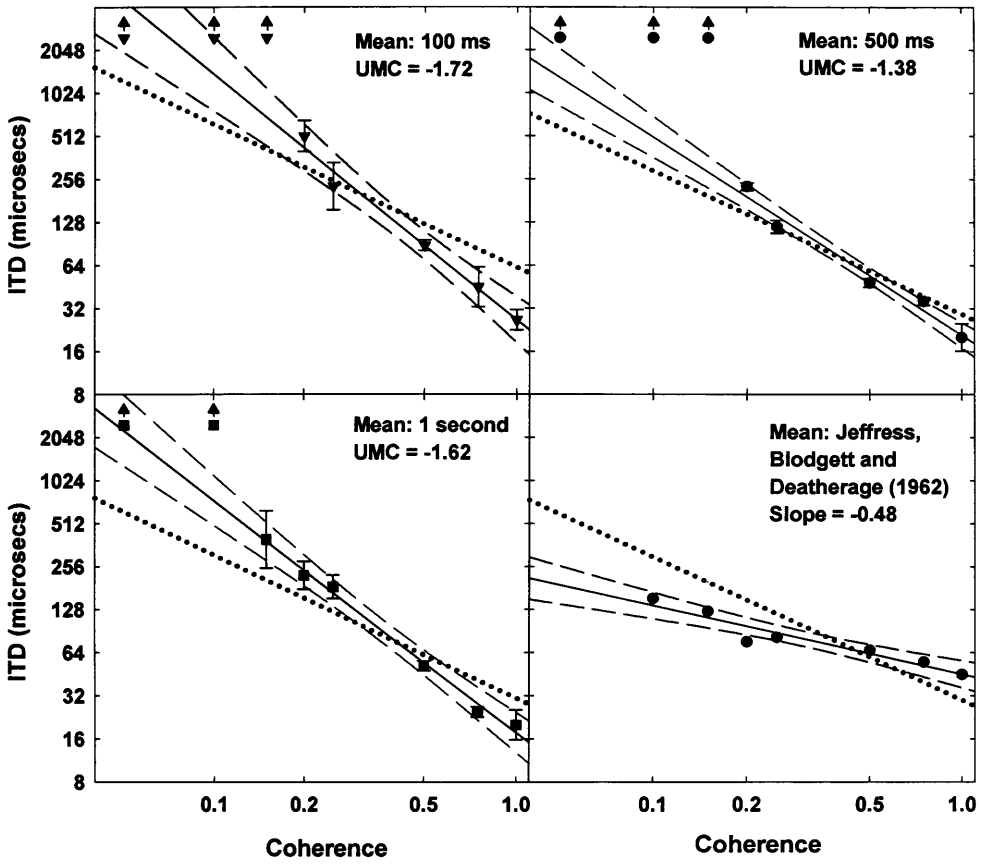

FIG. 3.6. Mean results for experiment 2 (Dilution with uncorrelated noise), and from Jeffress et al., (1962). Solid regression lines are plotted for each listener. The plotted thresholds are 71% points averaged from the last three experimental runs. The dotted line represents thresholds with a slope of -1. Dashed lines represent 95% confidence intervals.

Figure 3.6 describes the mean of the data across listeners for the three stimulus durations. The UMC of the thresholds at each stimulus duration was significantly

higher than -1, as the dotted line representing a slope of -1 lay outside the 95% confidence intervals of the observed data. The mean UMCs ranged from -1.38 to -1.72, which were substantially higher than the slope obtained by Jeffress et al., (1962), which was -0.48.

No consistent effect of stimulus duration was observed for coherences between 1 and 0.2 (i.e. the range of coherences where the 71% point of the psychometric function fell within the measured range for every datapoint, ($F(2,4) = 2.65$, $p = 0.19$)), and the range of UMCs obtained across stimulus duration and listener was wide (-1.48 to -1.86).

3.2.4 Discussion.

As coherence decreased, thresholds rose so that each time the coherence was halved, thresholds more than doubled. This finding contrasts with the diluting effect on ITD of correlated interfering noise observed in experiment 1, and invalidates the assumption made in the study of Bernstein et al. (2001) that ITD threshold is diluted by uncorrelated interfering noise in a linear 1:1 manner.

Individual UMCs ranged from -1.48 to -1.86, but did not show a trend of decreasing with stimulus duration, as they would if listeners were employing some form of long-term integration. If listeners were using long-term integration, thresholds at lower coherences would decrease, and thresholds at higher coherences would level off at a threshold of approximately 10 μs, the minimum value of ITD that the auditory system can discriminate (e.g. Mossop and Culling, 1998), with a net effect of reducing the UMC. Higher thresholds and steeper UMCs at all three stimuli durations were observed than those found by Jeffress et al. (1962, see Fig. 3.6). This may be partly attributable to the continuous stimuli used by Jeffress et al.; their participants appear

to have performed very long-term integration of a steady-state, partially correlated stimulus. The associated decrease in UMC with continuous stimuli indicates that there might be some form of integration process at work that allows the listener to optimize their performance. It is possible that listeners were employing a multiple looks strategy of the sort proposed by Viemeister and Wakefield (1991). By constantly resampling the correlation within the temporal window and comparing the outputs of temporal windows over time, listeners could be able to improve their performance (i.e. thresholds at lower coherences would decrease, and thresholds at higher coherences would level off at a threshold of approximately 10 μs, see above). Most probably the equipment they used may have been the main factor, as the highest threshold it was theoretically possible to obtain was approximately 280 μs, at a correlation of zero where participants were simply guessing.

The experiments described in this chapter indicate that interfering correlated and uncorrelated noises are integrated within the temporal window in different ways. Correlated interfering noise dilutes the ITD of delayed noise also present in the temporal window in a linear (1:1) manner, whereas the presence of an uncorrelated interferer is more debilitating to the listener. Modelling the temporal window will depend on the type of interferer incorporated in the design of the experiment. The next chapter attempts to model the temporal window using a lateralization task, and describes experiments that involve both correlated and uncorrelated interfering noise.

CHAPTER 4: TEMPORAL ANALYSIS

This chapter examines the limits of the binaural system in resolving temporal fluctuations. The experiments follow a progression that aims to accurately measure the ERD of the binaural temporal window, using a lateralization task developed from the design employed by Bernstein et al. (2001). The experiments are designed to shed light on why that study found temporal windows of a short duration (0.9-1.3 ms equivalent rectangular duration or ERD), a finding that runs contrary to the majority of the binaural temporal window literature (e.g. Grantham and Wightman, 1979; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998; Holube et al., 1998), which indicates ERDs of 40-200 ms. It will be argued that the short durations of the windows modelled in Bernstein et al.'s study result from confounding factors in the design of the experiment.

Bernstein et al. (2001) attempted to measure the binaural temporal window using a detection task in which participants were asked to make lateralization judgements based on stimuli with a total duration of 100, 40, or 20 ms. In their first experiment, they employed a stimulus design in which an ITD was applied to a probe section of noise with a duration of 2, 4, 8, 16, 32 or 64 ms, temporally centred within diotic noise (see Fig. 4.01).

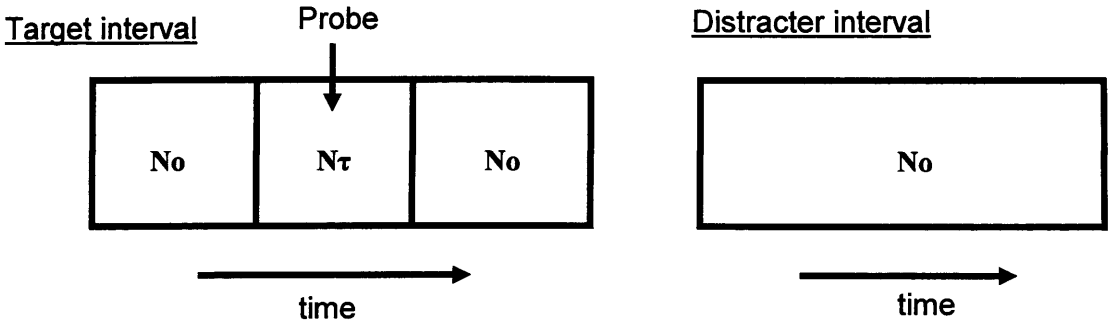

FIG. 4.01. A schematic of the stimuli presented in the detection task used by Bernstein et al. (2001), where diotic interfering noise (No) was presented contiguous to a delayed probe (Nτ). The same stimulus design is used in the detection tasks in experiment 3.

Following the simplifying assumptions made by Akeroyd and Summerfield (1999), Bernstein et al. assumed the window was symmetric, and that the listener detected the ITD imposed on the probe by centring a temporal window at the midpoint of the probe in order for the maximum amount of delayed noise to fall within the window. The window integrates together the ITD of the probe with the zero ITD conveyed by any temporally contiguous diotic noise that also falls within the window. The integration results in an internal, effective ITD that is lower than the external ITD imposed on the probe (Bernstein et al., 2001). As a result, the external ITD must be increased to a magnitude that will bring the internal ITD up to threshold. When the probe duration is long, ITD thresholds will be low as very little diotic noise enters the window, diluting the ITD of the probe. As the probe duration decreases, more diotic noise enters the window and is integrated with the probe ITD, increasing ITD threshold. This is the pattern of data obtained by Bernstein et al. (2001, see Fig. 4.02). According to the dilution assumption, if the listeners were simply integrating the entire stimulus, the slopes of the data at the three probe durations on this log-log plot

would be -1. As the slopes were shallower than -1, the data suggests that the listeners were using a temporal window to resolve the stimulus.

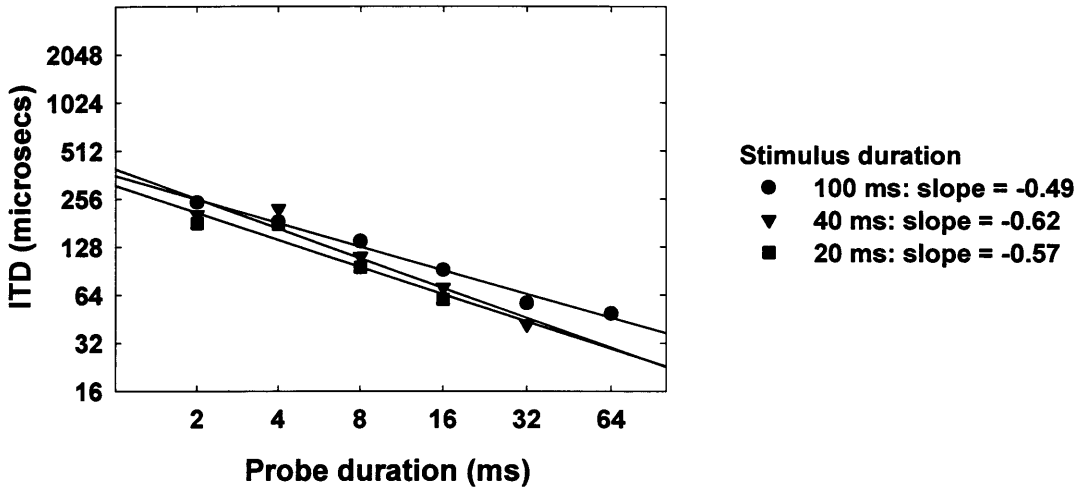

FIG. 4.02. Thresholds for detection of an ITD applied to the probe segment of a noise temporally flanked by interaurally-correlated noise taken from Fig. 1 of Bernstein et al. (2001), but with the x axis plotted logarithmically.

Best fitting temporal windows were found to be described by exponential-skirt functions. These windows were composed of two time constants; one of a short duration (between 0.02 and 0.12 ms) that described the central peak of the window, and a second longer time constant (between 7.48 and 64.21 ms) that described the window skirts. The weighting parameter *w* for the skirt ranged from -13.98 to -16.99 dB. These results, obtained with interfering diotic noise, were very similar to those found by Wagner (1991), who derived temporal window shapes for the barn owl using a similar design but with uncorrelated interfering noise and a discrimination task. The binaural temporal windows fitted to the barn owl data had a peak with a duration of approximately 3-5 ms, and a skirt with a duration of 30-50 ms.

These temporal windows were far shorter than previous estimates such as Culling and Summerfield (1998), who obtained a window 55 to 188 ms wide (see Fig. 4.03).

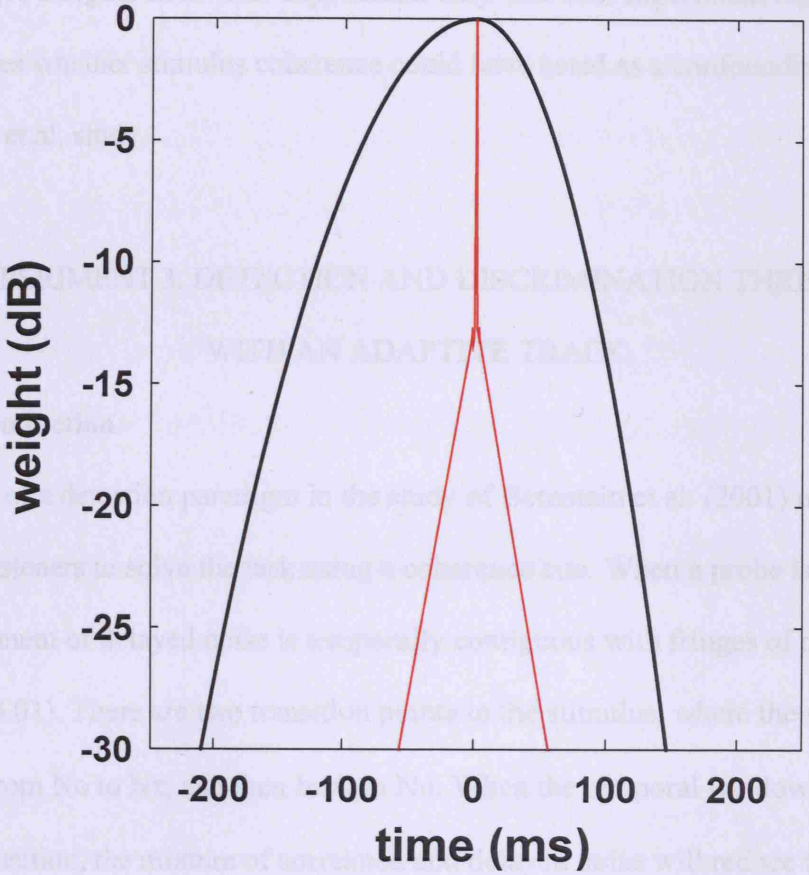

FIG. 4.03. Temporal window functions. The outer black window line plots the simple Gaussian window found by Culling and Summerfield (1998). The inner red line plots the narrow double-exponential window found in the study by Bernstein et al. (2001).

Investigation of the monaural system provided evidence that the monaural temporal window had an ERD of approximately 13 ms (Plack and Moore, 1990), thus the data produced in the study of Bernstein et al. suggests that the binaural system has an even finer temporal resolution than the monaural system. However, it should be noted that ERD can be misleading as a measure of temporal windows, as despite the mean ERD

of the window modelled by Bernstein et al. being approximately 0.48 ms, there is clearly an effect of probe duration up to at least 64 ms (see Fig. 4.02). In addition, the shape of the temporal window that is assumed influences the ERD (e.g. Gaussian fits always have a higher ERD than exponential fits). The first experiment reported here investigates whether stimulus coherence could have acted as a confounding cue in the Bernstein et al. study.

4.1 EXPERIMENT 3. DETECTION AND DISCRIMINATION THRESHOLDS WITH AN ADAPTIVE TRACK.

4.1.1 Introduction.

The use of a detection paradigm in the study of Bernstein et al. (2001) may have allowed listeners to solve the task using a coherence cue. When a probe is present, the probe segment of delayed noise is temporally contiguous with fringes of diotic noise (see Fig. 4.01). There are two transition points in the stimulus, where the noise changes from No to Nτ, and then back to No. When the temporal window straddles such a transition, the mixture of correlated and delayed noise will reduce the coherence of the noise within the window. The presence of the probe in the target interval will thus cause decorrelation, which can distinguish the target interval from the distracter intervals whose coherence is 1. Decorrelation could cause the of image width (Gabriel and Colburn, 1981) or loudness (Culling and Edmonds, 2006) of the target interval to change, allowing listeners to perform the task using width or loudness cues as well as the lateralization cue provided by the ITD of the probe. At the longest stimulus duration (100 ms), containing the shortest duration of probe (2 ms), the interaural coherence of the complete 100-ms stimulus can be calculated and was found to be approximately 0.98 (ITD = 255 μs). This value is close to that found

by Akeroyd and Summerfield (1999) for the just-noticeable decrement in coherence from unity: 0.975 for a 100-500 Hz wide band of noise. The similarity of these values suggests that the target stimuli used in the Bernstein et al. study demonstrated differences from the coherent distracter stimuli that could be detected without using fine temporal resolution. When temporal windows were fitted to the data of Bernstein et al. (2001), detection of changes in correlation was not modelled; the window was calculated by applying a temporal window to the values of the ITD carried by the probe and the interfering surrounding noise only.

In a 2AFC task, the problem of coherence can be addressed by introducing an ITD into both stimuli that the listeners hear (see Fig. 4.04). This controls for the reduction in coherence because it occurs in both intervals, and the listeners' task is no longer to detect a target interval, but to discriminate the contrasting direction of the ITD across both intervals.

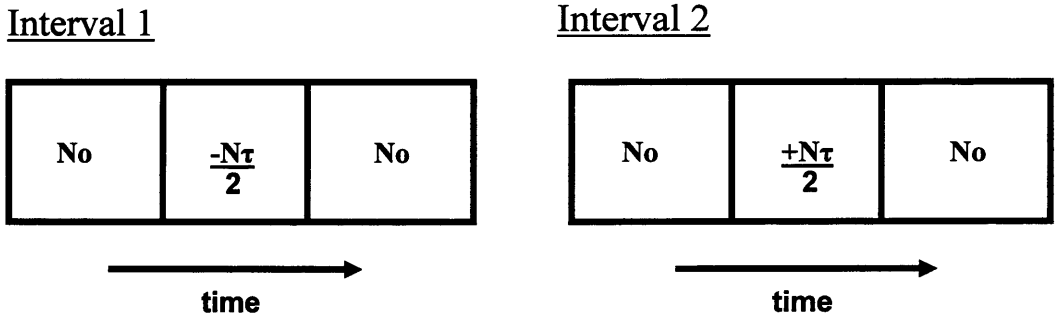

FIG. 4.04. A schematic of the stimuli presented in a discrimination task in experiment 3. The delayed portion of noise presented in the first interval is always in the opposite direction to that in the second interval.

The use of a discrimination task rather than a detection task with stimuli otherwise similar to those of Bernstein et al. (2001) should thus prevent listeners using width or

loudness to solve the task. Previous investigators have found that the temporal window has a duration of approximately 100 ms (e.g. Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998). If listeners were using a 100-ms long temporal window to perform the task, the window would have integrated most or all of the target interval for even the longest of Bernstein et al.'s stimuli. If the temporal window is in fact longer than 100 ms, the slope of the thresholds obtained in the discrimination condition should be approximately -1, as without coherence cues to help identify the target interval, listeners will not be able to use the window to resolve the stimulus and instead use simple temporal integration (e.g. Viemeister and Wakefield, 1991), so that for every halving of the signal duration, threshold doubles. In this event, windows fitted to the data will not be measurable, as listeners will be displaying temporal integration and not resolution.

Experiment 3 consisted of three conditions: a 4-interval detection task, a 2-interval detection task, and a 2-interval discrimination task. The 4-interval detection task was included as a direct replication of Bernstein et al. (2001). The 2-interval detection task was included to ensure that no significant differences occurred between the 2 and 4-interval detection tasks, thus allowing the 4-interval detection task to be directly compared with the 2-interval discrimination condition. Thresholds in both detection conditions were contrasted with discrimination thresholds. It was hypothesised that the slope of the thresholds in the discrimination condition would not be significantly different from -1, that the slopes of the data in the detection condition would be significantly less than -1, and that there would be no significant difference in performance using 4 or 2-interval detection tasks.

4.1.2 Method.

Participants

Four participants took part in the experiment. Two were male and two female aged between 18 and 35. Two of them (AK and SW) had taken part in experiments 1 and 2, and another (JC) was experienced in psychophysical experiments. The untrained participant (SP) received at least 7 hours of training before data collection. Participants were paid upon completion.

Apparatus/Materials

In the detection conditions, listeners were presented with stimulus intervals consisting of a single target interval and either 1 or 3 distracter intervals (2I and 4I-2AFC tasks). The total duration of each interval was 100 ms, with 10-ms gated onset and offset. The target interval consisted of a segment of noise with a probe ITD temporally centred within the 100-ms stimulus (see Fig. 4.01). The probe and surrounding diotic noise were temporally contiguous. Distracter intervals were composed entirely of diotic noise. The inter-stimulus interval was 500 ms. The discrimination condition consisted of a 2I-2AFC task. Listeners were presented with two target intervals both of which contained a temporally centred probe ITD, so that the first interval was lateralized to one side of the head, and the second lateralized to the other side of the head (see Fig. 4.04). The adapted variable was the difference in ITD between the two intervals. The duration of the probe was 2, 4, 8, 16, 32 or 64 ms.

Design

The experiment had a repeated measures design. The independent variables were probe duration (six levels) and condition (three levels). The six probe durations were

64, 32, 16, 8, 4 and 2 ms, and the three conditions were detect 4 (4-interval detection), detect 2 (2-interval detection) and discrimination (2-interval discrimination). The order in which participants performed the three conditions was randomised. The dependent variable was threshold ITD measured in microseconds.

Procedure

Listeners participated in each of the three conditions. In the detect 4 condition, the task was to identify the target interval among the three reference intervals. The target interval was presented randomly in either the second or third interval, and feedback was provided. The target ITD was varied adaptively in order to obtain a 70.7% correct estimate (Levitt, 1971). Each adaptive track started with the ITD set to 500 μs. Initially, the step size of the adaptive track corresponded to a factor of 0.2, and was reduced to a factor of 0.05 following two reversals. Following these two reversals, the track was terminated after 10 reversals. The last 10 reversals comprised the measurement phase, and the average of the reversals within the measurement phase was taken as threshold. Listeners responded using the keyboard (if the second interval contained the target, the participant pressed 1; if it occurred in the third interval, they pressed 2). The detect 2 condition was the same as the detect 4 condition, except that the first and last reference intervals were removed. In the discrimination condition, the listener's task was to identify which direction the two intervals moved in (left to right or right to left), corresponding to the ITDs embedded within each interval. Participants pressed 1 if the noise moved from right to left, and 2 if it moved left to right. The direction of sound movement was randomised, ITDs were varied adaptively and feedback was provided.

4.1.3 Results.

The results for the four participants are shown in Fig. 4.05. Each plotted threshold is the average derived from three experimental runs. All error bars represent ±1 standard error. In all three conditions, participants tended to produce the lowest thresholds for the longest probe durations, with thresholds increasing as probe duration decreased. For all four participants, thresholds for the discrimination condition were higher than for the detection conditions, especially at the short probe durations.

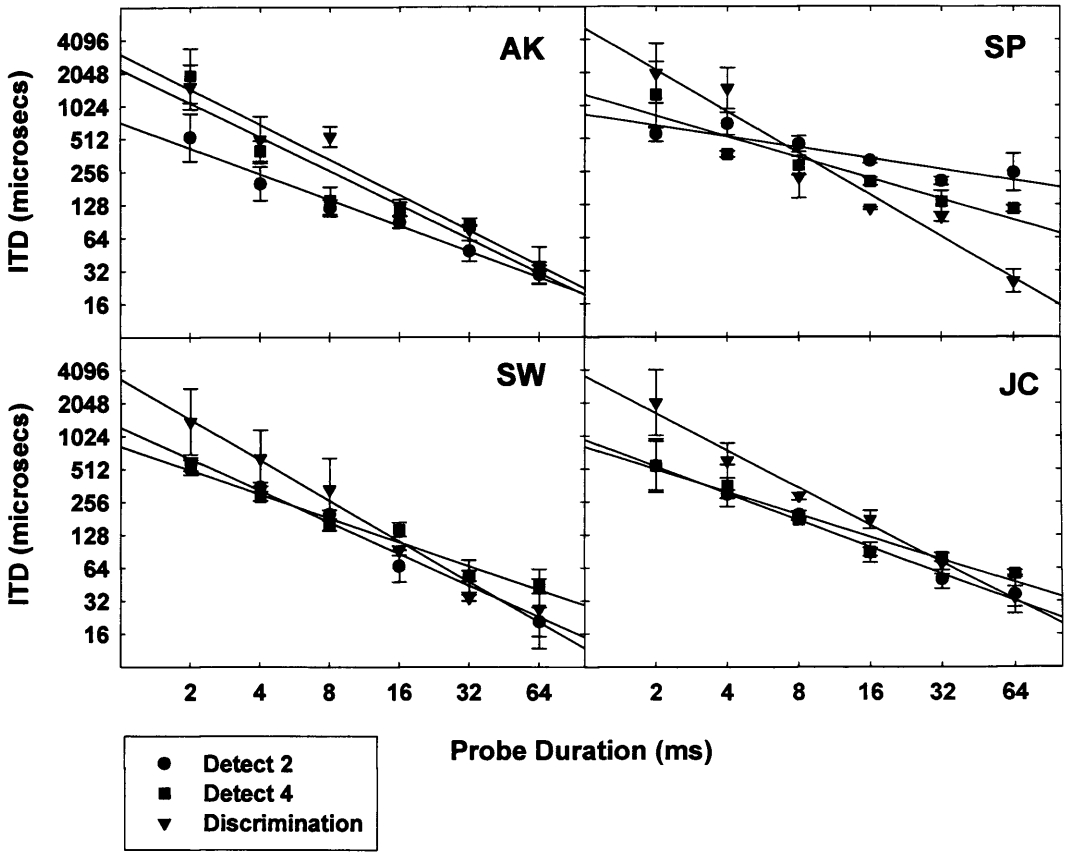

FIG. 4.05. Individual results for experiment 3 (Adaptive Track). Regression lines are plotted for the three conditions for each listener. Both axes are plotted logarithmically.

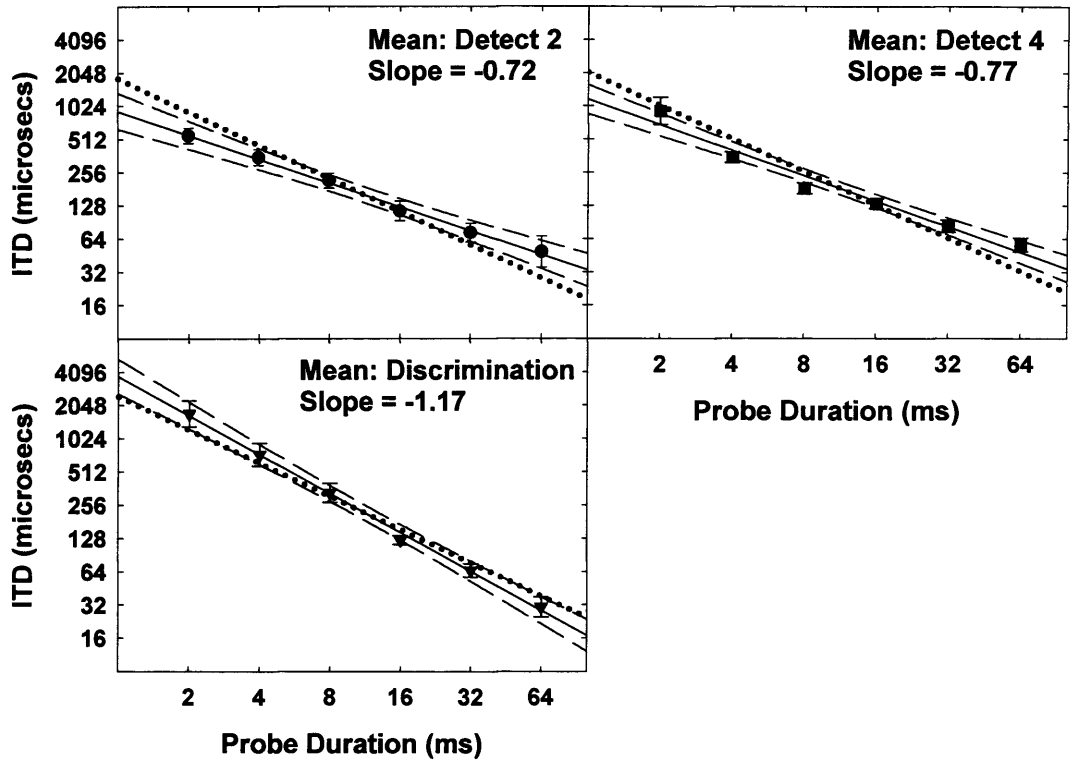

FIG. 4.06. Mean results for experiment 3 (Adaptive Track). Solid regression lines are plotted for each condition. Dotted lines represent a slope of -1, and dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

Fig. 4.06 shows that the slope of the mean data in the detect 2 condition was –0.72, $r^2 = 0.97$, in the detect 4 condition the slope was –0.77, $r^2 = 0.96$, and in the discrimination condition the slope was –1.17, $r^2 = 0.99$. The 95% confidence intervals of the regression encompassed a slope of -1 (dotted line in fig 4.06) in the discrimination condition, but not in the two detection conditions, where slopes of the data were significantly shallower than -1.

	Probe duration (ms)	Mean (ms)	Mean 95% Confidence intervals		Participant	Slope
			lower bound (ms)	upper bound (ms)		
Detect 2	64	48.83	32.16	65.67	AK	-0.78
	32	72.78	60.98	94.00	SP	-0.33
	16	114.13	99.65	156.11	SW	-0.96
	8	214.78	173.29	243.65	JC	-0.81
	4	349.68	282.18	406.09	Mean	-0.72
	2	552.61	468.44	663.93		
Detect 4	64	56.60	41.42	56.82	AK	-1.07
	32	83.56	71.29	95.52	SP	-0.63
	16	133.45	125.06	157.59	SW	-0.73
	8	183.69	210.49	270.96	JC	-0.68
	4	351.70	358.96	459.79	Mean	-0.77
	2	924.17	505.96	943.98		
Discrimination	64	30.72	22.62	36.05	AK	-1.03
	32	65.43	55.12	74.94	SP	-1.26
	16	123.59	130.79	160.00	SW	-1.23
	8	330.75	261.55	405.32	JC	-1.13
	4	736.21	564.04	952.10	Mean	-1.17
	2	1727.49	1222.90	2224.56		

Table 4.01. Mean thresholds, 95% confidence intervals from the regression line, and slopes for experiment 3 (Adaptive Track).

Table 4.01 shows a summary of individual and mean slopes from Figs 4.05 and 4.06, and shows mean thresholds at each probe duration and numerical values for the 95% confidence intervals plotted in Fig. 4.06. A within-subjects ANOVA demonstrated a significant main effect of condition ($F(2,6)= 12.48$, $p<0.05$), a significant main effect of probe duration ($F(5,15)= 64.88$, $p<0.001$), and a significant interaction between condition and probe duration ($F(10,30)= 4.88$, $p<0.001$). Pairwise comparisons run with a Bonferroni correction demonstrated a significant difference between thresholds in the detect 2 and discrimination conditions at probe durations of 4 ms ($p = 0.03$) and 2 ms ($p = 0.01$). No significant differences were observed between the detect 2 and detect 4 conditions at any probe duration, or between the detect 4 and discrimination conditions.

A temporal window was modelled to the dataset provided by each participant in each condition. In addition to modelling the data with a CMC (Correlated Masking

Coefficient) of -1 consistent with the dilution assumption, individual CMC's can be modelled to the data, provided the listener also took part in experiment 1, as did AK and SW (N.B. the χ^2 value does not include the CMC (or, when an uncorrelated interferer is applied the UMC) as a free parameter). Temporal windows were fitted to the data using simple and skirt Gaussian, exponential and rounded-exponential fits (Tables 4.02 to 4.07). Time constants were defined as indeterminate if they were larger than 1000 ms. The cutoff of 1000 ms was chosen as the length of the stimulus was considerably shorter (i.e. 100 ms), thus time constants longer than the stimulus duration could not be measured with accuracy.

Listener	Fitting function (simple functions)								
	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	20.94	41.89	0.0361	9.96	35.38	0.0440	18.46	37.00	0.0394
JC	20.43	40.86	0.0129	9.45	33.49	0.0088	17.97	35.97	0.0097
SP	2.62	5.21	0.1169	1.72	6.09	0.1574	2.92	5.84	0.1378
SW	107.49	215.04	0.0291	30.44	107.92	0.0312	66.37	133.78	0.0308

Listener	Fitting function (skirt functions)														
	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)
AK	9.60	**1000+**	-5.06	**1000+**	0.0462	3.11	400.92	-3.56	673.06	0.0399	6.76	776.21	-3.94	635.30	0.0418
JC	20.40	20.56	-8.11	40.85	0.0172	9.45	9.54	-10.76	33.50	0.0117	17.96	18.22	-11.05	35.96	0.0130
SP	0.07	8.60	-16.79	0.49	0.0643	0.14	6.13	-12.92	1.57	0.0516	0.12	10.33	-15.29	0.92	0.0579
SW	0.06	**1000+**	-7.79	**1000+**	0.0290	0.15	**1000+**	-3.13	**1000+**	0.0290	0.11	**1000+**	-5.37	808.41	0.0290

Table 4.02. Window fitting parameters for the detect 2 condition in experiment 3 (Adaptive Track). The CMC was set to -1 for all fits, in accord with the dilution assumption. *tp* is the time constant for the peak of the window, *ts* is the time constant of the skirt of the window, *w* is the weighting applied to the skirt of windows that have a skirt fitting function, and the ERD is the integral of the entire window. The goodness of fit (χ^2) is expressed for each individual window fit, where lower values of χ^2 indicate closer fits to the empirical data. Best fits are highlighted in yellow. Parameters are reported to 2 d.p., χ^2 values to 4 d.p.

Table 4.02 shows the results for windows fitted with the CMC parameter set to -1 for the detect 2 condition. Individual differences are large, with best fits and shapes of the window varying across listeners. However, the pattern of data is inconsistent; in several cases the skirt parameter grew beyond any reasonable value. The best fit to listener AK's data was a simple exponential, for JC a simple Gaussian, for SP a skirt Gaussian, and for SW a skirt fit of indeterminate shape that was essentially rectagular. ERDs of the best fitting windows ranged from 1.57 ms (SP) to an indeterminate duration (SW).

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	9.98	19.92	0.0399	5.58	19.78	0.0756	10.02	20.04	0.0570
SW	47.03	94.06	0.0361	17.82	63.20	0.0436	35.87	71.73	0.0416

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)
AK	4.16	**1000+**	-7.23	645.00	0.0296	2.10	**1000+**	-6.08	**1000+**	0.0189	4.02	**1000+**	-6.44	642.12	0.0230
SW	0.11	351.69	-7.14	136.17	0.0285	0.18	693.71	-4.23	929.63	0.0288	0.02	159.98	-13.55	11.70	0.0287

Table 4.03. Window fitting parameters for the detect 2 condition in experiment 3 as for Table 4.02, but all fits were modelled with the CMC parameter set according to the data obtained for each individual participant from experiment 1 (-1.24 for AK, and -1.06 for SW).

Table 4.03 shows the results for windows fitted with individual CMC parameters for the detect 2 condition. As with window fits to the data with the CMC set to -1, individual differences are again large, with best fits and shapes varying between the two listeners. The best fit to listener AK's data was a skirt Gaussian, and for SW a skirt exponential. ERDs of the best fitting windows ranged from 136.17 ms (SW) to an indeterminate duration (AK).

Listener	Fitting function (simple functions)								
	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	**1000+**	**1000+**	0.3248	47.78	170.04	0.3246	126.86	253.73	0.3246
JC	13.07	26.18	0.0269	6.77	23.99	0.0220	12.53	25.06	0.0217
SP	10.29	20.59	0.1535	5.57	19.76	0.1469	10.18	20.35	0.1452
SW	13.32	26.64	0.0507	7.65	27.05	0.0775	13.32	26.75	0.0666

Listener	Fitting function (skirt functions)														
	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)
AK	**1000+**	**1000+**	-0.04	**1000+**	0.4331	36.00	192.95	-3.59	370.96	0.4328	59.00	301.97	-1.13	484.65	0.4328
JC	13.12	12.86	-10.07	26.18	0.0358	5.15	644.91	-8.37	347.97	0.0250	11.44	889.28	-12.81	114.88	0.0287
SP	10.26	10.56	-10.39	20.56	0.2046	3.75	13.87	-6.20	23.55	0.1903	8.84	148.17	-12.04	34.91	0.1928
SW	1.74	63.87	-4.41	48.49	0.0376	1.26	21.08	-4.02	32.42	0.0330	2.10	47.17	-4.20	38.67	0.0352

Table 4.04. Same as Table 4.02, but for the detect 4 condition in experiment 3 (Adaptive Track).

Table 4.04 shows window fits with the CMC parameter set to -1 for the detect 4 condition. The 'best' fit to listener AK's data was either a simple Gaussian or simple roex shape, for JC a simple roex, for SP a simple roex, and for SW a skirt Gaussian. However, the goodness-of-fit for listener AK is somewhat misleading, as the χ^2 values for the three simple fits only differ in the fourth decimal place, a level of accuracy higher than that allowed for by the data. ERDs of the best-fitting windows ranged from 20.35 (SP) to 253.73 ms (AK).

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	23.98	47.78	0.2459	10.31	36.55	0.2486	19.55	39.09	0.2418
SW	11.11	22.23	0.0592	6.54	23.23	0.0961	11.37	22.73	0.0803

	Fitting function (skirt functions)															
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential					Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)		of-fit (χ^2)
AK	19.91	206.68	-9.00	86.88	0.3278	4.05	916.51	-3.76	1000+	0.3003	9.04	564.05	-4.23	436.83		0.3099
SW	1.61	55.22	-5.06	37.07	0.0368	1.19	19.74	-4.62	26.76	0.0322	1.96	43.19	-4.83	31.22		0.0343

Table 4.05. Window fitting parameters for the detect 4 condition in experiment 3. As for Table 4.02, but all fits were modelled with the CMC parameter set according to the data obtained for each individual participant from experiment 1.

For the detect 4 condition with windows fitted with individual CMC parameters (see Table 4.05), the best fit to listener AK's data was given by a simple roex, and for SW a skirt Gaussian. ERDs of the best-fitting windows ranged from 26.73 (SW) to 39.09 ms (AK).

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	1000+	1000+	0.1168	1000+	1000+	0.1168	1000+	1000+	0.1168
JC	1000+	1000+	0.2155	1000+	1000+	0.2155	1000+	1000+	0.2155
SP	1000+	1000+	0.4251	1000+	1000+	0.4251	1000+	1000+	0.4251
SW	1000+	1000+	0.4326	1000+	1000+	0.4326	1000+	1000+	0.4326

	Fitting function (skirt functions)															
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential					Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)		of-fit (χ^2)
AK	1000+	1000+	-0.04	1000+	0.1557	1000+	1000+	-0.04	1000+	0.1557	1000+	1000+	-0.04	1000+		0.1557
JC	1000+	1000+	-0.04	1000+	0.2874	1000+	1000+	-0.04	1000+	0.2874	1000+	1000+	-0.04	1000+		0.2874
SP	1000+	1000+	-0.04	1000+	0.5668	1000+	1000+	-0.04	1000+	0.5668	1000+	1000+	-0.04	1000+		0.5668
SW	1000+	1000+	-0.04	1000+	0.5768	1000+	1000+	-0.04	1000+	0.5768	1000+	1000+	-0.04	1000+		0.5768

Table 4.06. Same as Table 4.02, but for the discrimination condition in experiment 3 (Adaptive Track).

The ERD of the window obtained from the discrimination task with the CMC parameter set to -1 (see Table 4.06) was large and indeterminate for each listener for all fits. All simple fits had an equal χ^2 value, and all skirt fits had an equal χ^2 value, although χ^2 values were lower for simple fits. We believe that the equal χ^2 values are related to the large time constants that describe both the peak and (for the skirt fits) the skirt values, and gives little room for the base function to have an influence. The shape of the window with the best fit to the data is inconsistent across listeners, condition, and dependent on the setting of the CMC parameter (individual or dilution: see Table 4.07). Essentially, all fits are rectangular.

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)	tp(ms)	ERD(ms)	of-fit (χ^2)
AK	39.65	78.99	0.0859	14.36	50.91	0.0870	29.25	58.73	0.0867
SW	**1000+**	**1000+**	0.3760	33.09	117.31	0.3745	123.42	246.84	0.3754

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD(ms)	of-fit (χ^2)
AK	39.08	75.73	-18.26	80.27	0.1145	7.19	259.47	-2.95	510.09	0.1149	17.51	344.51	-4.01	299.56	0.1151
SW	**1000+**	**1000+**	-0.04	**1000+**	0.5013	32.23	148.52	-18.21	119.73	0.4993	163.08	121.43	-0.31	232.07	0.5005

Table 4.07. Window fitting parameters for the discrimination condition in experiment 3. Same as for Table 4.02, but all fits were modelled with the CMC parameter set according to the data obtained for each individual participant from experiment 1.

Table 4.07 shows windows fitted to the discrimination data with individual CMC parameters. The best fit to listener AK's data was given by a simple exponential, and for SW a simple Gaussian. ERDs of the best-fitting windows ranged from 78.99 (AK) to 117.31 ms (SW).

Both participants who participated in experiment 1 had individual CMC's steeper than -1 (AK: -1.24, SW: -1.06). For both listeners, windows fitted with individual CMCs resulted in a decrease in the ERD of the temporal window; the steeper the CMC, the lower the ERD. In the detect 4 condition (compare Tables 4.04 and 4.05), with a dilution CMC of -1, the ERD of the best fits to AK's data is 170.04 (Gaussian) and 253.73 ms (roex). This reduces to 39.09 ms (roex) when the CMC is set to -1.24. Compared to AK, the smaller disparity between individual (-1.06) and dilution CMC in the case of SW results in a lower magnitude in the reduction of ERD when windows are fitted using the individual CMC parameter. With a CMC of -1, the best fit ERD is 32.42 ms. This reduces to 26.76 ms when the CMC is set to -1.06.

A summary of the best fits to the data is shown in Table 4.08.

Listener	Task	CMC	Best fit	tp (ms)	ts (ms)	w (dB)	ERD (ms)	Goodness-of-fit (χ^2)
AK	Detect 2	-1	simple exponential	20.94	N/A	N/A	41.89	0.0361
		-1.24	skirt Gaussian	2.1	**1000+**	-6.08	**1000+**	0.0189
	Detect 4	-1	simple Gaussian	47.78	N/A	N/A	170.04	0.3246
		-1.24	simple roex	19.55	N/A	N/A	39.09	0.2418
	Discrimination	-1	simple indeterminate	**1000+**	N/A	N/A	**1000+**	0.1168
		-1.24	simple exponential	39.65	N/A	N/A	78.99	0.0859
SW	Detect 2	-1	skirt indeterminate	0.15	**1000+**	-3.13	**1000+**	0.0290
		-1.06	skirt exponential	0.11	351.69	-7.14	136.17	0.0285
	Detect 4	-1	skirt Gaussian	1.26	21.08	-4.02	32.42	0.0330
		-1.06	skirt Gaussian	1.19	19.74	-4.62	26.76	0.0322
	Discrimination	-1	simple indeterminate	**1000+**	N/A	N/A	**1000+**	0.4326
		-1.06	simple Gaussian	33.09	N/A	N/A	117.31	0.3745
JC	Detect 2	-1	simple Gaussian	9.45	N/A	N/A	33.49	0.0088
	Detect 4	-1	simple roex	12.53	N/A	N/A	25.06	0.0217
	Discrimination	-1	simple indeterminate	**1000+**	N/A	N/A	**1000+**	0.2155
SP	Detect 2	-1	skirt Gaussian	0.14	6.13	-12.92	1.57	0.0516
	Detect 4	-1	simple roex	10.18	N/A	N/A	20.35	0.1452
	Discrimination	-1	simple indeterminate	**1000+**	N/A	N/A	**1000+**	0.4251

Table 4.08. Best fits to the detect 2, detect 4 and discrimination data for all listeners. For a description of the parameters, see Table 4.02.

4.1.4 Discussion.

Detection and discrimination thresholds were found to be significantly different, especially at short probe durations. Threshold slope for the discrimination condition

was significantly steeper than detection slopes, supporting the hypothesis that there was a coherence cue present in the detection tasks that listeners were able to use to increase their performance, especially for low probe durations (see Figs. 4.05 and 4.06). Fig 4.06 indicates that when coherence was controlled for by employing a discrimination task, the mean steepness of the slope increased to -1.17 from -0.72 (detect 2) and -0.77 (detect 4). The slopes of the data for the detection conditions were significantly shallower than -1, as the plotted hypothetical slope of -1 was steeper and fell outside the 95% confidence intervals in both conditions (see Fig. 4.06). The slope of the discrimination data was not significantly different from -1, which suggests that when the coherence cue was removed by employing a discrimination task, participants act in a manner consistent with integration of the entire stimulus instead of using a temporal window. This is supported by the results that were obtained when attempts were made to fit temporal windows to the discrimination dataset; all windows were indeterminate and effectively rectangular.

No consistent best-fitting window shape was evident for each task, but from a skirt Gaussian most often provides the best fit to the detect 2 data, a simple roex to the detect 4 data, and a simple fit of indeterminate shape to the discrimination data (see Table 4.08). However, due to the occurrence of rectangular fits and parameter inflation beyond 1000 ms throughout the dataset, no conclusions can be drawn about the shape of the best fitting window at this point in the study. When windows were fitted to the discrimination data, the fitting algorithm increased the duration parameters indefinitely; the data did not constrain the fitted window size. A possible explanation was that the window was as long as or longer than the stimulus duration and the data consequently reflected only the complete integration of all the stimulus. If the binaural temporal window is approximately 100 ms long (e.g. Culling and

Summerfield, 1998; Akeroyd and Summerfield, 1999), a stimulus longer than 100 ms is required so that its duration extends for a time period that is great enough to encompass the entire window.

However, the thresholds for the detect 4 condition do not replicate those found using identical stimuli by Bernstein et al. (2001); the thresholds obtained in the current experiment are far higher than previous comparable results e.g. at a probe duration of 2 ms, the mean ITD of the four listeners was 924.17 μs (see Fig. 4.06 and Table 4.01), whereas Bernstein et al. obtained a value of approximately 255 μs with a stimulus duration of 100 ms (see Fig 4.02). Of the four listeners, only SW's data was better fit by a skirt function like that favoured by Bernstein et al. (for both an individual CMC and a CMC of -1; see Tables 4.04 and 4.05), but the durations of both the peak and the skirt were longer than those obtained by Bernstein et al.'s listeners. Our experience of replicating the experiment was that performance was rather unstable. An adaptive track methodology was used to gather thresholds in order to match the methodology used previously by Bernstein et al. However, when employed in ITD lateralization tasks, the adaptive track method has several problems associated with it that were highlighted in the discussion of experiment 1. The use of ITD (at least for delays lower than approximately 700 μs) as the dependent variable in an adaptive track violates assumption #1 of Levitt (1971)[1], and the step sizes used in the adaptive track may have caused the increase in thresholds. The step size was 0.2, which was lowered to 0.05 after two reversals. The large magnitude of the initial step size made the results strongly dependent on the listener's initial performance before the first two reversals. Due to the difficulty of the task at the shortest probe durations,

[1] '...adaptive psychophysical procedures should *never* be used when it is known that changes in performance are *not* monotonically related to changes in the independent variable' (Trahiotis, Bernstein, Buell and Spektor, 1990).

participants were more likely to make an error early in the track, reducing the step size and making it harder to obtain a low threshold. An example of this is the results obtained for participant AK in the detect 4 condition, where the steep slope and indeterminate window stem from a high threshold at a probe duration of 2 ms (see Fig. 4.05).

An alternative method of gathering data to produce temporal windows is to measure the psychometric functions of participants by producing stimuli with a range of fixed ITDs and measuring the percentage of correct responses at each ITD. This method is preferable to the adaptive track as it accounts for the instability of performance at short probe durations. By taking the 75% point of the psychometric function of each probe duration as threshold, a temporal window can be constructed. 75% was chosen as the slope of the psychometric function is steepest at this point, and the thresholds did not have to be compared to adaptive track data (which would require the 71% point to be taken as threshold in order to allow equivalent comparison for a percentage correct estimate).

4.2 EXPERIMENT 4. PSYCHOMETRIC FUNCTIONS.

4.2.1 Introduction.

Experiment 4 again contrasted performance in detection and discrimination tasks, but thresholds were gathered by obtaining six-point psychometric functions. As the results of experiment 3 indicated there was no significant difference between 2 and 4 interval detection tasks, the 4-interval detection task was removed from the procedure, thus the fourth experiment obtained psychometric functions for 2-interval detection and 2-interval discrimination tasks only.

4.2.2 Method.

Participants

Three participants, two of whom had participated in experiment 1, 2 and 3 (AK and SW), took part in the present experiment. The new participant (HM) had participated in experiment 1. Two were female and one was male. The subjects were between 18 and 25 years of age, and were paid for participating in the experiment.

Apparatus/Materials

The experiment consisted of 2I-2AFC detection and discrimination tasks, and used stimuli with the same design as those used in the previous experiment. The total duration of each interval was 100 ms, with 10-ms gated onset and offset. The inter-stimulus interval was 500 ms. The duration of the probe noise was 2, 4, 8, 16, 32 or 64 ms. Instead of varying the ITD adaptively, the probe was presented with a delay of 1024, 512, 256, 128, 64 or 32 microseconds.

Design

The experiment manipulated three independent variables within-subjects: probe duration (64, 32, 16, 8, 4 and 2 ms), ITD (1024, 512, 256, 128, 64, 32 microseconds), and condition (detection and discrimination). ITDs were presented randomly. The dependent variable was percentage of correct responses.

Procedure

In this experiment, participants performed a 2I-2AFC detection task, and a 2I-2AFC discrimination task. Participants responded using the keyboard as in the first experiment, and the same procedure was followed for both tasks.

Initially, a block of trials was presented with probe durations of 64 ms, and the order in which the six probe ITDs were presented within the block was randomised. Blocks of trials with probe durations of 32, 16, 8, 4 and 2 ms followed. The procedure was repeated so that 20 trials were repeated for each ITD and probe duration. Each run therefore consisted of 6 ITDs x 6 probe durations x 20 repetitions = 720 trials in all.

4.2.3 Results.

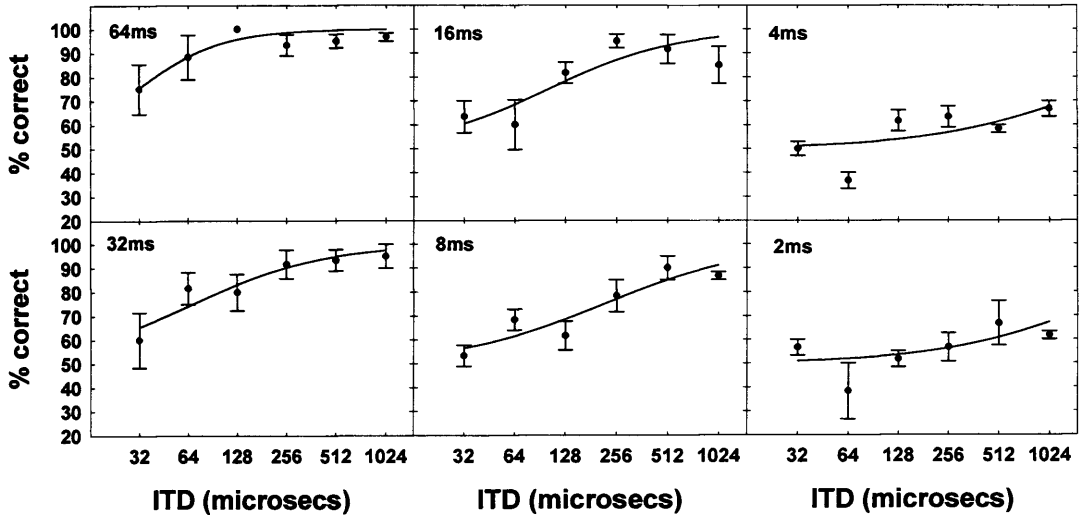

FIG. 4.07. Psychometric functions taken from listener HM performing the detection task. The different panels indicate results at different probe durations. The symbols show the observed threshold data. The lines show the closest match to this data achieved by scaling the predicted data to the empirically averaged data. The x axis is plotted logarithmically. Error bars represent 1 standard error.

Fig. 4.07 shows psychometric functions plotted at each probe duration for listener HM. At short probe durations a shallower psychometric function was observed as performance even at the highest ITDs did not reach threshold.

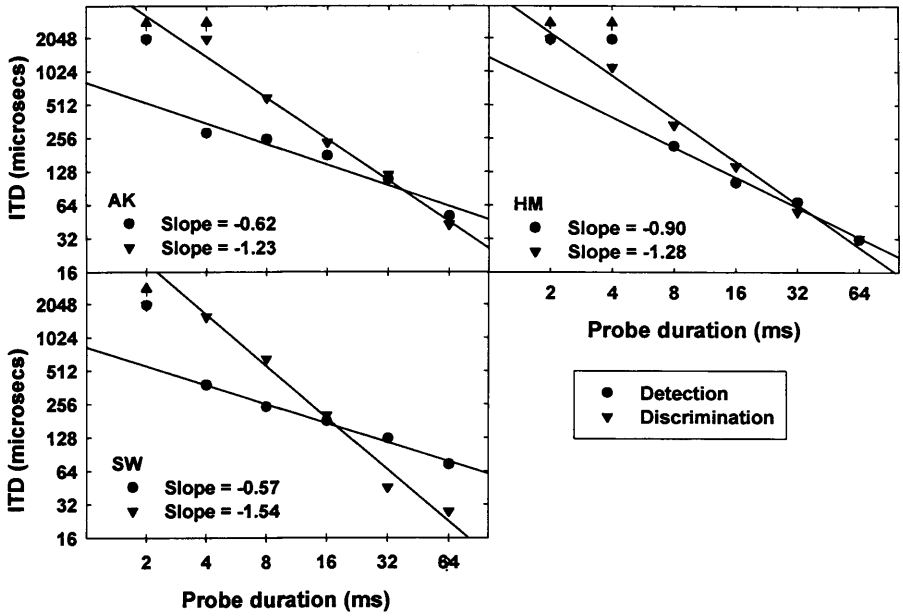

FIG. 4.08. Individual results for experiment 4 (Psychometric Functions). Regression lines are plotted for the two conditions for each listener. Arrows above data points arbitrarily plotted at a value of 2 ms indicate that threshold fell outside the measured range of the psychometric function. Both axes are plotted logarithmically.

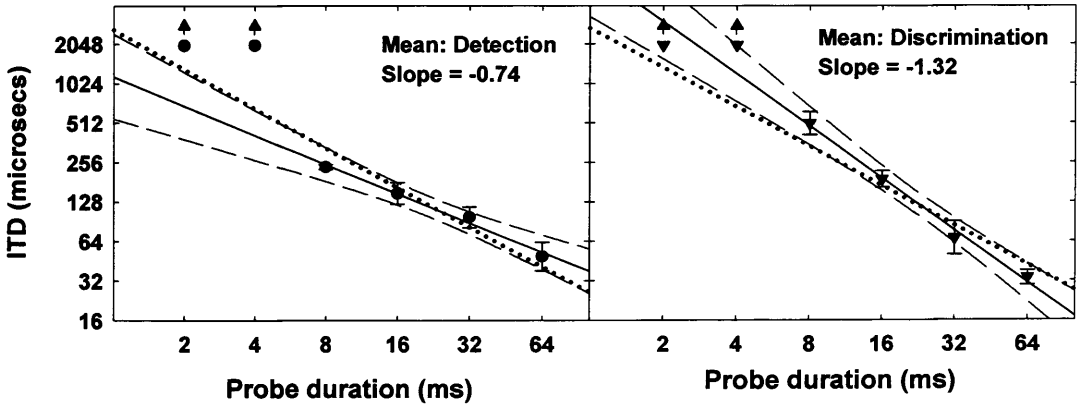

FIG. 4.09. Mean results for experiment 4 (Psychometric Functions). Solid regression lines are plotted for both conditions. Dotted lines represent a slope of -1, dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

The results for the three participants are shown in Fig. 4.08. Each plotted threshold is derived by fitting a psychometric function to the average of the listener's last three experimental runs. None of the participants were able to reach a measurable threshold at the shortest probe duration (2 ms). Participant AK was unable to produce measurable thresholds for the discrimination task at 4 ms probe duration, and HM was not able to produce measurable thresholds at a probe duration of 4 ms in the detection task (a surprising result, as she was capable of producing a measurable threshold at 4 ms probe duration in the discrimination task which often yielded higher thresholds). The mean slope of the detection data was - 0.74, $r^2 = 0.99$, and the slope of the discrimination data was - 1.32, $r^2 = 0.99$ (see Fig. 4.09). A within-subjects ANOVA demonstrated a significant interaction between condition and probe duration ($F(3,6) = 9.05$, $p<0.05$). Pairwise comparisons run with a Bonferroni correction demonstrated a significant difference between thresholds between the two conditions at a probe duration of 16 ms. The difference between thresholds at a probe duration of 8 ms approached but did not reach significance ($p=0.08$). Thresholds predicted by a slope of -1 were not significantly different to those predicted by the empirical data, as the slope of -1 was encompassed by the confidence intervals in both the detection and discrimination conditions (see Fig. 4.09).

Temporal windows were fitted to the data in the same format as in the previous experiment.

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)
AK	13.24	26.48	0.0972	8.42	29.86	0.1376	13.95	27.89	0.1207
HM	71.70	143.95	0.0103	23.01	81.55	0.0125	48.91	97.43	0.0118
SW	9.47	18.95	0.0321	5.60	19.84	0.0603	9.98	19.92	0.0465

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)
AK	0.04	**1000+**	-20.66	**1000+**	0.0174	0.02	**1000+**	-22.49	110.46	0.0174	0.23	**1000+**	-13.44	**1000+**	0.0174
HM	4.71	**1000+**	-1.73	**1000+**	0.0245	3.17	23.49	-0.04	**1000+**	0.0222	5.82	**1000+**	-1.46	**1000+**	0.0236
SW	1.87	**1000+**	-7.90	**1000+**	0.0012	1.43	**1000+**	-6.57	**1000+**	0.0004	2.34	**1000+**	-7.00	**1000+**	0.0008

Table 4.09. Same as Table 4.02, but for the detection condition in experiment 4 (Psychometric Functions). The CMC was set to -1 for all fits, in accord with the dilution assumption.

Table 4.09 shows the results for windows fitted with the CMC parameter set to -1 for the detection condition. Individual differences are large, with best fits and shapes of the window varying across listeners. The best fit to listener AK's data was a simple roex, for HM a simple exponential, and for SW a Gaussian skirt fit. ERDs of the best fitting windows ranged from 27.89 ms (AK) to an indeterminate duration (SW).

	Fitting function (simple functions)								
Listener	Exponential		Goodness-of-fit (χ^2)	Gaussian		Goodness-of-fit (χ^2)	Rounded Exponential		Goodness-of-fit (χ^2)
	tp(ms)	ERD		tp(ms)	ERD		tp(ms)	ERD	
AK	7.64	15.22	0.1381	4.51	15.98	0.2030	8.03	16.05	0.1727
HM	79.68	159.38	0.0101	24.43	86.64	0.0119	52.61	105.22	0.0114
SW	8.34	16.68	0.0369	4.95	17.54	0.0681	8.87	17.67	0.0529

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-of-fit (χ^2)	Gaussian				Goodness-of-fit (χ^2)	Rounded Exponential				Goodness-of-fit (χ^2)
	tp(ms)	ts(ms)	w(dB)	ERD		tp(ms)	ts(ms)	w(dB)	ERD		tp(ms)	ts(ms)	w(dB)	ERD	
AK	0.06	**1000+**	-21.62	**1000+**	0.0080	0.04	**1000+**	-20.28	265.39	0.0080	0.04	**1000+**	-23.61	154.24	0.0080
HM	4.71	**1000+**	-1.58	**1000+**	0.0247	3.63	25.05	-0.04	**1000+**	0.0226	5.85	**1000+**	-1.33	**1000+**	0.0238
SW	1.82	**1000+**	-8.52	**1000+**	0.0009	1.40	**1000+**	-7.13	**1000+**	0.0002	2.28	**1000+**	-7.58	**1000+**	0.0005

Table 4.10. Same as Table 4.02, but for the detection condition in experiment 4 (Psychometric Functions). All fits were modelled with the CMC parameter set according to the data obtained for each individual participant from experiment 1 (-1.24 for AK, -0.99 for HM, and -1.06 for SW).

Table 4.10 shows the results for windows fitted with individual CMC parameters for the detection condition. As with window fits to the data with the CMC set to -1, individual differences are again large, with best fits and shapes varying between listeners. The best fit to listener AK's data was a skirt fit of indeterminate shape, for HM a simple exponential, and for SW a Gaussian skirt. ERDs of the best fitting windows ranged from 154.24 ms (AK) to an indeterminate duration (AK and SW).

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)
AK	**1000+**	**1000+**	0.0498	**1000+**	**1000+**	0.0498	**1000+**	**1000+**	0.0498
HM	**1000+**	**1000+**	0.1246	**1000+**	**1000+**	0.1246	**1000+**	**1000+**	0.1246
SW	**1000+**	**1000+**	0.4111	**1000+**	**1000+**	0.4111	**1000+**	**1000+**	0.4111

	Fitting function (skirt functions)															
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-	
	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	
AK	**1000+**	**1000+**	-0.04	**1000+**	0.1494	**1000+**	**1000+**	-0.04	**1000+**	0.1494	**1000+**	**1000+**	-0.04	**1000+**	0.1494	
HM	**1000+**	**1000+**	-0.04	**1000+**	0.2493	**1000+**	**1000+**	-0.04	**1000+**	0.2493	**1000+**	**1000+**	-0.04	**1000+**	0.2493	
SW	**1000+**	**1000+**	-0.04	**1000+**	0.8223	**1000+**	**1000+**	-0.04	**1000+**	0.8223	**1000+**	**1000+**	-0.04	**1000+**	0.8223	

Table 4.11. Same as Table 4.02, but for the discrimination condition in experiment 4 (Psychometric Functions). The CMC was set to -1 for all fits, in accord with the dilution assumption.

The ERD of the window obtained from the discrimination task with the CMC parameter set to -1 was large and indeterminate for each listener for all fits (see Table 4.11). As for the previous experiment, all simple fits had an equal χ^2 value, and all skirt fits had an equal χ^2 value, although χ^2 values were lower for simple fits. The time constants that describe both the peak and (for the skirt fits) the skirt values are again large and indeterminate, thus providing little room for the base function to have an influence. All the fits are rectangular, thus the χ^2 value cannot be used to assess goodness-of-fit.

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)	tp(ms)	ERD	of-fit (χ^2)
AK	**1000+**	**1000+**	0.0071	**1000+**	**1000+**	0.0071	**1000+**	**1000+**	0.0071
HM	**1000+**	**1000+**	0.1316	**1000+**	**1000+**	0.1316	**1000+**	**1000+**	0.1316
SW	**1000+**	**1000+**	0.3366	**1000+**	**1000+**	0.3366	**1000+**	**1000+**	0.3366

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD	of-fit (χ^2)
AK	**1000+**	**1000+**	-0.04	**1000+**	0.0213	**1000+**	**1000+**	-0.04	**1000+**	0.0213	**1000+**	**1000+**	-0.04	**1000+**	0.0213
HM	**1000+**	**1000+**	-0.04	**1000+**	0.2632	**1000+**	**1000+**	-0.04	**1000+**	0.2632	**1000+**	**1000+**	-0.04	**1000+**	0.2632
SW	**1000+**	**1000+**	-0.04	**1000+**	0.6732	**1000+**	**1000+**	-0.04	**1000+**	0.6732	**1000+**	**1000+**	-0.04	**1000+**	0.6732

Table 4.12. Same as Table 4.02, but for the discrimination condition in experiment 4 (Psychometric Functions). The CMC parameter was set according to the data obtained for each individual participant from experiment 1 (-1.24 for AK, -0.99 for HM, and -1.06 for SW).

For the discrimination condition (Table 4.12) with individual CMC parameters, as when the CMC parameter was set to -1, the best fit in all cases was given by a simple fit of indeterminate shape and ERD (compare Tables 4.11 and 4.12).

The effect of the CMC parameter on the temporal window is now described. Two of the three participants who participated in experiment 1 had individual CMC's steeper than -1 (AK: -1.24, SW: -1.06), and the third was slightly shallower (HM: -0.99). Modelling of the results from experiment 3 demonstrated that the steeper the CMC, the lower the ERD (e.g. compare Tables 4.04 and 4.05 for modelling of the detect 4 data). In the detection condition of the current experiment (compare Tables 4.09 and 4.10), in the case of AK, the ERD was indeterminate when windows were fitted using both individual and dilution (i.e. -1) CMC parameters. For listener HM, when windows were fitted using the dilution CMC parameter, the ERD of the best fit was 143.95 ms. This increased to 159.38 ms when the CMC was set to -0.99. In the case

of SW, the ERD was indeterminate when windows were fitted using both individual and dilution (i.e. -1) CMC parameters.

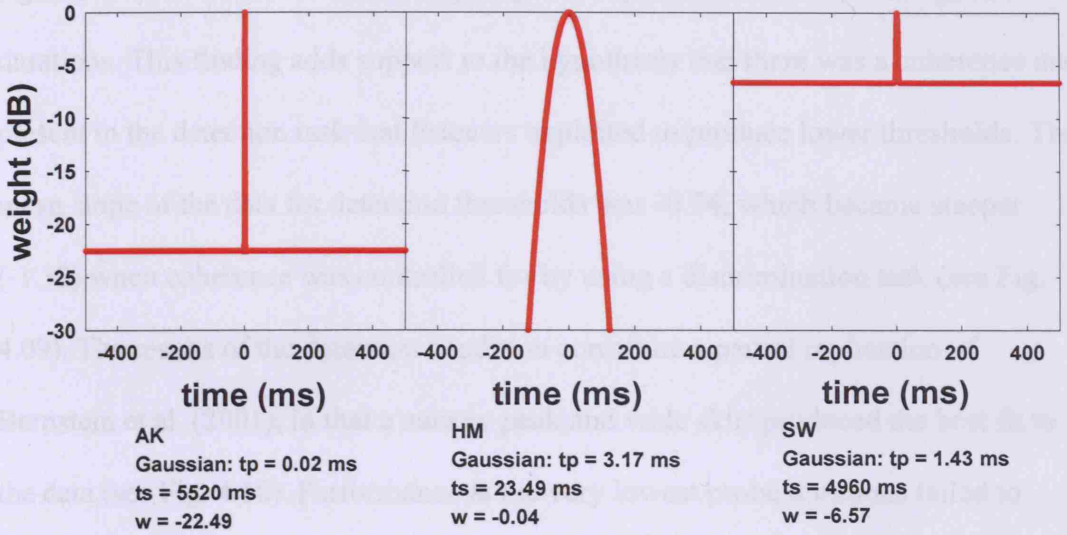

FIG. 4.10. Best skirt-function window fits to individual data in the detection condition with the CMC parameter set to -1 for all fits. Although the three fits to AK's data all produced equal values of χ^2, only Gaussian fits are illustrated to compare to the best fits of the other listeners. Parameters are obtained from Table 4.09.

The best skirt-function window fits to individual data in the detection condition are graphed in Fig. 4.10. Both AK and SW's fits are well described by a narrow peak and an indeterminate skirt function. HM's data is best fitted with a narrow peak and a wider skirt, although the majority of the fit is made up of the skirt function (w = -0.04). The Gaussian fit to AK's data produces an ERD of 110.46 ms, however this value is not constrained by the stimulus duration. The ERD of the window within the 100-ms stimulus is only 17.1 ms. The different shapes of window fitted to listener AK's data all produced equal values of χ^2, and the exponential and rounded-exponential fits produced indeterminate ERDs. The ERDs of the best skirt fits to HM and SW's data were both indeterminate.

4.2.4 Discussion.

As in experiment 3, detection and discrimination thresholds were found to be significantly different, where the disparity was especially marked at short probe durations. This finding adds support to the hypothesis that there was a coherence cue present in the detection task that listeners exploited to produce lower thresholds. The mean slope of the data for detection thresholds was -0.74, which became steeper (-1.32) when coherence was controlled for by using a discrimination task (see Fig. 4.09). The results of the detection condition constitute a partial replication of Bernstein et al. (2001), in that a narrow peak and wide skirt produced the best fit to the data (see Fig. 4.10). Performance at the very lowest probe durations failed to replicate, as (unlike Bernstein et al.'s listeners) none of the participants were able to produce measurable thresholds at a probe duration of 2 ms, and participant HM was unable to produce a threshold at a probe duration of 8 ms (see Fig. 4.08). Indeterminate windows for all listeners were obtained in the discrimination condition (see Tables 4.11 and 4.12).

Unlike the slope of the discrimination data in the previous experiment, the slope of the present discrimination data was steeper than -1 (i.e. mean: -1.32), which suggests that listener performance is worse than what would be expected if their performance was determined by simple long-term temporal integration. This effect could be attributable to decorrelation at the boundaries between the probe and the diotic noise distracting the listeners. In the detection task, participants could use the decorrelation cue to help them solve the task, which improved performance at the lowest probe durations and produced the shallow slopes seen for detection thresholds (mean: -0.74). The same cue could hamper listener's performance if they persevered with this cue when they performed the discrimination task, producing the steeper slope.

The next experiment attempted to measure the temporal window by taking psychometric functions for 500-ms-long stimuli. To prevent any confound from coherence cues associated with a detection task, participants only performed a discrimination task.

4.3 EXPERIMENT 5. LONG STIMULI WITH DIOTIC INTERFERING NOISE.

4.3.1 Introduction.

When coherence was controlled for by using a discrimination task, indeterminate windows were obtained for stimuli with a duration of 100 ms. In order to measure the temporal window, the length of the stimuli must be extended for a time period that is great enough to encompass the window in its entirety. The fifth experiment utilised a 2-interval discrimination task for stimuli with a duration of 500 ms, to obtain psychometric functions for a wider range of probe durations than used in the preceding experiment.

4.3.2 Method.

Participants

Three listeners took part. Two of the listeners (AK and SW) had taken part in all the preceding experiments took part in the present experiment. Another participant, RH, had taken part in experiments 1 and 2. They were paid upon completion.

Apparatus/Materials

The experiment consisted of a 2I-2AFC discrimination task. The total duration of each interval was 500 ms, with 10-ms gated onset and offset. The inter-stimulus interval was 500 ms. The durations of the probe noises were 362, 256, 181, 128, 90,

64, 45, 32, 22.5 and 16 ms. The probes were presented with delays of 1024, 512, 256, 128, 64, 32, 16 and 8 microseconds.

Design

The experiment manipulated two independent variables within subjects: probe duration (362, 256, 181, 128, 90, 64, 45, 32, 22.5 and 16 ms), and ITD (1024, 512, 256, 128, 64, 32, 16 and 8 microseconds). The dependent variable was percentage of correct responses.

Procedure

Participants performed a 2I-2AFC discrimination task, responding using the keyboard as in the previous experiment.

Initially, a block of trials was presented with probe durations of 256 ms, and the order in which the eight probe ITDs were presented within the block was randomised. Blocks of trials with probe durations of 128, 64, 32 and 16 ms followed. After these blocks were completed, participants completed blocks of trials at probe durations of 362, 181, 90, 45 and 22.5 ms. The procedure was repeated so that 20 trials were repeated for each ITD and probe duration. Each run therefore consisted of 8 ITDs x 10 probe durations x 20 repetitions = 1600 trials in total.

4.3.3 Results.

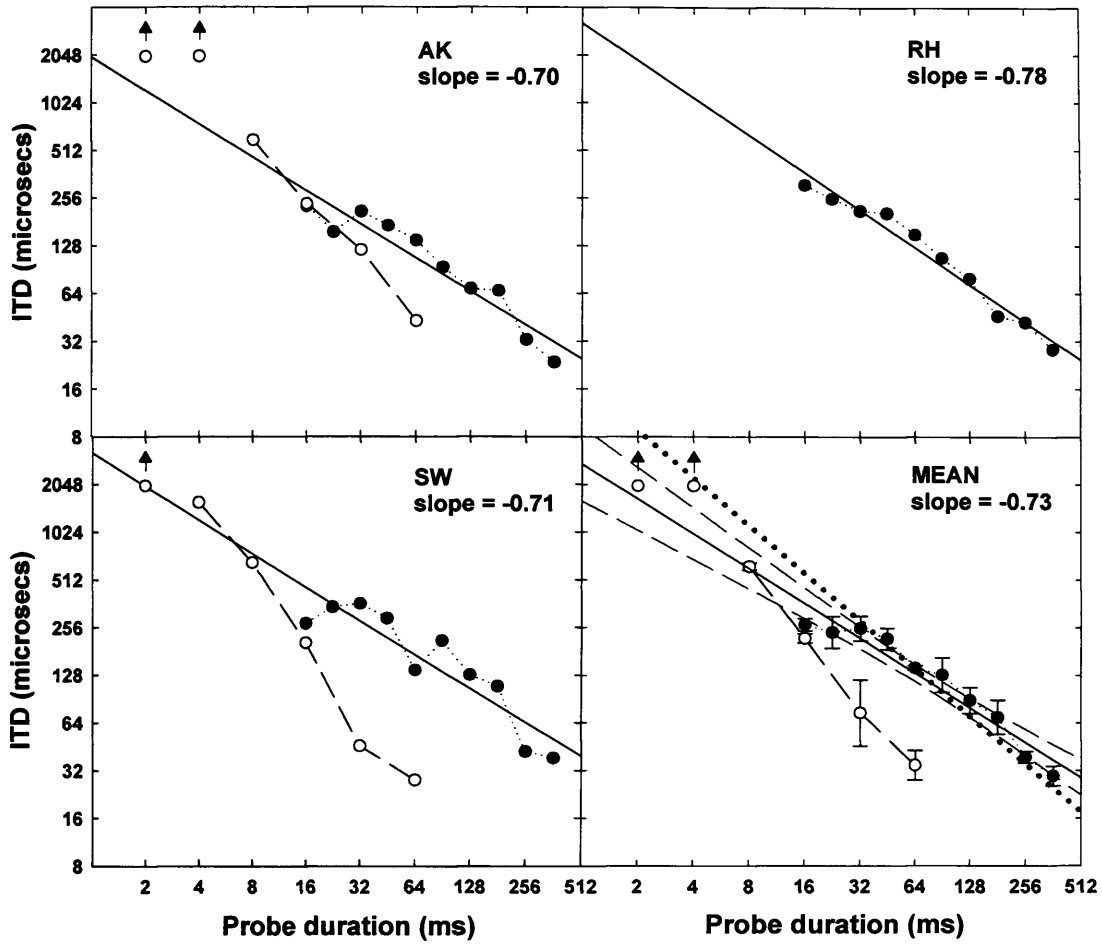

FIG. 4.11. Results for experiment 5 (Long stimuli with diotic interfering noise), indicated by closed circles joined by thin dotted lines. Solid regression lines are plotted for each listener. Dashed lines represent 95% confidence intervals. The thick dotted line represents a slope of -1. Open circles joined by thin dashed lines indicate thresholds taken with 100 ms long stimuli (taken from experiment 4). Arrows above data points arbitrarily plotted at an ITD of 2 ms indicate that threshold fell outside the measured range of the psychometric function. Mean thresholds from experiment 4 are the average of listeners AK and SW only. Both axes are plotted logarithmically.

The results for the three participants are shown in Figure 4.11. Each plotted threshold is derived by fitting a psychometric function to the average of the listener's last three experimental runs. The mean slope of the empirical data was –0.73, $r^2 = 0.93$. The slope of the linear regression line was significantly lower than -1, as indicated by the dotted line that signifies a slope of -1 falling outside the confidence intervals of the regression. However, there appears to be a non-linear trend of steepness with duration; at short durations, thresholds levelled off at approximately 250 μs for listeners AK and SW. This pattern was not observed for the 100 ms long stimuli. At the longer probe durations, thresholds fall on a slope of approximately -1. The triangles indicate thresholds obtained in experiment 4 with 100 ms long stimuli. Mean thresholds for the 100 ms long data are the average of listeners AK and SW only, as listener RH did not participate in experiment 4.

Simple and skirt Gaussian, exponential and roex windows were fitted to the data in the same format as in the previous two experiments.

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD (ms)	of-fit (χ^2)	tp(ms)	ERD (ms)	of-fit (χ^2)	tp(ms)	ERD (ms)	of-fit (χ^2)
AK	104.24	206.81	0.1205	61.95	219.62	0.1497	105.88	210.94	0.1401
RH	143.16	286.31	0.0486	70.23	247.46	0.0617	130.60	262.23	0.0580
SW	116.12	232.23	0.1905	71.04	251.82	0.2185	120.67	241.35	0.2106

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)
AK	0.78	**1000+**	-13.25	**1000+**	0.0404	1.39	**1000+**	-8.69	**1000+**	0.0404	0.83	**1000+**	-13.01	**1000+**	0.0404
RH	0.29	**1000+**	-15.36	**1000+**	0.0125	0.64	**1000+**	-9.72	**1000+**	0.0125	0.08	**1000+**	-20.88	748.74	0.0125
SW	0.09	**1000+**	-22.48	**1000+**	0.1014	0.21	**1000+**	-16.46	**1000+**	0.1014	1.14	**1000+**	-11.71	**1000+**	0.1014

Table 4.13. Window fitting parameters for the discrimination task with long stimuli with diotic interfering noise in experiment 5 (as for Table 4.02). The CMC parameter was set to -1 for all fits, in accord with the dilution assumption.

	Fitting function (simple functions)								
Listener	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
	tp(ms)	ERD (ms)	of-fit (χ^2)	tp(ms)	ERD (ms)	of-fit (χ^2)	tp(ms)	ERD (ms)	of-fit (χ^2)
AK	50.46	100.92	0.1733	30.53	108.22	0.2421	52.96	105.51	0.2120
RH	90.55	181.10	0.0654	50.65	179.54	0.0925	89.93	179.86	0.0827
SW	89.39	180.18	0.2060	57.43	202.36	0.2464	93.49	188.45	0.2326

	Fitting function (skirt functions)														
Listener	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)	tp(ms)	ts(ms)	w(dB)	ERD (ms)	of-fit (χ^2)
AK	0.28	**1000+**	-20.11	**1000+**	0.0312	0.03	**1000+**	-27.12	177.82	0.0312	1.32	**1000+**	-13.55	**1000+**	0.0312
RH	0.73	**1000+**	-13.17	**1000+**	0.0105	0.12	**1000+**	-18.19	285.79	0.0105	0.29	**1000+**	-16.96	317.33	0.0105
SW	0.68	**1000+**	-14.57	**1000+**	0.0971	1.07	**1000+**	-10.39	**1000+**	0.0971	1.13	**1000+**	-12.46	**1000+**	0.0971

Table 4.14. As for Table 4.13, but all fits were modelled with the CMC parameter set according the data obtained for each individual participant from experiment 1 (-1.24 for AK, -1.11 for RH, and -1.06 for SW).

When windows were fitted to the data for both individual and dilution CMC parameters (see Tables 4.13 and 4.14), the best fit was found to be a skirt window for all three listeners. The peak parameter was short (less than 2 ms), and the skirt of the window obtained was large and indeterminate for each listener, with the result that the shape of the best-fitting window could not be determined. The narrow peak and long skirt, similar to that obtained in the Bernstein et al. (2001) replications described in experiments 3 and 4, meant that the ERD of the fitted windows was large, and in many cases indeterminate. The equal χ^2 values are related to the narrow time constants that describe the peak and the very large time constant describing the skirt, preventing the base function having an influence.

4.3.4 Discussion.

Unlike the discrimination data in the preceding experiments, the mean slope of the data (-0.73) was significantly shallower than -1 (see Fig. 4.11), suggesting that the

participants were not integrating the entire stimulus, and were instead performing temporal resolution using a temporal window. However, when windows were fitted to the data, in all cases they demonstrated a narrow peak and indeterminate skirt (see Tables 4.13 and 4.14).

A possible explanation for this may be that listeners could have been susceptible to the distracting effects of auditory events that occurred at the transitions between interferer and probe. Because the change from noise with zero delay to noise with an ITD was instantaneous, listeners may not have been able to track the change in perceived location, and will instead have been distracted by a blurred or wide noise image at these points (see Fig. 4.12).

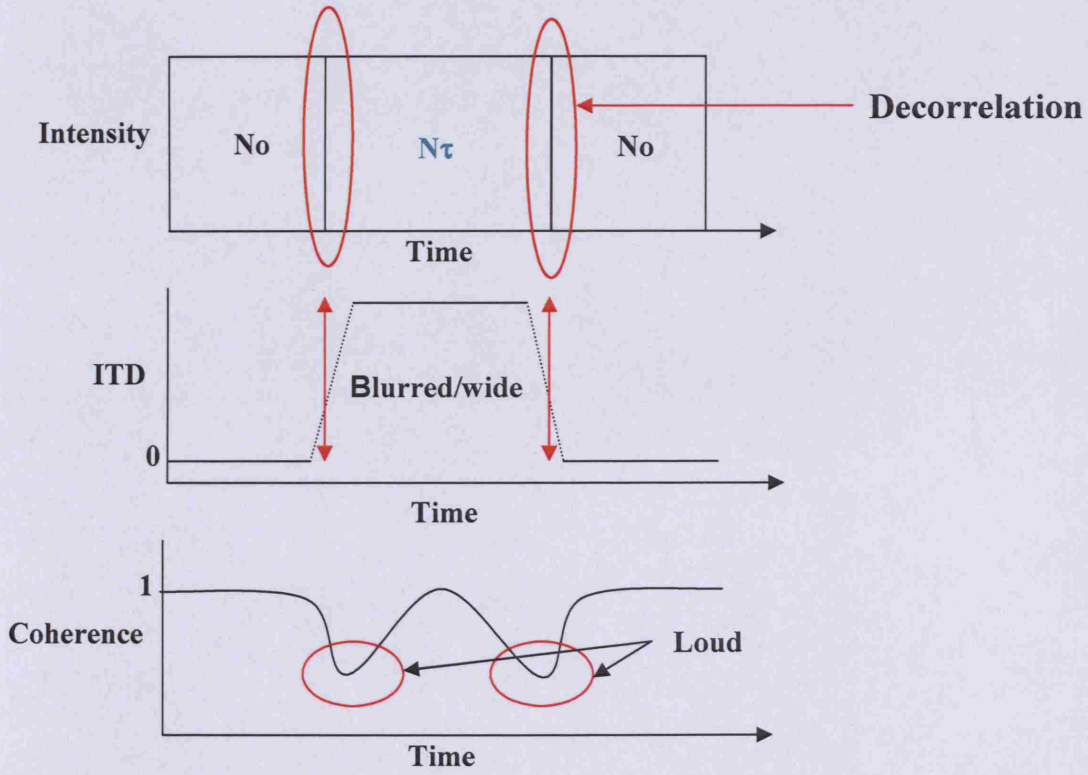

FIG. 4.12. Schematic illustration of the effects of distraction. Changes in ITD and coherence when the probe begins and ends may result in changes in image width and loudness, distracting the listener.

Distraction effects caused by abrupt transitions from No to Nπ noise were described by Culling and Summerfield (1998, see their footnote 3). In that study, a short out-of-phase signal was presented against a masker whose interaural correlation changed from uncorrelated (Nu) noise, to correlated (No) noise, then back to uncorrelated (Nu) noise. The design of the experiment was a refinement from a pilot study in which Nπ noise replaced the uncorrelated sections of the masker. When Nπ noise was employed, listeners often produced higher thresholds for brief No durations than they did when the No duration was zero, a finding that would seem to be inconsistent with that predicted by the temporal window model. However, the change in perceived location at the transition points between Nπ and No noise in the masker was found to have a detrimental distracting effect, hence the use of uncorrelated noise in the final study. A similar problem was highlighted by Kollmeier and Gilkey (1990, footnote 3), who measured the threshold of tones masked by broadband noise that changed in configuration from No to Nπ at a point close to the presentation of the tone.

In the present study, although the magnitude of the decorrelation caused by transition from No to Nτ at the boundaries between the probe and the contiguous diotic noise is not as large as that of No to Nπ, the change may still affect the listener's thresholds. At each stage of the stimulus, the following percept of the noise image can be expected: centralised and narrow during the No portion, wide at the first transition to the probe, a narrow percept lateralised to one side during the Nτ portion, wide at the second transition, and centralised and narrow during the last portion of No noise. Although the transitions from stage to stage are too rapid to perceive independently they could be a source of distraction for the listener, providing that the distraction effect was duration-specific (i.e. the distraction effect did not appear to be present in the previous discrimination tasks as the 100-ms long stimuli were too brief

for the transitions to be a source of distraction for the listener). If the image width (Gabriel and Colburn, 1981)/loudness change (Culling and Edmonds, 2006) at the two transitions was having a detrimental effect on the listener, a way to reduce the impact of these events could be to use uncorrelated or semi-correlated noise instead of correlated noise at the start and end of each interval. Uncorrelated noise produces a percept that has a wide spatial extent, whereas correlated noise sounds narrow and centralised in the head. By incorporating Nu noise instead of No noise, the width percept would be more stable across the interval and less distracting, as the decorrelation at the transition points would no longer be a sudden 'blip'. In contrast to the present experiment, where the task is to discriminate a small change in ITD in the context of two large changes in spatial extent, when an uncorrelated interferer is employed there is only one event, the change in ITD in both intervals (see Akeroyd and Summerfield, 1999).

For the 500-ms data, as probe duration increased beyond 32 ms, thresholds decreased in a linear fashion (see Fig. 4.11). At some point, thresholds should stop improving with increasing probe duration, as at that point the temporal window should fill with delayed noise and no further improvement in temporal resolution should occur. Tobias and Zerlin (1959) measured ITD thresholds for white noise stimuli bursts at durations ranging from 10 to 1940 ms. Thresholds were observed to show an improvement as burst duration increased up to approximately 700 ms, beyond which an asymptote of approximately 6 μs was reached. In the current experiment, as the ERD of the temporal window has been found to be approximately 100 ms (e.g. Culling and Summerfield, 1998) it was thought that the lowest thresholds would be observed at the longest probe durations (256 and 362 ms), as only delayed noise would enter the window. However, this did not occur. It is possible that the

auditory events occurring at the transition points between diotic and delayed noise may be distracting the listener even at the longest probe durations. This may explain why threshold does not flatten off as probe duration is increased, and may be a contributing factor in explaining the disparity between thresholds for the 100-ms data from experiment 4 and the higher thresholds for the 500-ms data obtained at comparable probe durations. At probe durations of 16, 32 and 64 ms, the thresholds previously obtained for the 100-ms stimuli in experiment 4 are lower than thresholds obtained for 500 ms long stimuli, and thus increasing the overall duration of the interval results in a rightward shift in thresholds (see Fig. 4.11).

To investigate the effect of the narrow peak on the efficiency of the window-fitting function, windows were modelled to the mean data with the parameters set to $tp = 1$ ms and $ts = 10$ s. w was set to either 0 dB, -5 dB or -10 dB. From the fits to the data, is appears that the peak is accounting for the 'levelling off' of the data at very short probe durations (see Fig. 4.13).

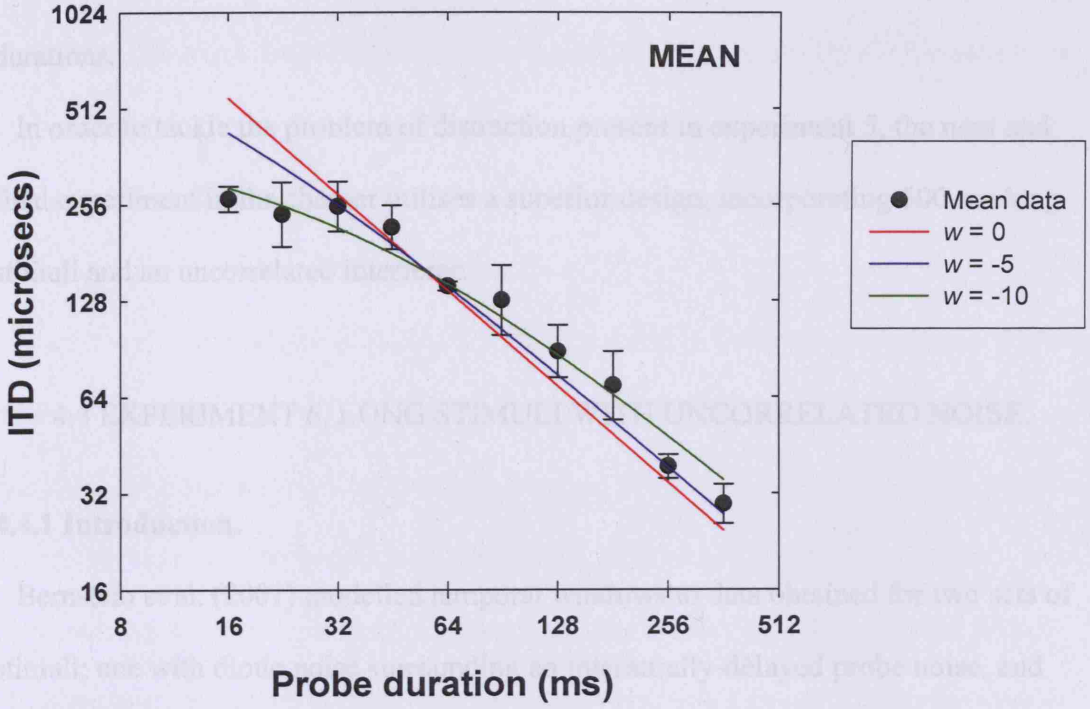

FIG. 4.13. The effect of window peak on goodness of fit. The lines represent window fits with a narrow peak and long skirt to the mean data shown in Fig. 4.11. The red line indicates a fit where the window is composed only of the skirt, the blue line represents the fit when the time constant describing the skirt parameter is only moderately weighted, and the green line represents the fit when the skirt is less heavily weighted. Both axes are plotted logarithmically.

When $w = 0$ dB and the window is composed of the skirt function only, threshold slope is -1, consistent with complete integration of the stimulus for a rectangular window. As w decreases (lowering the skirt), a better fit to the thresholds at low probe durations is obtained. It appears that the narrow peak is accounting for the thresholds at the shortest probe durations, which do not demonstrate improvement as probe duration is increased. Thus, simple windows produced a worse fit to the data than skirt windows (see Table 4.13), as simple windows do not account for the lack of

improvement as probe duration is increased for thresholds at the shortest probe durations.

In order to tackle the problem of distraction present in experiment 5, the next and final experiment in the chapter utilises a superior design, incorporating 500-ms long stimuli and an uncorrelated interferer.

4.4 EXPERIMENT 6. LONG STIMULI WITH UNCORRELATED NOISE.

4.4.1 Introduction.

Bernstein et al. (2001) modelled temporal windows to data obtained for two sets of stimuli; one with diotic noise surrounding an interaurally-delayed probe noise, and one with uncorrelated noise surrounding the probe. Their stimuli had a total duration of 100, 40 or 20 ms, and both sets of data produced short windows with an ERD of about a millisecond. Fig. 4.14 is a plot of the results obtained for the uncorrelated noise case.

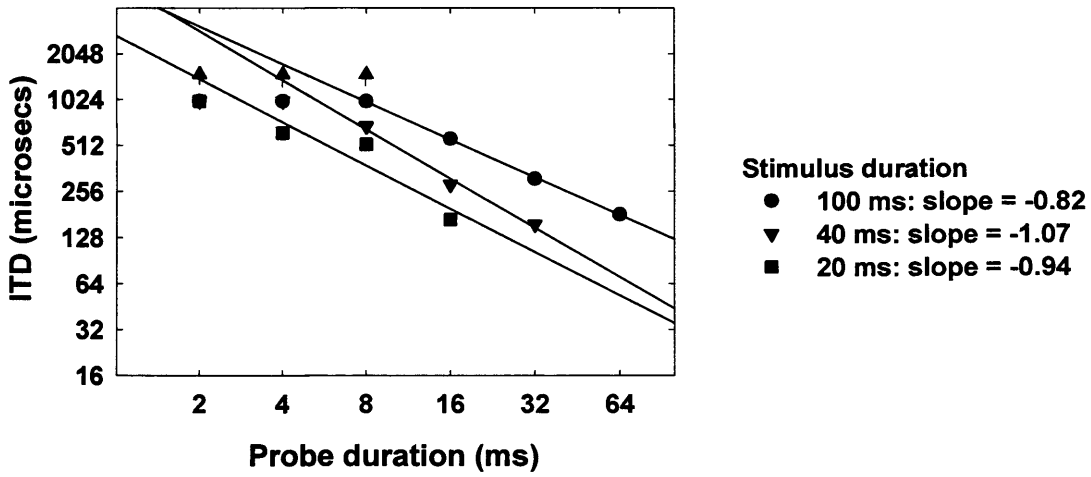

FIG. 4.14. Thresholds for detection of an ITD applied to the probe segment of a noise temporally flanked by interaurally-uncorrelated noise taken from Fig. 3 of Bernstein et al. (2001), but with both axes plotted logarithmically.

The design of the experiment with uncorrelated noise contiguous to the probe was superior to the experiment in which diotic noise was employed. This is because there was a change in the interaural parameters between the probe and masker in both the target and distracter intervals, so the listener cannot have used width or loudness changes as a cue to identifying the target.

The lowest Uncorrelated Masking Coefficient (UMC) observed in experiment 2 was approximately -1.4 (see Fig. 3.5). The slopes of the data for the three stimulus durations are considerably shallower (they range from -0.82 to -1.07, see Fig. 4.14), suggesting that listeners were using a temporal window to resolve the stimulus. However, because the overall coherence of the interval is proportional to the probe duration, as probe duration decreases, thresholds may have risen because the overall coherence of the intervals falls, not because the listeners were employing temporal resolution. The lower the coherence, the higher threshold ITD becomes, as demonstrated by experiment 2 (see Figs. 3.4 and 3.5), thus what determines listeners' performance is the overall coherence of the interval, and this is proportional to the probe duration.

An attempt was made to see to what extent the data of Bernstein et al. (2001) for uncorrelated noise preceding and lagging the probe could be predicted on the basis of coherence judgements alone. For each probe duration and stimulus duration, the average long-term interaural coherence was derived by calculating the area of a temporal window occupied by the probe and dividing that value by the total integral of the temporal window (limited to the overall duration of the stimulus interval). The temporal window used to calculate the various coherences was the simple Gaussian window obtained by Culling and Summerfield (1998) for a centre frequency of 500 Hz at a level of 40 dB. Thresholds were predicted from the data of Jeffress et al.

(1962) by taking the standard deviation of the listener's centring judgement as equivalent to ITD threshold change, and reading off the threshold from Fig. 1 of Jeffress et al. for the coherence of the stimulus at each probe duration. Thresholds were also predicted by reading off the ITD threshold from the 100-ms stimulus duration data from experiment 2 (see Fig. 4.15).

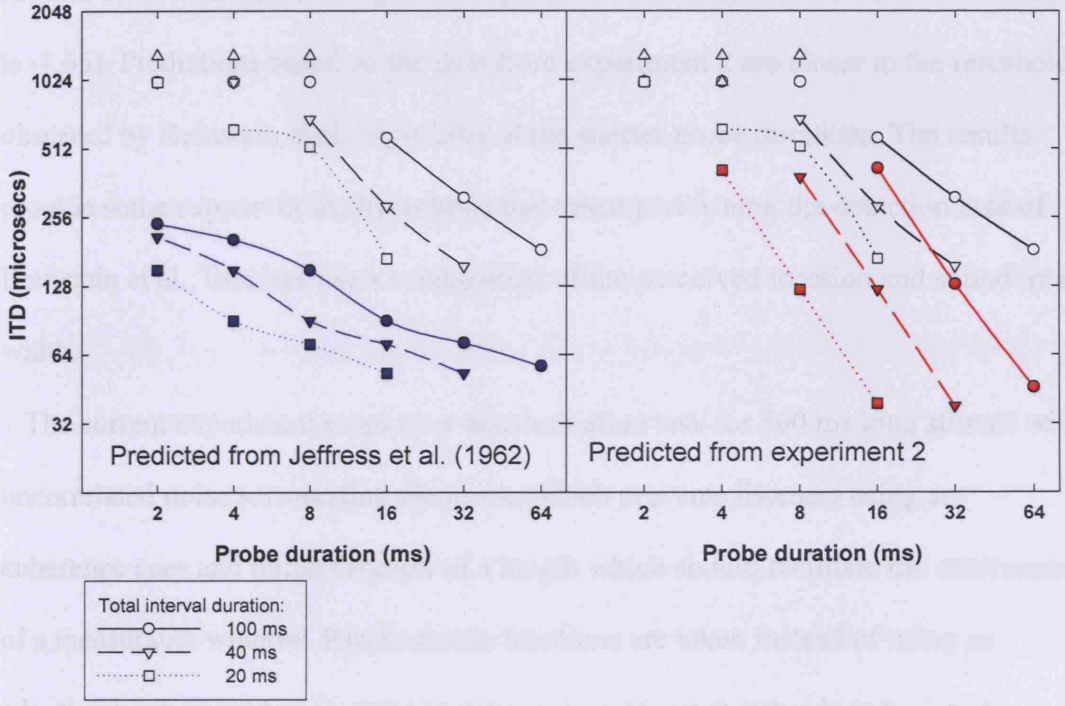

FIG. 4.15. Predicted thresholds for detection of a delayed probe noise in the presence of uncorrelated interfering noise. Circles joined by solid lines indicate thresholds for 100-ms stimuli, triangles joined by long-dashed lines for 40-ms stimuli, and squares joined by dotted lines for 20-ms stimuli. Open symbols in both graphs represent thresholds taken from Fig. 3 of Bernstein et al. (2001). Closed blue symbols represent predicted thresholds read off from Fig. 1 of Jeffress et al. (1962). Closed red symbols represent predicted thresholds predicted by the 100-ms duration data from experiment 2 with a UMC of -1.72. Both axes are plotted on a logarithmic scale.

The pattern of thresholds predicted by Jeffress et al. (1962) and the data from experiment 2 is similar to that produced by the data of Bernstein et al. (2001), in that at all three total interval durations, thresholds increase as probe duration decreases. However, the thresholds are substantially lower, and the slopes of the data are shallower when predicted from Jeffress et al. (the mean slope of Bernstein et al.'s data at the three probe durations is -0.94, and -0.48 for the predictions obtained from Jeffress et al.'s data), and steeper when predicted from experiment 2 (the mean slope is -1.65). Predictions based on the data from experiment 2 are closer to the thresholds observed by Bernstein et al., especially at the shorter probe durations. The results provide some support to the hypothesis that when performing the detection task of Bernstein et al., listeners use a combination of the perceived location and sound image width.

The current experiment employs a discrimination task for 500 ms long stimuli with uncorrelated noise surrounding the probe, which prevents listeners using any coherence cues and utilises stimuli of a length which should facilitate the attainment of a measurable window. Psychometric functions are taken instead of using an adaptive track procedure in order to accurately measure thresholds at low probe durations. Results are compared to those found in experiment 5. UMCs (Uncorrelated Masking Coefficients) obtained in experiment 2 are employed in the modelling procedure.

4.4.2 Method.

Participants

The three listeners that participated in experiment 5 took part in the present experiment. They were paid upon completion.

Apparatus/Materials

The experiment consisted of a 2I-2AFC discrimination task, and used stimuli with the same design as the discrimination stimuli used in the previous experiments. The total duration of each interval was 500 ms, with 10-ms gated onset and offset. The inter-stimulus interval was 500 ms. The duration of the probe noises were 362, 256, 181, 128, 90, 64, 45, 32, 22.5 and 16 ms. The probes were presented with delays of 1024, 512, 256, 128, 64, 32, 16 and 8 microseconds.

Design

The experiment manipulated two independent variables within subjects: probe duration (362, 256, 181, 128, 90, 64, 45, 32, 22.5 and 16 ms), and ITD (1024, 512, 256, 128, 64, 32, 16 and 8 microseconds). The dependent variable was percentage of correct responses.

Procedure

Participants performed a 2I-2AFC discrimination task, responding using the keyboard as in the previous experiment.

Initially, a block of trials was presented with probe durations of 256 ms, and the order in which the eight probe ITDs were presented within the block was randomised. Blocks of trials with probe durations of 128, 64, 32 and 16 ms followed. After these blocks were completed, participants completed blocks of trials at probe durations of 362, 181, 90, 45 and 22.5 ms. The procedure was repeated so that 20 trials were repeated for each ITD and probe duration. Each run therefore consisted of 8 ITDs x 10 probe durations x 20 repetitions = 1600 trials in total.

4.4.3 Results.

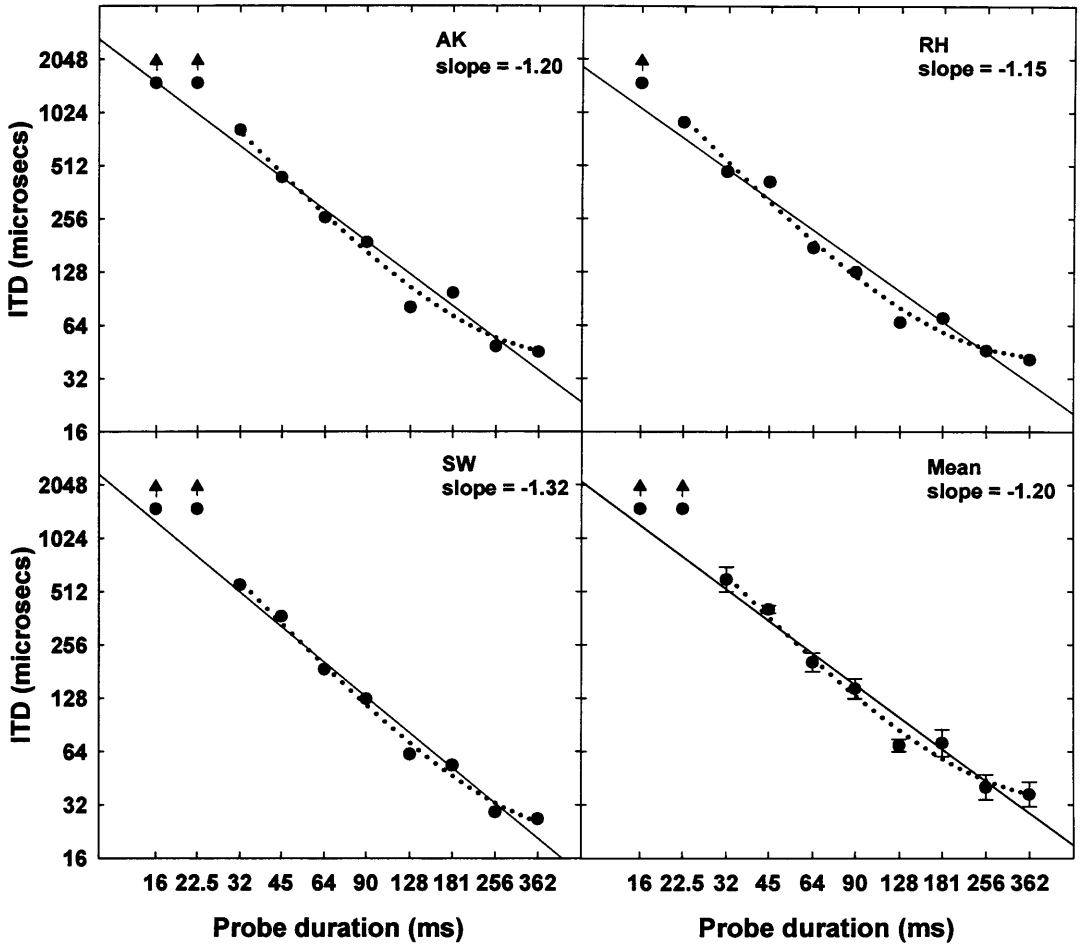

FIG. 4.16. Results for experiment 6 (long stimuli with uncorrelated interfering noise). Dotted lines indicate predicted thresholds when the data is fitted by the 'best' fitting roex window with a UMC of -1.64. Mean predicted thresholds were obtained by taking the mean of the three individual predicted thresholds. Arrows above data points arbitrarily plotted at an ITD of 2 ms indicate that threshold fell outside the measured range of the psychometric function. The slope of the mean regression line is -1.20, r^2 = 0.97, and both axes are plotted logarithmically.

Fig 4.16 shows individual and mean results for experiment 6. For all listeners, ITD threshold fell as probe duration increased, and flattened off at the longest probe

durations (256 and 362 ms). These results contrast with those obtained in the previous experiment for 500 ms long stimuli with a diotic interferer, where at short probe durations thresholds levelled off at approximately 250 μs, and did not flatten off at the longest probe durations (see Fig. 4.11). A measure of threshold could not be obtained at the shortest probe durations as the 75% point of the psychometric function fell outside the measured range.

1 and 3-parameter Gaussian, roex, and exponential temporal windows were modelled to the dataset provided by each participant with UMCs ranging from -1.4 to -1.9. The range was chosen to encompass the variety of UMCs obtained in experiment 2 for various probe durations and participants. The best fits are plotted in Fig. 4.17:

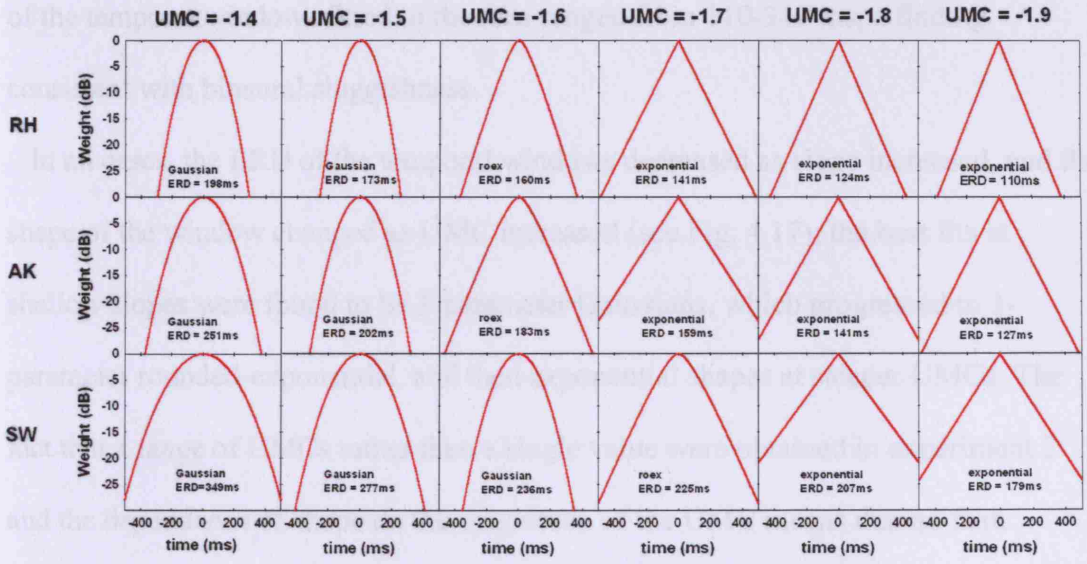

FIG. 4.17. Window fits for individual participants, experiment 6 (long stimuli with uncorrelated interfering noise). The different panels show best-fitting windows plotted for individual participants at UMCs ranging from -1.4 to -1.9.

In all cases, the best fits were found to be simple 1-parameter fits. ERD was found to decrease as the UMC parameter became steeper. The ERD of listener AK's temporal

window ranged from 251 ms at a slope of -1.4 to 127 ms at a slope of -1.9. Listener RH's ERD ranged from 198 to 110 ms, and listener SW's ERD ranged from 349 to 179 ms. The shape of the best fit was seen to change as a function of UMC. At shallower UMCs, best fits were obtained for Gaussian windows. At moderate UMCs (-1.6 for listeners AK and RH and -1.7 for SW) best fits were roex, and at steep UMCs the best fit was found to be an exponential.

4.4.4 Discussion.

For all participants, the slope of the data was shallower than any given UMC (see Fig. 4.16), suggesting that the participants were not integrating the entire stimulus, and instead were performing temporal resolution using a temporal window. The ERDs of the temporal windows fitted to the data ranged from 110-349 ms, a finding consistent with binaural sluggishness.

In all cases, the ERD of the temporal windows decreased as slope increased, and the shape of the window changed as UMC increased (see Fig. 4.17); the best fits at shallow slopes were found to be 1-parameter Gaussians, which progressed to 1-parameter rounded-exponential, and then exponential shapes at steeper UMCs. The fact that a range of UMCs rather than a single value were obtained in experiment 2 and the dependence of shape on the magnitude of the UMC means that no firm conclusions as to the shape of the binaural temporal window can be made from this study. The range of UMCs obtained in experiment 2 also makes it difficult to conclusively state the ERD. Instead, a wide range of ERDs were obtained by modelling various UMCs. Although the range (110-349 ms) is large compared to other studies, these ERD values are of the same order of magnitude as values measured in previous studies of binaural sluggishness, such as Grantham and

Wightman (1979), who obtained ERDs between 44 and 243 ms, and Culling and Summerfield (1998), who obtained ERDs ranging from 55 to 188 ms. In order to obtain a 'best' window, the mean of the UMCs measured for individual participants in experiment 2 (-1.64, see Fig. 3.4) can be used in the window fitting procedure. When temporal windows were modelled with a UMC of -1.64 for the three listeners, the best window in every case was a roex fit, and the mean ERD of the individual windows was 197 ms. This 'best' window is shown in Fig 4.18.

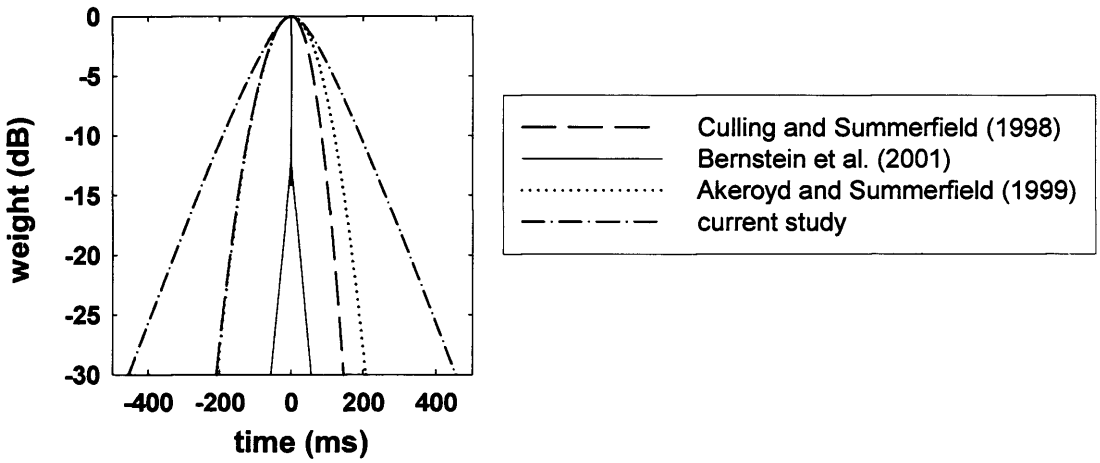

FIG. 4.18. Temporal window functions. The dash-dotted line plots the 'best' symmetric roex window obtained in the current study. The dotted line plots the symmetric Gaussian window obtained by Akeroyd and Summerfield (1999). The dashed window line plots the simple Gaussian window found by Culling and Summerfield (1998). The solid line plots the narrow double-exponential window found in the study by Bernstein et al. (2001).

The findings of the current study support the hypothesis that the binaural system is sluggish. The time constants for the temporal window, though obtained using a lateralization task, are comparable to those obtained in previous studies using a

detection task with BMLD as the dependent variable (Grantham and Wightman, 1979; Culling and Summerfield, 1999). The use of an adaptive track was found to be a problematic method when used to measure ITD threshold change due to an associated cue change from lateralised and fused to lateralised and diffuse at large ITDs. The measurement of psychometric functions was found to ameliorate the cue change problem associated with the adaptive track methodology, and was essential when an uncorrelated interfering noise was present due to the large magnitude of ITD thresholds at low stimulus coherences. The best fit was observed to be a symmetric roex fit with an ERD of 197 ms.

CHAPTER 5. SPECTRAL ANALYSIS

The preceding chapters examined the limits of the binaural system when attempting to resolve temporal fluctuations. Now, the experimental paradigm is transformed from the time to the frequency domain, in order to examine how capable the binaural system is at resolving spectral features. The aim of the experiment described in this chapter was to attempt to resolve some of the inconsistencies in the auditory filter literature regarding the width of the filter, as some studies have found evidence of narrow filter bandwidths (e.g. Kohlrausch, 1988), while others have indicated wider bandwidths (e.g. Sondhi and Guttman, 1966). Binaural filter bandwidths were measured using a binaural analogue of the notched-noise technique used in the measurement of monaural filter bandwidths. The design is similar to that of Sondhi and Guttman (1966).

Hall et al. (1983) contrasted NoSo and NoSπ BMLD thresholds using both bandlimited masking noise and a notched-noise technique. The bandlimited noise data demonstrated that noise beyond the monaural critical band centred on the signal affected thresholds, whereas the notched-noise data indicated that the binaural auditory filter was comparable to the monaural auditory filter. This led them to suggest that in the wideband noise conditions the auditory system is influenced by filters adjacent to the filter centred at the signal frequency. In broadband noise, the adjacent filters contain little or no interaural differences, which Hall et al. attributed to being detrimental to the detection of the out-of-phase signal. When the bandwidth of the masker is narrowed to the monaural critical bandwidth, relatively fewer critical band filters in the binaural array will display no interaural difference in the presence of the signal, thus the signal-to-noise ratio at the critical band centred at the signal will be lower at threshold (Hall and Fernandes, 1984).

Sondhi and Guttman (1966) used broadband noise to mask a low-frequency tone in a variety of configurations, at two centre frequencies (250 and 500 Hz). The masker was divided into three successive bands, where the interaural phase of the middle band was antiphasic with respect to the lower and upper bands. The following notation similar to that used in binaural unmasking is used to describe the stimuli. The suffix following N symbol describes the interaural phase at the two ears of successive frequency bands of the noise, and the suffix following S symbol describes the interaural phase of the signal. For example, in the NπoπSπ configuration (see Fig. 5.1.), the interaural phase of the masker was set to zero for the frequency band surrounding an out-of-phase tone (Sπ), while noise components outside the band were set to π (Nπoπ). The four stimulus configurations that were tested were NoπoSo, NπoπSo, NoπoSπ, and NπoπSπ, and masked thresholds were obtained for a range of bandwidths (*B*) of the No portion of the masker.

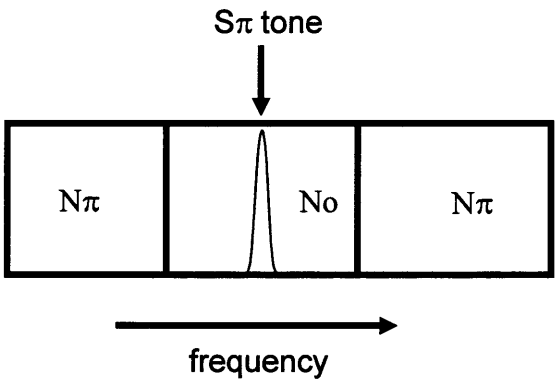

FIG. 5.1. A schematic of the stimuli used by Sondhi and Guttman (1966). The stimulus configuration in this example is NπoπSπ.

As their study predated the use of the ERB as a measure of filter bandwidth, Sondhi and Guttman defined the binaural critical bandwidth as the value of *B* at which the

release of masking in the antiphasic inner-band conditions was one-half the maximum BMLD, and obtained values of 125 Hz at a centre frequency of 250 Hz, and 200 Hz at a centre frequency of 500 Hz.

In order to derive bandwidth estimates comparable with modern monaural bandwidth measurements, Sondhi and Guttman's data was modelled with simple and skirt (1 and 3-parameter respectively) symmetric Gaussian, exponential, and roex auditory filters. The results are shown in Table 5.1:

	Fitting function (no skirt)								
Signal	Exponential		Goodness-	Gaussian		Goodness-	Rounded Exponential		Goodness-
frequency (Hz)	tp(Hz)	ERB	of-fit (χ^2)	tp(Hz)	ERB	of-fit (χ^2)	tpl(Hz)	ERB	of-fit (χ^2)
250	19.88	39.76	0.38	19.27	68.32	0.18	27.71	55.42	0.25
500	51.46	102.92	0.55	51.32	181.94	0.43	72.88	145.75	0.51

	Fitting function (skirt functions)														
Signal	Exponential				Goodness-	Gaussian				Goodness-	Rounded Exponential				Goodness-
frequency (Hz)	tp(Hz)	ts(Hz)	w(dB)	ERB	of-fit (χ^2)	tp(Hz)	ts(Hz)	w(dB)	ERB	of-fit (χ^2)	tp(Hz)	ts(Hz)	w(dB)	ERB	of-fit (χ^2)
250	19.90	157.66	-97.27	39.79	0.51	19.27	183.51	-96.36	68.29	0.24	27.69	287.77	-99.54	55.37	0.34
500	51.47	119.43	-97.01	102.93	0.73	51.27	125.20	-99.54	181.76	0.58	72.86	145.73	-99.41	145.73	0.67

Table 5.1. Auditory filters fitted to the data of Sondhi and Guttman (1966). *tp* is the parameter that describes the symmetric branch of the filter, *ts* is the skirt parameter, and *w* is the weight applied to the skirt of filters that have a skirt fitting function. The ERB is the integral in Hz of the entire filter. The goodness of fit (χ^2) is expressed for each individual window fit, where lower values of χ^2 indicate more efficient fits.

Best fits were found to be simple Gaussian fits, with an ERB of 69 Hz at a centre frequency of 250 Hz, and an ERB of 182 Hz at a centre frequency of 500 Hz. When skirt fits were modelled, the weighting of the skirt was very low for all fits, indicating that the skirt was exerting virtually no influence on the fit. The ERB of the symmetric exponential fit (102 Hz) is in rough agreement with Holube et al. (1998), who fitted a

double-sided exponential filter to the data of Sondhi and Guttman at a centre frequency of 500 Hz and obtained an ERB value of 92 Hz. The ERB should not be used as a statistic to compare filters of different shapes. When comparing auditory filters (monaural and binaural, and across studies) in terms of equivalent rectangular bandwidth, the same shape of filter must be compared, as the magnitude of the ERB is dependent on the shape of the filter that is assumed. To demonstrate this, figure 5.2 shows plots of the simple fits to Sondhi and Guttman's data from Table 5.1:

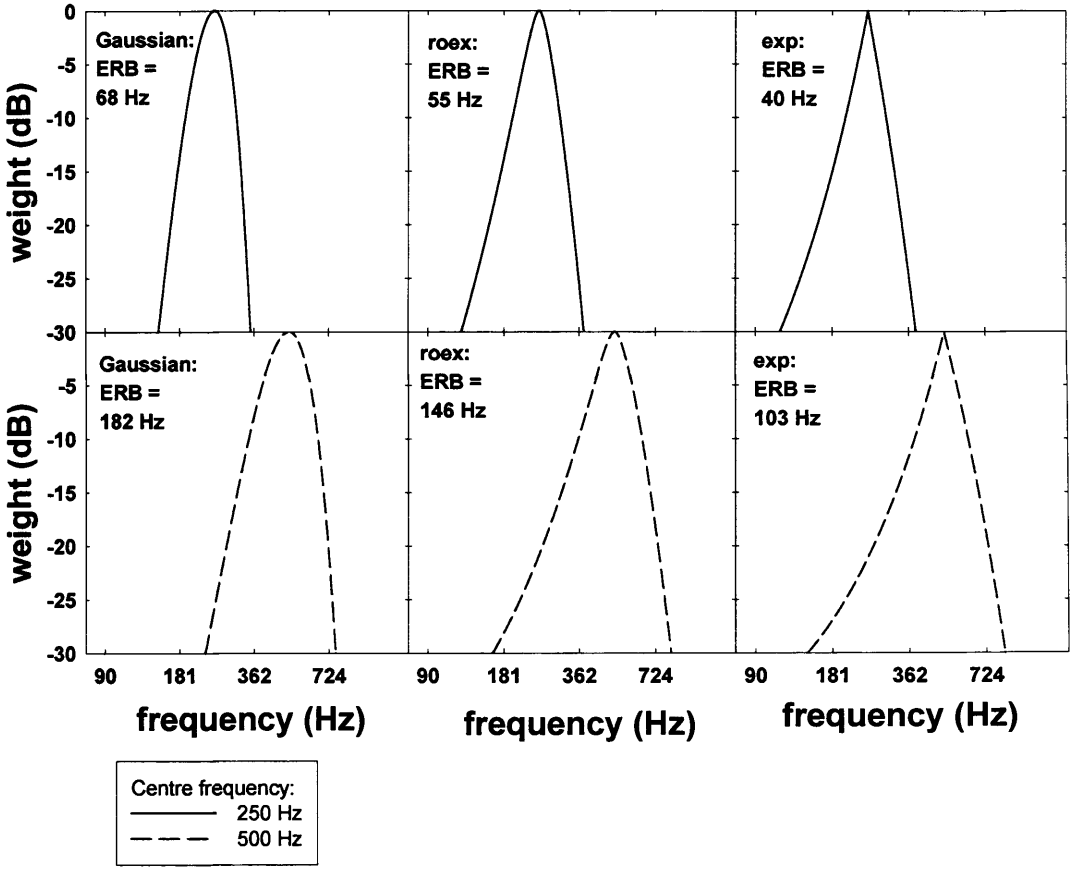

FIG. 5.2. Attenuation characteristics of binaural filters measured by Sondhi and Guttman (1966) at centre frequencies of 250 and 500 Hz. The x axis is plotted logarithmically.

Figure 5.2 illustrates that when different filter shapes are fit to the same data, the exponential shape produces the narrowest ERB values, and the Gaussian shape the widest ERB values. The ERB of the roex fit is between the exponential and Gaussian values. Figure 5.3 shows NπoπSπ data at the two centre frequencies tested by Sondhi and Guttman (1966), and with the data fitted using a roex filter superimposed.

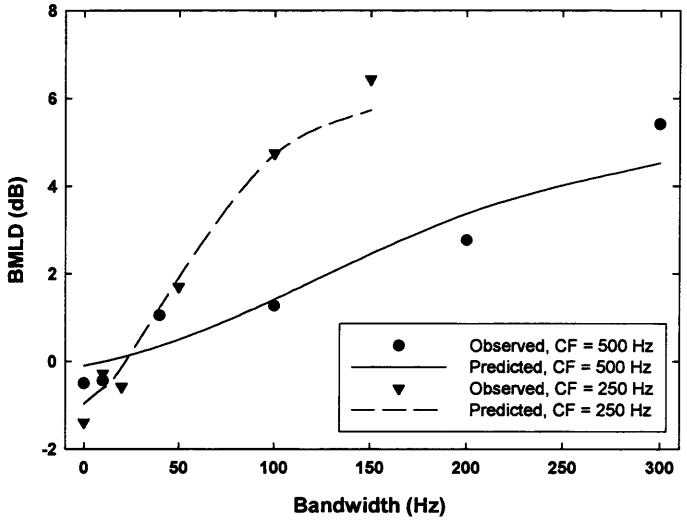

FIG. 5.3. Data observed in the NπoπSπ condition from Sondhi and Guttman (1966) at centre frequencies (CFs) of 500 Hz (closed circles) and 250 Hz (closed triangles). The solid line shows predicted values based on fitting a roex filter to Sondhi and Guttman's data at a CF of 500 Hz, and the dashed line shows predicted values based on a roex fit at a CF of 250 Hz.

Previous studies of the monaural filter have favoured a roex shape in their measurements, such as Glasberg and Moore (1990), who obtained ERB values of 67 Hz at a centre frequency of 250 Hz, and 94 Hz at a centre frequency of 500 Hz. The ERB values of the roex filters fitted to the data of Sondhi and Guttman (1966) were 55 Hz at a centre frequency of 250 Hz, and 146 Hz at a centre frequency of 500 Hz

(see Table 5.1), suggesting that the binaural auditory filter is wider than the monaural filter at a centre frequency of 500 Hz, but not at 250 Hz.

Kohlrausch (1988) gave listeners a task in which they had to detect a signal presented either monaurally (Sm), binaurally in-phase (So), or binaurally out-of-phase (Sπ), within a broadband masker that was in-phase below 500 Hz and out-of-phase at higher frequencies, denoted NoπSπ (or vice versa, NπoSπ). Signal thresholds were measured at frequencies between 200 and 800 Hz for a range of binaural configurations (see Fig. 5.4).

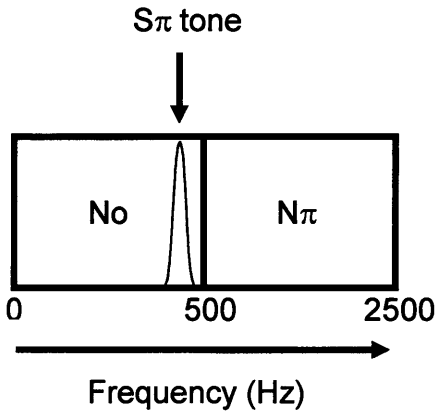

FIG. 5.4. A schematic example of the stimuli used by Kohlrausch (1988). In this example, the stimulus configuration is NoπSπ.

Kohlrausch used this design in order to measure the shape of the binaural auditory filter, as it could be calculated from the frequency-dependent values of binaural masked thresholds around the phase transition frequency. The best fit to the data was found to be made by a trapezoidal filter with an ERB of 80-84 Hz. These values contrast with the wider bandwidth of the Gaussian filter (182 Hz) that produced the best fit (i.e. lowest χ^2 value, see Table 5.1) to the data obtained by Sondhi and Guttman (1966) at a centre frequency of 500 Hz (see Table 5.1 and Fig. 5.2), and

would seem to suggest that binaural and monaural filter widths are comparable. Table 5.3 summarises findings concerning auditory filter bandwidth in previous studies.

5.1 EXPERIMENT 7. BANDWIDTH.

5.1.1 Introduction.

When dichotic broadband noise switches from in-phase to out-of-phase over a narrow frequency region, a percept called the binaural edge pitch (BEP) is created (Klein and Hartmann, 1981). This is a sensation of pitch at the frequency where the phase transition occurs, similar to the Huggins pitch (Cramer and Huggins, 1958), and is strongest for frequency regions between 350 and 800 Hz (Klein and Hartmann, 1981). The presence of the BEP in a tone-detection task might affect the level of masked thresholds, as, perceptually, listeners are experiencing two tones and have to distinguish between them. Thus, the presence of a BEP could have a distracting effect on the listener similar to the 'auditory events' described in chapter 4, which occurred when in-phase noise was presented contiguous to delayed noise, causing the spatial extent of the percept to fluctuate.

In the Sondhi and Guttman (1966) study, the transitions in the masker from Nπ to No and back to Nπ create two BEPs, which could have affected listeners' judgements. In their paper, Sondhi and Guttman recognised the presence of a 'noise artefact' similar to the Huggins pitch when the inner band of their stimulus was narrow. They attempted to ameliorate the impact of the artefact by pulsing the signal at 2 Hz, the intention being to allow the listener to distinguish the signal from the artefact.

A single BEP was also present in the stimuli presented by Kohlrausch (1988), at the transition at 500 Hz from in-phase to out-of-phase masking noise. An additional attribute of the stimuli employed by Kohlrausch was that it provided listeners with the

possibility of using an off-frequency listening strategy (Patterson, 1976). If the listener is able to use a range of filters with centre frequencies close to that of the signal, a better signal-to-noise ratio may be found by listening to a filter centred just above or below the signal frequency, thus producing lower thresholds and narrower filter bandwidth estimates (Patterson, 1976). As the fitting procedure did not explicitly model off-frequency listening, the procedure used by Kohlrausch may have underestimated the ERB. Off-frequency listening was accounted for in the Sondhi and Guttman (1966) study, as the phase-inversion of noise at frequencies above and below the centre frequency (in relation to the noise band surrounding the centre frequency) prevented the listener gaining any substantial advantage in signal-to-noise ratio by listening to an off-frequency filter.

The present experiment used a design that closely matched the study of Sondhi and Guttman (1966), but replaces the $N\pi$ sections of the masker with Nu noise in order to remove the BEP pitches from the stimulus (see Fig. 5.5).

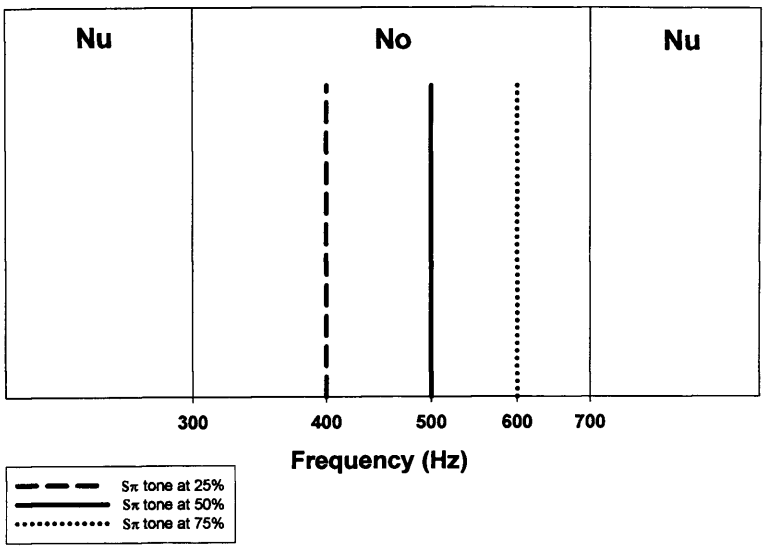

FIG. 5.5. Schematic illustration of the stimulus used in experiment 7. In this example, the bandwidth of the No noise is 400 Hz, and the centre frequency is 500 Hz. For illustrative purposes, the three tone positions are shown on the same stimulus. The x axis is plotted logarithmically and to scale.

As well as accounting for any confounds due to BEPs, the stimulus design limits the extent to which listeners can use off-frequency listening to improve their thresholds. Positioning their filter at slightly lower or higher frequencies than the signal will not produce a substantially higher proportion of No noise in the window, due to the presence of uncorrelated interfering noise at both higher and lower frequencies in the spectrum. A minor disadvantage of this stimulus configuration is that the maximum masked level difference is reduced; at a signal frequency of 500 Hz, NuSπ vs NoSπ produces a BMLD of 10-12 dB, whilst NπSπ vs NoSπ gives a BMLD of 15 dB (Robinson and Jeffress, 1963). As well as measuring the ERB, the shape of the filter can also be described using this method. By taking thresholds at the 25, 50, and 75% points of the No portion of the masker, the amount of variation in the level of binaural

unmasking at these positions will reveal any asymmetry in the filter. Binaural auditory filters were measured at centre frequencies of 250, 500, and 750 Hz, in order to examine how the ERB changed with centre frequency, and to compare the results to previous experiments that have examined binaural filter bandwidths (Sondhi and Guttman, 1966, centre frequencies at 250 Hz and 500 Hz; Kohlrausch, 1988, a centre frequency of 500 Hz).

5.1.2 Method.

Participants

Three participants, all previously trained in psychophysical experiments, took part in the experiment. One was male and two female aged between 18 and 25. They were paid upon completion.

Apparatus/Materials

Participants were given a 2I-2AFC task. To create the stimuli, two independent noises (N_1 and N_2) were generated digitally using Matlab. The correlated band was generated by band-pass filtering N_1, and presenting the filtered waveform in both channels. To produce the uncorrelated noise below and above the No band, two independent noises were band-pass filtered rectangularly with an FFT between 1 Hz and the high-pass cut-off of the No band, and between the low-pass cut-off of the No band and 3 kHz. The two noises were subsequently presented to separate channels. In the target interval, a tone was generated at the appropriate frequency and presented interaurally out of phase. Both the noise and tone had gated onset and offset; a 20-ms ramp for the tone, and a 10-ms ramp for the noise. The duration of the tone and noise

was 0.5 s, and the interstimulus interval was 0.5 s. The order of target and distracter was randomised for each trial.

Design

The experiment had a repeated measures design. The first independent variable was the bandwidth of the No noise (0, 100, 141.42, 200, 282.84, 400 Hz). The range of bandwidths was chosen following pilot testing to encompass the dynamic range over which the BMLD at all three noise positions was observed to change. The second and third independent variables were centre frequency (250, 500 and 750 Hz), and tone position (25%, 50%, and 75% points of the No band). The dependent variable was signal magnitude measured in dB.

Procedure

The listener's task was to identify the sound interval containing the tone. The target interval was presented randomly in either interval, and feedback was provided. The magnitude of the tone was varied adaptively in order to obtain a 70.7% correct estimate (Levitt, 1971). The initial level of the tone was chosen after pilot testing in order to be clearly audible for the first steps of the adaptive track. Initially, the step size of the adaptive track was set to 4 dB, and was reduced to 2 dB following two reversals. The last 10 reversals comprised the measurement phase, and the average of the reversals within the measurement phase was taken as threshold. If the listener made a reversal within the first ten trials, data from the run was excluded and the run repeated[2]. After the stimuli were presented, listeners were presented with a Matlab

[2] As discussed in chapter 3, a reversal made early on in an adaptive track can have a disproportionate effect on the final threshold, as the step size is reduced after the first reversal, making it harder for the listener to obtain a low threshold. 10 trials was chosen as a cut-off point because the tone is clearly

figure with two buttons corresponding to the two sound intervals, which allowed them to respond with the mouse. Thresholds were taken sequentially at each bandwidth from 0-400 Hz. The experiment was arranged in blocks so that participants finished taking thresholds at 500 Hz centre frequency before taking part in the 750 and 250 Hz centre frequency conditions. Each participant provided thresholds at 3 tone positions x 5 bandwidths + 1 bandwidth at an No bandwidth of zero[3] = 16 conditions. There were 4 repetitions of each condition, making 64 thresholds in all.

audible in the first 10 trials in each condition, thus any reversal is almost certainly attributable to the listener accidentally pressing the wrong response key.

[3] When the No bandwidth is zero, signals presented at the three tone positions are all at the same frequency, thus only one set of four repetitions was taken in these conditions.

5.1.3 Results.

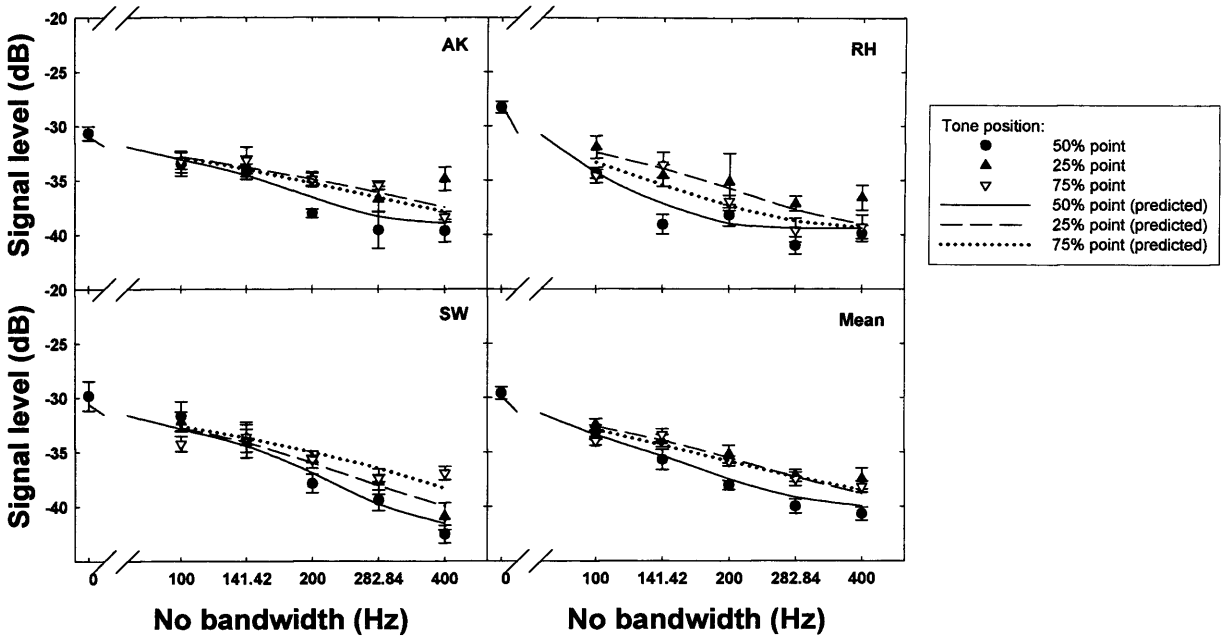

FIG. 5.6. Individual and mean threshold data at a centre frequency of 250 Hz. Thresholds are plotted in relation to an arbitrary zero. Symbols show observed mean thresholds. Filled circles indicate thresholds for tones placed symmetrically within the No band (50% point), upwards pointing filled triangles show thresholds for tones at the 25% point, and downwards pointing open triangles show thresholds for tones at the 75% point. Thresholds predicted by best-fitting windows for the 50%, 25% and 75% points are denoted by the solid, dashed, and dotted lines respectively. Mean predicted thresholds were plotted by averaging the three listeners' predicted thresholds. The x axis is plotted logarithmically.

Figs. 5.6 shows individual and mean thresholds at a centre frequency of 250 Hz at various No bandwidths for the three tone positions. The three listeners produced a broadly similar pattern of results. At all tone positions, thresholds fall as No bandwidth increases. Lowest thresholds were obtained for the 50% noise position,

where thresholds decreased by 11.1 dB between No bandwidths of zero and 400 Hz. Overall improvement at the other tone positions was less than at the 50% point: 7.9 dB for the 25% point and 8.6 dB at the 75% point.

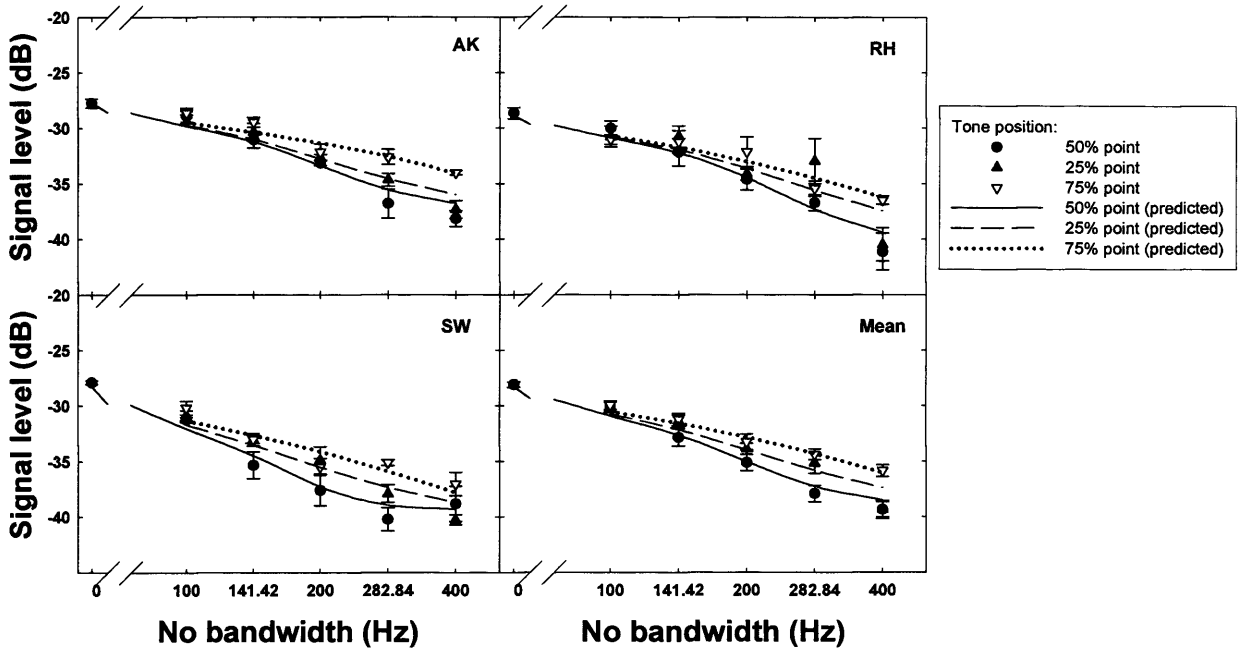

FIG. 5.7. As for Fig. 5.6, but for a centre frequency of 500 Hz.

Fig 5.7 shows thresholds at a centre frequency of 500 Hz. At all tone positions, thresholds decrease as No bandwidth increases. Unlike the pattern of results at a centre frequency of 250 Hz, overall threshold improvement occurred at both the 25% and 50% points: 11.3 dB at the 50% point, and 11.2 dB at the 25% point. Thresholds at the 75% tone position were found to decrease least as No bandwidth increased; overall improvement was 7.8 dB.

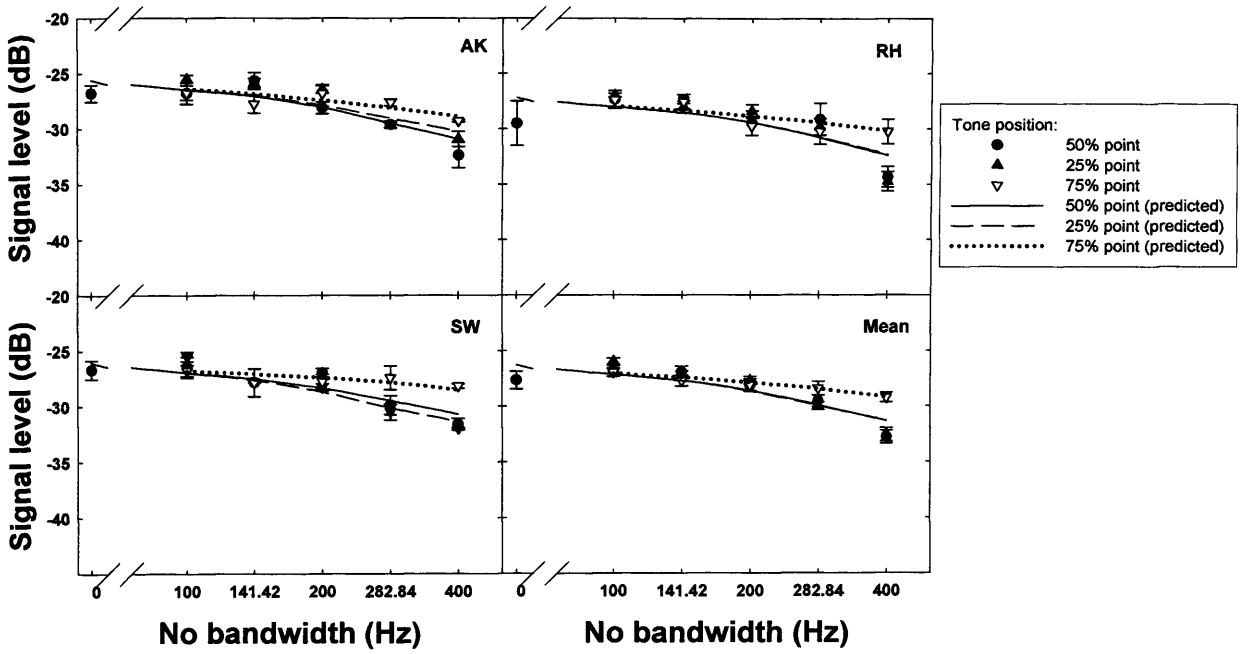

FIG. 5.8. As for Fig. 5.6, but for a centre frequency of 750 Hz.

The pattern of threshold change is similar at a centre frequency of 750 Hz to that at 500 Hz (compare Fig. 5.7 and Fig. 5.8). Overall improvement at the 25% point (4.9 dB) is comparable to that at the 50% point (5.1 dB), and worse at the 75% point (1.6 dB).

		Fitting function (no skirt)											
Listener	Signal	Exponential			Goodness-	Gaussian			Goodness-	Rounded Exponential			Goodness-
	frequency (Hz)	tpl(Hz)	tpu(Hz)	ERB	of-fit (χ^2)	tpl(Hz)	tpu(Hz)	ERB	of-fit (χ^2)	tpl(Hz)	tpu(Hz)	ERB	of-fit (χ^2)
AK	250	53.32	48.09	101.41	1.42	48.64	44.68	165.41	1.12	72.14	65.90	138.04	1.27
	500	45.86	66.92	112.78	1.48	43.23	59.57	182.21	0.74	63.42	89.99	153.42	1.11
	750	67.74	91.84	159.58	1.34	59.46	81.22	249.35	0.97	90.63	123.57	214.20	1.16
RH	250	38.40	28.37	66.77	1.62	31.06	25.34	99.97	1.54	49.85	38.49	88.34	1.55
	500	59.96	71.20	131.17	2.50	55.91	65.52	215.23	1.83	82.53	97.00	179.53	2.14
	750	68.35	114.02	182.37	2.46	60.53	100.14	284.79	1.97	91.51	152.04	243.55	2.23
SW	250	53.43	66.09	119.52	1.36	50.00	61.66	197.90	0.91	73.60	90.69	164.29	1.11
	500	38.73	49.72	88.45	1.53	35.34	42.73	138.37	0.93	52.65	65.48	118.14	1.22
	750	40.34	103.87	144.21	0.88	39.14	90.26	229.36	0.59	56.60	139.05	195.65	0.74

		Fitting function (skirt functions)																	
Listener	Signal	Exponential					Goodness-	Gaussian						Goodness-	Rounded Exponential					Goodness-
	frequency (Hz)	tpl(Hz)	tpu(Hz)	tsl(Hz)	w(dB)	ERB	of-fit (χ^2)	tpl(Hz)	tpu(Hz)	tsl(Hz)	w(dB)	ERB	of-fit (χ^2)	tpl(Hz)	tpu(Hz)	tsl(Hz)	w(dB)	ERB	of-fit (χ^2)
AK	250	53.31	48.06	170.88	-98.83	101.36	1.53	48.61	44.63	245.81	-98.60	165.26	1.22	72.22	65.84	113.40	-98.77	138.06	1.38
	500	45.85	66.85	80.09	-99.52	112.70	1.60	43.19	59.60	131.27	-100.34	182.20	0.81	63.42	89.88	127.51	-97.64	153.30	1.21
	750	67.96	91.85	138.01	-99.07	159.81	1.45	59.38	81.17	107.45	-96.52	249.11	1.05	90.66	123.57	226.50	-98.73	214.23	1.25
RH	250	37.71	27.77	56.87	-32.37	65.50	1.76	24.10	24.91	48.20	-5.69	98.67	1.53	37.69	37.17	75.39	-5.88	84.76	1.65
	500	59.97	71.18	122.02	-98.22	131.15	2.71	55.95	65.48	101.39	-101.43	215.23	1.99	82.42	96.82	164.84	-99.34	179.24	2.32
	750	68.18	113.78	136.99	-99.67	181.95	2.66	60.60	100.05	122.11	-98.21	284.74	2.13	91.98	151.93	318.68	-98.92	243.91	2.41
SW	250	53.47	66.03	106.14	-99.56	119.50	1.48	49.97	61.59	87.60	-98.67	197.73	0.98	73.53	90.60	142.03	-97.37	164.12	1.21
	500	38.75	49.68	154.50	-100.79	88.44	1.66	35.33	42.74	184.51	-97.62	138.36	1.00	52.65	65.56	151.72	-100.70	118.21	1.32
	750	40.41	103.90	80.68	-99.78	144.31	0.95	39.11	90.26	148.95	-100.06	229.30	0.64	56.73	138.82	201.70	-97.39	195.55	0.80

Table 5.2. Fitting parameters for individual participants, for each of the three centre frequencies and window shapes. *tpl* is the parameter that describes the lower frequency lobe of the filter peak, *tpu* describes the higher frequency lobe, *tsl* is the skirt parameter, and *w* is the weight applied to the skirt of filters that have a skirt fitting function. The ERB is the integral in Hz of the entire filter. The goodness of fit (χ^2) is expressed for each individual window fit, where lower values of χ^2 indicate more efficient fits.

Simple and skirt asymmetric roex, exponential and Gaussian filters were fitted to the data. The filter fits to the data are described in Table 5.2, and plotted in Fig. 5.9. A simple asymmetric Gaussian shape was generally found to give the best fit to the data, and the equivalent rectangular bandwidth of the filter was found to increase with centre frequency. When skirt fits were modelled, the weighting of the skirt was very low in the majority of cases, showing that the skirt was exerting almost no influence on the fit. When described on a linear scale, at a CF of 500 Hz, the filters were

approximately asymmetric. Asymmetry became more evident at CFs of 500 and 750 Hz, with narrower lower lobes than upper lobes, and asymmetry increasing with centre frequency (see Table 5.2). The asymmetry was especially marked for all listeners at a centre frequency of 750 Hz. There were two exceptions to this pattern: at a centre frequency of 250 Hz, a very slightly better fit was observed for a Gaussian skirt function for listener RH, and listener SW demonstrated an ERB of 138.37 Hz at 500 Hz, which was lower than her ERB at 250 Hz (197.90 Hz). A within-subjects ANOVA demonstrated significant main effects of centre frequency ($F(2,22)=277.51$, $p<0.001$), tone position ($F(2,22)=26.51$, $p<0.001$) and No bandwidth ($F(4,44)=223.46$, $p<0.001$). Significant interactions were demonstrated between centre frequency and tone position ($F(4,44)=7.041$, $p<0.001$), centre frequency and No bandwidth ($F(8,88)=5.98$, $p<0.001$), tone position and No bandwidth ($F(8,88)=5.62$, $p<0.001$), and between centre frequency, tone position and No bandwidth ($F(16,176)=1.90$, $p<0.05$). N.B. datapoints at an No bandwidth of zero were not included in the analysis, as only one set of points was included for the three tone positions.

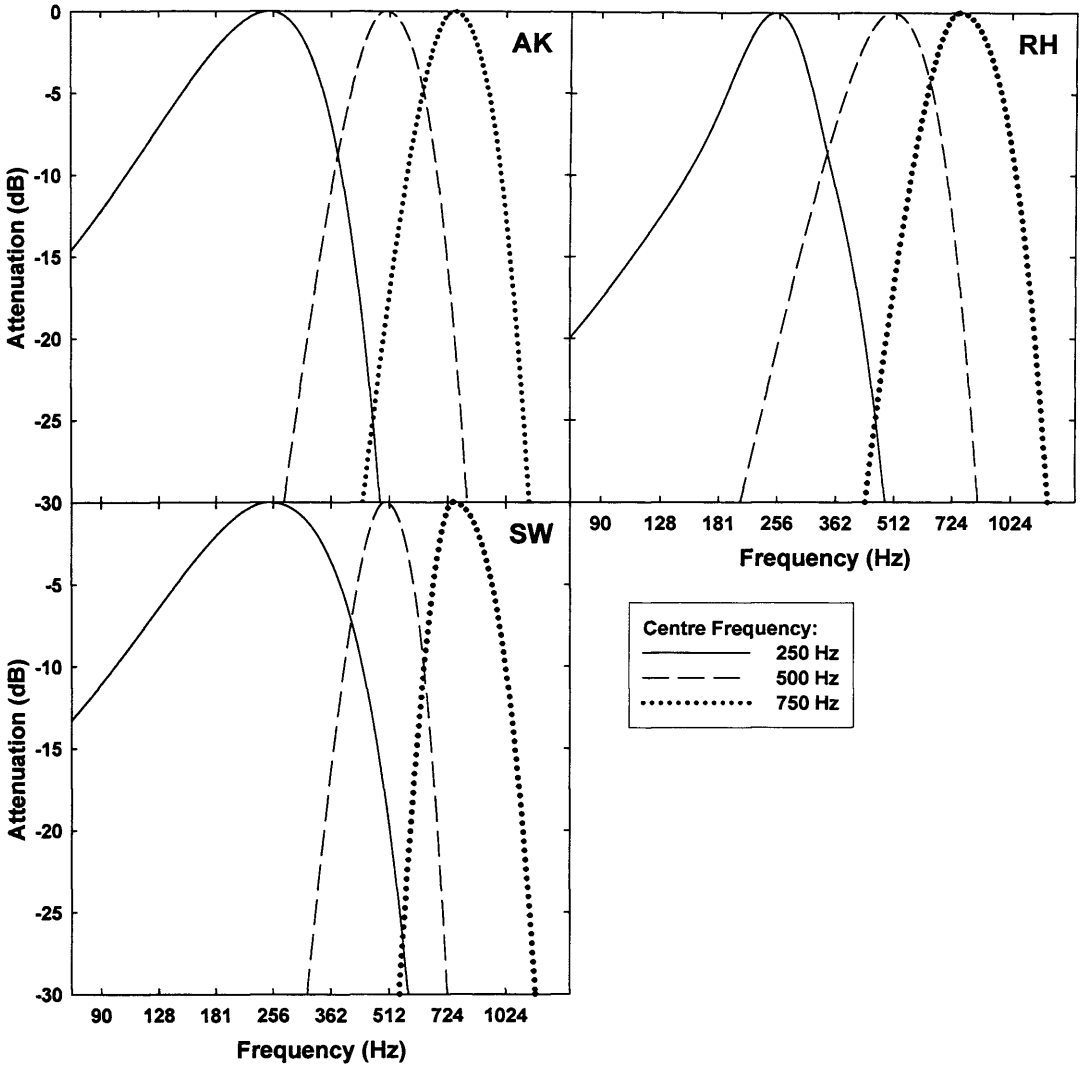

FIG. 5.9. Attenuation characteristics of binaural filters for individual listeners at centre frequencies of 250, 500 and 750 Hz. The x axis is plotted logarithmically.

Fig. 5.9. shows plots of the best fitting filter shapes for individual datasets at the three centre frequencies. When plotted on a logarithmic scale, the filters were asymmetric with wider lower lobes than upper lobes, and asymmetry decreasing with centre frequency. The opposite pattern of asymmetry is observed when described on a linear scale (see Table 5.2).

5.1.4 Discussion.

The binaural filter bandwidth was found to increase with centre frequency, and was found to be wider than comparable monaural bandwidths. Glasberg and Moore (2000) obtained monaural roex filters with ERB values of 67 ms at a centre frequency of 250 Hz, and 94 ms at a centre frequency of 500 Hz, and Moore, Peters, and Glasberg (1990), again using roex filter shapes, obtained an ERB value of 147 Hz at a centre frequency of 800 Hz. Roex fits to the data in the current study displayed ERB values ranging from 88-164 Hz at a centre frequency of 250 Hz, 118-180 Hz at a centre frequency of 500 Hz, and 196-244 Hz at a centre frequency of 750 Hz (see Table 5.2). These findings are summarised in Table 5.3.

		Centre Frequency (Hz)		
Study	Type	250	500	750
Glasberg and Moore (2000)	Monaural	67	94	N/A
Moore et al. (1990)	Monaural	N/A	N/A	147
Sondhi and Guttman (1966)	Binaural	68	182	N/A
Kohlrausch (1988)	Binaural	N/A	80-84	N/A
Current study	Binaural	88-164	118-180	196-244

Table 5.3. Auditory filter bandwidth (Hz) derived by fitting roex filters to monaural and/or binaural data from a range of studies. Note the following exceptions: Moore et al.'s result is obtained using a centre frequency of 800 Hz, and Kohlrausch's result was obtained using a trapezoidal filter. The fits to Sondhi and Guttman were modelled in the current study (see Table 5.1).

The best filter fit to the data in the current study at all centre frequencies was generally obtained using a simple asymmetric Gaussian fit, with (on a linear scale) narrow lower lobes and wide upper lobes (see Table 5.2). At a centre frequency of 250 Hz, ERB values ranged from 99 to 198 Hz. At a centre frequency of 500 Hz, the ERB ranged from 138 to 215 Hz, and at a centre frequency of 750 Hz, the ERB

ranged from 229 to 285 Hz. The asymmetry of the filters was found to increase with centre frequency. On a logarithmic frequency axis, the upper lobe tends to be steeper than the lower lobe (see Fig. 5.9), an observation also seen in measurements of the monaural filter (Moore et al., 1990). The data is in quantitative agreement with the binaural bandwidth obtained by Sondhi and Guttman (1966) at a centre frequency of 500 Hz, which when modelled demonstrates a best fit with a simple Gaussian filter with an ERB of 182 Hz, but not at 250 Hz, where a simple Gaussian filter with a narrower ERB value of 69 Hz provided the best fit (see Table 5.1). The findings in this study depart from the findings of Kohlrausch (1988), who obtained much narrower bandwidths with a trapezoidal filter (80-84 Hz at a centre frequency of 500 Hz).

Figures 5.6, 5.7 and 5.8 demonstrate that the threshold change observed across datasets is in accordance with the auditory filter model proposed by Fletcher (1940). Minimal binaural unmasking occurs when the bandwidth of the No band is zero (equivalent to an NuSπ configuration), where the highest thresholds are observed. When the No bandwidth is narrow, thresholds at the three noise positions are approximately equal and the magnitude of the unmasking relatively small, as only a small portion of the filter will contain No noise. As the No bandwidth increases, thresholds at all three tone positions decrease as more No noise enters the filter, leading to greater unmasking. Thresholds are lowest for the 50% tone position because it is assumed that the listener will position the centroid of their filter at the midpoint of the No band in order to obtain the best signal-to-noise ratio. The filter will attenuate binaural information at frequencies remote from the centre frequency, leading to less unmasking at the 25% and 75% tone positions. At large bandwidths, thresholds are lower at the 25% tone position than the 75% position at centre

frequencies of 500 and 750 Hz (see Figs. 5.7 and 5.8), which, when modelled, result in asymmetry in the filter. It is noted that this pattern is not observed for listeners AK and RH at a centre frequency of 250 Hz (see Fig. 5.6). Although listener SW displays thresholds in a pattern consistent with those observed at higher frequencies, listeners AK and RH display lower thresholds at the 75% point than at the 25% point at the widest No bandwidths. As a result, AK and RH's filters are roughly symmetric on a linear scale at a centre frequency of 250 Hz, whereas SW's filter is asymmetric (see Table 5.2). The pattern of thresholds observed at the 25% and 75% points of the No bandwidth may be explained in terms of changes in the overall BMLD across frequency.

The overall BMLD falls as centre frequency is increased (see Figs. 5.6-5.8). This is because the magnitude of the BMLD is greatest for the frequency region around 200 Hz, and decreases slowly as centre frequency is increased. Overall BMLD is observed to fall rapidly as centre frequency is decreased below approximately 200 Hz (Hirsh, 1948). Equal BMLD magnitude at each tone position was assumed in the modelling procedure. The alteration in magnitude may have affected measurement of the asymmetry of the filter, as the magnitude of the BMLD at the 25 and 75% points will be slightly different to that at the 50% point. The change in BMLD magnitude would be minimal at small No bandwidths, but at the larger bandwidths the magnitude change may begin to have an effect. At an No bandwidth of 400 Hz, the 25% tone position would be held at 150 Hz, a frequency region where the magnitude of the BMLD is reduced. In comparison, the 75% tone position would be held at 350 Hz, in a frequency region where the magnitude of the BMLD is largest. This could account for the elevated thresholds observed for listeners AK and RH at the 25% tone position at a centre frequency of 250 Hz (see Fig. 5.6).

The change in BMLD magnitude could be modelled using EC theory (Durlach, 1972), or accounted for by keeping the tone position at a constant frequency, and moving the No band position relative to the tone position, so that although the signal was kept at a constant frequency, it would occur at the 25, 50, and 75% point of the No band, and the magnitude of maximum BMLD change would be kept constant.

Though the findings of this study and the bandwidths reported by Sondhi and Guttman (1966) would seem to indicate that binaural auditory filters are wider than their monaural counterparts, the viewpoint that the binaural filter bandwidth is wider than the monaural bandwidth has been challenged (e.g. Hall, Tyler, and Fernandes, 1983; van der Par and Kohlrausch, 1999). The prevailing view is that binaural processing is constrained by the same peripheral processes as in the monaural case, but binaural processing is influenced by a wider range of frequencies that includes information from filters adjacent to the filter centred at the signal frequency, leading to a wider operational bandwidth depending on the spectral content of the masker (Bernstein, Trahiotis and Freyman, 2006).

van de Par and Kohlrausch (1999) assumed that the binaural system attends to one or more auditory filters depending on which arrangement results in optimal detection. In narrow-band noise, both central and adjacent filters provide binaural information that can be combined for signal detection, improving the detection threshold given that the internal noise of each filter is independent. Combining binaural information from several filters reduces the internal noise compared to using only the central filter (van der Heijden and Trahiotis, 1998). This can be contrasted with the situation when broadband noise is employed. In this situation, only the central filter provides useful binaural information and the magnitude of the internal noise would be higher. This reasoning was extended to the experiment conducted by Kohlrausch (1988), which

measured thresholds for signals presented within broadband masking noise with a correlation of 1 below 500 Hz and a correlation of -1 at the higher frequencies. Van de Par and Kohlrausch (1999) argued that in the study of Kohlrausch (1988), only the single auditory filter centred at the signal frequency provided useful information for the binaural processor, and integration across other filters could not create apparently larger binaural auditory filters. The binaural filter measured in that study had an ERB comparable to that of the monaural filter, providing support to the notion that binaural frequency selectivity is constrained by the same peripheral processes as monaural listening.

However, the same argument can be applied all experiments that use analogues of the notched-noise technique and broadband masking noise, including that of Sondhi and Guttman (1966) and the present experiment. Because the notched-noise technique utilises broadband masking noise across the spectrum, independent of the specific condition, only the filter centred at the test frequency provides useful binaural information and combining information across filters reduces sensitivity. The results of Sondhi and Guttman (1966) at a centre frequency of 500 Hz, and the present study at all centre frequencies thus suggests that the binaural auditory filter is wider than its monaural counterpart, rather than manifesting itself as a wider operational binaural bandwidth stemming from across-frequency integration of monaural filters.

Findings from the study indicated that the ERB of the binaural auditory filter is dependent on centre frequency, and ranges from 88-164 Hz at a centre frequency of 250 Hz, from 118-180 Hz at a centre frequency of 500 Hz, and 196-244 Hz at a centre frequency of 750 Hz. These ERB estimates are considerably larger than comparable monaural filters. The filter was best fitted by a simple function with a rounded peak

(e.g. Gaussian). No firm conclusion concerning the asymmetry of the auditory filter was drawn from this study, due to the problem of BMLD change across frequency.

CHAPTER 6. GENERAL DISCUSSION

The experiments conducted throughout this thesis attempted to measure the temporal and spectral resolution of the binaural system. Previous studies that investigated temporal resolution have assumed that interfering noise dilutes delayed noise within the temporal window. Two experiments described in this thesis have demonstrated that the dilution concept is valid for correlated interfering noise, but not for uncorrelated interfering noise. A study by Bernstein et al. (2001) suggested that the binaural temporal window is considerably smaller than estimates made in previous studies (e.g. Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998). Results from this thesis are different to those of Bernstein et al., and suggest that several factors led to their findings, in particular the lack of control over the coherence of the stimulus due to the use of a detection task, the overall duration of the stimuli, and the use of a correlated interferer.

In the frequency domain, a study by Sondhi and Guttman (1966) that investigated the frequency selectivity of the binaural system found evidence that suggested that binaural auditory filters are substantially wider than monaural auditory filters. Conversely, Kohlrausch (1988) measured auditory filters that were comparable to monaural filters. The results from the experiment conducted in this thesis found that binaural auditory filters are substantially wider than monaural auditory filters. These findings are summarised in section 6.1. Possible further research is described in section 6.2, in which the thesis is brought to a close.

6.1 SUMMARY OF RESULTS.

Chapters 3 and 4 described eight experiments that investigated temporal resolution within the binaural system, and chapter 5 described one experiment designed to measure binaural auditory filter bandwidth. The findings are summarised below.

6.1.1 The effects of delay masking with different interferers.

As a precursor to measuring the temporal window, the effect of the presence of interfering noise within the temporal window on the perception of interaural delay needed to be established. Experiments 1 and 2 contrasted threshold ITD change in the presence of interfering diotic and uncorrelated noise respectively.

Experiment 1 measured ITD discrimination threshold for broadband delayed noise in the presence of interaurally correlated interfering noise. Six listeners performed an adaptive 2I-2AFC discrimination task with 100-ms stimuli composed of delayed and diotic noises bandlimited between 0 and 3 kHz mixed in various proportions (1, 0.75, 0.563, 0.422, 0.316, 0.237, and 0.178). A regression analysis showed that thresholds doubled for every halving of the proportion of delayed noise power in the stimulus, consistent with the hypothesis that the diotic noise was 'diluting' the ITD (i.e. the linear (1:1) dilution hypothesis).

Experiment 2 measured ITD discrimination thresholds for a range of interaural coherences obtained by mixing correlated and uncorrelated noise before applying the delay. Three listeners performed 2I-2AFC discrimination tasks at coherences of 1, 0.75, 0.5, 0.25, 0.2, 0.15, 0.1, and 0.05, and durations of 100, 500, and 1000 ms. Psychometric functions were used to obtain thresholds by taking the 71% correct point. In contrast to experiment 1, thresholds more than doubled for each halving of the coherence at all three stimuli durations tested.

6.1.2 Temporal resolution.

Sensitivity to small ITDs in detection and discrimination tasks was measured using a lateralization task developed from the design of Bernstein et al. (2001), in an attempt to measure the temporal resolution of the binaural system and explain the short temporal windows obtained in that study.

Experiment 3 was composed of three tasks. The first task consisted of a replication of Bernstein et al. (2001); a 4I-2AFC adaptive detection task. All stimuli consisted of 100 ms noise bursts. Target stimuli consisted of delayed sections of noise temporally fringed with diotic noise. Thresholds were measured for probe durations of 64, 32, 16, 8, 4, or 2 ms. Reference stimuli consisted of diotic noise only. The second task consisted of a 2I-2AFC adaptive detection task. The third task consisted of a 2I-2AFC adaptive discrimination task which controlled for coherence changes across presentation intervals. Half the total delay was imposed on each interval, so that the intervals were lateralized to opposing sides of the head, and the participant's task was to choose the direction of sound movement across the two intervals. The results were contrasted with the 2I-2AFC adaptive detection task, and the Bernstein et al. replication. Thresholds for the discrimination task were higher than for the two detection tasks, especially at short probe durations. Results from the study of Bernstein et al. (2001) were partially replicated, in that a window with a narrow peak and wide skirt was obtained in the majority of cases.

In experiment 4, psychometric functions were obtained from 3 participants for the six probe durations using the same 2I-2AFC detection and discrimination tasks. Binaural temporal windows were fitted to the data using a variety of fitting functions. Fits to the detection task data demonstrated narrow tips but indeterminate skirts. In the

discrimination task neither parameter could be accurately measured, suggesting that the overall stimulus duration was too short to encompass the window.

Experiment 5 increased the stimulus length to 500 ms and used probe durations of 256, 128, 64, 32 and 16 ms. Fits with narrow peaks and indeterminate skirts were again obtained, probably due to the distracting effects of 'auditory events' occurring at the transition points between the masking diotic noise and the probe.

Experiment 6 used 500 ms long stimuli with a discrimination task and uncorrelated interfering noise to ameliorate the distracting effect of the auditory events. Simple fits with an ERD of between 110 and 349 ms were found to produce the best fit to the data, providing supporting evidence for sluggishness in the binaural system.

In conclusion, it was found that the short temporal windows obtained in the study by Bernstein et al. (2001) probably resulted from the design of their experiments. When Bernstein et al. employed No interfering noise, the use of a detection task instead of a discrimination task made available an additional coherence cue that the listeners were able to use to lower their thresholds. When Bernstein et al. employed Nu interfering noise, the assumption was made that the interferer would dilute the ITD in a linear (1:1) manner. This assumption was demonstrated to be inaccurate by the results of experiment 2. Longer stimuli, a discrimination task, and an uncorrelated interferer produced temporal windows with an ERD greater than 100 ms, consistent with the notion that the binaural system is sluggish.

6.1.3 Spectral resolution.

The spectral resolution of the binaural system was measured using a tone detection task in a binaural analogue of the notched-noise technique. Three listeners performed 2I-2AFC tasks with a 500-ms out-of-phase signal within 500-ms of broadband

masking noise consisting of an 'outer' band of uncorrelated noise, and an 'inner' band of diotic noise. Three centre frequencies were tested (250, 500 and 750 Hz), and the asymmetry of the filter measured by placing the tone at the 25, 50, and 75% points of the inner band of No noise. Thresholds were taken for various bandwidths of the No noise (0, 100, 141.42, 200, 282.84, and 400 Hz). A 2-parameter asymmetric Gaussian shape was consistently found to give the best fit to the data. The equivalent rectangular bandwidth of the filter was found to increase with centre frequency, and was consistently larger than monaural bandwidths.

6.2 LIMITATIONS TO THE EXPERIMENTS AND FURTHER RESEARCH.

The studies presented within this thesis have investigated the spectral and temporal resolution of the binaural system. However, there are some limitations to the experiments. A range of questions remain unanswered regarding the resolution of the binaural system, and there remain inconsistencies among the numerous studies that have investigated this area of auditory perception.

Implications and applications of the work presented in the thesis are first described, followed by a section that describes the limitations to some of the existing data concerning temporal resolution. Finally, the results are compared to those of physiological studies, and the thesis brought to a close. Further research that is related to the work presented in this thesis, but beyond the scope of the current investigation, is discussed throughout this section.

6.2.1 Limitations, implications and applications of the results of the experiments conducted within the temporal domain.

In the thesis, the temporal window was measured using a lateralization task. One limitation involved with using this technique is that it is not possible to measure any asymmetry in the temporal window using this method. Previous research has used a tone detection paradigm to measure asymmetry in the temporal window (Culling and Summerfield, 1998). This form of analysis is not possible with a lateralization paradigm.

The results of the experiments conducted within the temporal domain add to the literature that has found evidence that the binaural system is sluggish (e.g. Grantham and Wightman, 1978; Culling and Summerfield, 1998). The sluggishness of the binaural system limits our abilities when we attempt to localise sounds (e.g. Perrott and Pacheco, 1989; Chandler and Grantham, 1992), and may impact on our ability to perceive speech. The implications of binaural sluggishness for speech perception were studied by Culling and Colburn (2000), who investigated the effect of speech rate on speech-reception thresholds (SRTs) in NoSo and NoSπ configurations. The dependent measure was BILD (the difference in SRT between the NoSo and NoSπ configurations), and it was assumed that correlation change detection was the basis of the BILD. It was predicted that high articulation rates would reduce the binaural intelligibility difference, as the binaural system responds sluggishly to changes in interaural correlation caused by fluctuations in speech energy in the NoSπ configuration. SRTs were measured using articulation rates of 1, 1.5 and 2 times the original articulation rate, which was increased using a phase-vocoder technique that increased the modulation frequencies present within the speech but did not alter the pitch. At an articulation rate of 1, BILD thresholds were 5.2 dB lower in the NoSπ

condition than in the NoSo condition. Doubling the articulation rate failed to abolish the binaural advantage, but lowered the BILD to 2.8 dB. The much larger 6-8 dB increase in thresholds for both the NoSo and NoSπ configurations when the articulation rate was increased demonstrated that accelerated speech was more difficult to understand in noise at increased articulation rates regardless of the binaural configuration. It was concluded that information at low-modulation frequencies (below 5 Hz) within the speech signal was sufficient to overcome the effects of binaural sluggishness even at increased articulation rates. These results indicate that binaural sluggishness may have an effect when listeners attempt to temporally resolve fine temporal detail from signals in noise, however binaural sluggishness probably has no impact on speech intelligibility in noise at normal articulation rates, and only a limited impact at increased articulation rates.

6.2.2 Limitations, implications, and applications of the results of the experiment conducted within the frequency domain.

In experiment 7 (as described in the discussion section of that experiment), changes in the magnitude of the BMLD across the frequency spectrum may have adversely influenced the measurement of the asymmetry of the auditory filter. The design of the experiment could be altered so that the tone position was kept at a constant frequency, and the position of the No band moved relative to the tone position. Thus, although the signal was kept at a constant frequency, it would occur at the 25, 50, and 75% point of the No band, keeping the magnitude of the BMLD change constant and accurately measuring the asymmetry of the filter.

The findings that the binaural auditory filter is wider than the monaural auditory filter (experiment 7) may have implications for research concerning auditory

streaming. When rapid sequences of sound are heard, sounds with components in a similar range of frequencies tend to be grouped together, a phenomenon called fusion or coherence where the sounds are perceived to originate from a single source. Sounds with components in different frequency ranges tend to be heard as different streams, and heard as if originating from different sources; a phenomenon called fusion or segregation (Bregman, 1990; Rose and Moore, 1997; 2000). When successive tones, separated in frequency, are presented to listeners and the frequency separation is less than a critical value called the fission boundary, a single stream is heard. This is distinct from the temporal coherence boundary, which is the value of frequency separation above which two streams are heard.

Research has provided support to the proposition that the percept of one or more streams is at least partly dependent on the overlap of the excitation patterns evoked by sounds in the cochlea, and that successive tones are integrated into a single stream if they excite the same peripheral channels (Hartmann and Johnson, 1991). The frequency difference between successive tones at the fission boundary was found to significantly increase with increasing level, consistent with the broadening of monaural auditory filters with increasing level (Rose and Moore, 2000). As the overlap of the excitation patterns depends on the sharpness of the auditory filters (Moore and Glasberg, 1983), the finding that the binaural auditory filter is wide compared to the monaural auditory filter would predict that the frequency separation between tones at the fission boundary would be wider for tones presented binaurally (e.g. in the NoSπ configuration) compared to tones presented monaurally (e.g. at the same sensation level in NoSo).

It has been argued that reduced frequency selectivity plays a role in hearing impairment, leading to difficulty understanding speech in the presence of background

noise (Plomp, 1994; Moore, 1998). A number of studies have examined speech intelligibility in background noise within this framework. Baer and Moore (1993) used a spectral smearing technique to simulate the effects of broadened auditory filters in hearing-impaired listeners. The results indicated that spectral smearing had little effect on speech intelligibility in quiet listening conditions, but a large effect on speech intelligibility in noise. A similar study of speech intelligibility in the presence of a single interfering talker demonstrated similar results (Baer and Moore, 1994), and was in agreement with a study by ter Keurs, Festen, and Plomp (1993) that also examined the effects of spectral smearing on speech intelligibility. The wide binaural bandwidth measured in experiment 7 for normally hearing listeners may have implications for studies and models of auditory bandwidth measurement for both normal hearing and hearing-impaired listeners in binaural conditions. However, it should be noted that reduced frequency selectivity has not been conclusively demonstrated to be responsible for reduced speech perception in noise among listeners with sensorineural hearing loss, and is likely to be only one of several factors that contribute.

6.2.3 Limitations to the literature regarding temporal resolution, and unanswered questions regarding the temporal window model.

Several studies that have investigated the temporal resolution of the binaural system have used detection tasks in their design and found evidence suggesting a lack of sluggishness. The results of Bernstein et al. (2001) have already been discussed as probably stemming from the use of a detection task that allows listeners to use cues other than fine temporal resolution (e.g. coherence changes across intervals) to increase their performance. The experiments described in chapter 4 suggest that the

use of a discrimination task controls for these changes, and accounts for the lack of sluggishness found in these studies. Similarly, the results of Witton et al. (2000; 2003) were discussed in the General Introduction as probably stemming from the use of a detection task. More recently, Siveke, Ewert and Wiegrebe (2006) tested listeners' sensitivity to sinusoidally amplitude modulated (SAM) noise and two types of noise stimuli with sinusoidal correlation modulation (SCM). An adaptive 2AFC detection task was given to the listeners, and the dependent variable was uncorrelated noise added to the SAM or SCM stimulus. No evidence of binaural sluggishness was found in the case for one of the types of SCM stimulus. Once again, it is possible that the use of a detection paradigm may have allowed the listeners to use alternative cues such as width or loudness cues instead of laterality. Future research investigating temporal resolution involving ITD as the dependent measure should employ a discrimination task rather than a detection task to prevent listeners using cues not anticipated by the investigators.

Although most estimates of binaural integration times are within the same order of magnitude, the range of estimates across studies is wide (Holube et al., 1998). It is possible that there is not a single time constant that describes the binaural system's ability to process dynamic changes in interaural parameters (Kollmeier and Gilkey, 1990), and the wide range of binaural time constants found across experiments may be attributable to different experiments probing different binaural abilities, such as localization (Blauert, 1968), lateralization (Blauert, 1972; Grantham and Wightman, 1978; Pollack, 1978; Grantham, 1982), and detection (Grantham and Wightman, 1979; Kohlrausch, 1986). Task dependent factors may act jointly with the temporal window, or determine the strategy with which the temporal window is applied (Bernstein et al., 2001). However, it is possible that the wide range of time constants

arises from the different methodologies used in the experiments that measure the time constants, rather than the listener employing a different time constant depending on the task. Different experimental designs may introduce artefacts that increase the time constant, and the magnitude of the time constant is dependent on the shape of the window that is assumed (e.g. a Gaussian fit always gives a larger ERD than an exponential when fitted to the same data).

The action of the temporal window (which averages temporal disparities occurring within the window) has been successful at modelling much of the data obtained in studies of temporal resolution. However, some studies have indicated that the temporal window does not average all portions of the stimuli equally (e.g. Zurek, 1980, 1987; Akeroyd and Bernstein, 2001). Zurek (1980) presented listeners with stimuli composed of a 5-ms burst of probe noise which had either an ITD or an ILD imposed upon it embedded within 50-ms of diotic noise. Thresholds were found to increase markedly when the onset of the probe occurred between 1 and 10 ms after the onset of the diotic noise. From the data, it appeared that sensitivity to the delay of the probe segment was inhibited for a brief period by the leading portion of the diotic noise. Akeroyd and Bernstein (2001) used a similar paradigm to Zurek and observed similar results. They conducted an experiment with four conditions: 'both-fringe', 'forward-only,' 'backward-only,' and 'probe alone.' The both-fringe condition was a replication of Zurek (1980). The forward-only condition was identical to the both fringe condition, but the diotic noise occurring after the probe was removed. The backward-only condition removed the leading portion of diotic noise, and a condition where only the probe was presented was employed. The duration of the forward and backward fringes were varied between 0 and 45 ms. Threshold ITDs and ILDs in the backward-only and forward-only conditions were substantially lower than

corresponding thresholds in the both-fringe condition, however the presence of diotic noise before the probe in the forward-only condition produced a relatively steep increase in thresholds as a function of the duration of the fringe. Attempts to model the data with a temporal window failed to account for the rise in thresholds as the duration of the forward fringe increased up to about 10 ms. However, when the data was modelled using a temporal window that incorporated a weighting function that described a brief 'post-onset' loss of binaural sensitivity, a better fit that conformed to the rise in thresholds was observed. This model constitutes an implementation of the precedence effect described by Zurek (1987), which suppresses that weight of binaural disparities occurring after the onset of the stimulus.

The temporal window model has been successful in accounting for a wide range of data, however some studies have obtained data that are not amenable to explanations based on the temporal window model. McFadden (1966), demonstrated that the BMLD for brief stimuli in which both signal and masker were simultaneously gated ('pulsed' stimuli) was smaller than for a continuous masker, and showed that the masker must be on for at least 500 ms prior to the signal for the maximum BMLD to be obtained, the 'forward fringe' effect. This finding is inconsistent with the predictions of the temporal window model, which would predict that the introduction of interfering noise within the window in the continuous masking condition would result in a smaller BMLD. Trahiotis, Dolan, and Miller (1972) confirmed the effect and found that the forward temporal fringe had a greater impact than the backward fringe. Yost (1985) also confirmed that the BMLD was lower when signal and masker were both pulsed compared to conditions in which the masker was continuous or a forward fringe of masking noise was present, and also showed that differences between the fringe and masker resulted in decreased signal detection. For example,

while an No fringe was beneficial to the listener, an Nπ fringe had no effect on thresholds. These findings are difficult to interpret in terms of a temporal window model.

A possible explanation for Yost's results may stem from evidence that the binaural system can be 'reset'. When listeners are presented with trains of high-frequency clicks, their ability to make lateralization judgements based on ITD or level differences declines following the signal's onset (Hafter and Dye, 1983). Hafter and Buell (1990) demonstrated that a temporal gap inserted within the series of clicks presented at high rate was sufficient to restart the adapted binaural system, which had adapted to the acoustic environment. The stimuli used by Yost (1985) contained fringes that were gated off using a 5-ms linear ramp before the start of the onset of the masker. This resulted in a temporal gap with an ERD of 6 ms, which may have reset the binaural system, thus preventing an Nπ fringe from having a masking effect. However, this does not explain why an No fringe was beneficial to the listener. If the binaural system can be reset, the resetting would have to be incorporated into the temporal window model, and limits the applicability of the temporal window obtained in experiment 6 to stimuli that do not contain onsets. Further research could be conducted as to what constitutes an onset and how long the duration of the temporal gap needs to be in order to restart the binaural system. The onset detection mechanism could be investigated, in order to see whether detection is based on the magnitude of the change in intensity at the temporal gap, or the rate of change of intensity. The detection of intensity changes within the monaural temporal window has been investigated by Plack, Gallun, Hafter, and Raimond (2006), and concluded that the detection mechanism used the maximum magnitude of change as the decision statistic

for both increments and decrements in intensity. A similar study could be conducted for binaural listening.

In conclusion, the experiments conducted within this thesis contribute to the considerable literature concerning auditory temporal and spectral resolution. They have demonstrated that the presence of interfering diotic noise dilutes the ITD of a given stimulus in a linear (1:1) fashion, and that the presence of uncorrelated interfering noise is more debilitating for the listener than a diotic interferer. The experiments support existing research that has provided evidence that the binaural system is sluggish (e.g. Grantham and Wightman, 1978; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998), and that the binaural auditory filter is wider than the monaural auditory filter (Sondhi and Guttman, 1966).

References.

Akeroyd, M. A., and Bernstein, L. R. (2001). The variation across time of sensitivity to interaural disparities: Behavioural measurements and quantitative analyses. J. Acoust. Soc. Am., 110, 2516-2526.

Akeroyd, M. A., and Summerfield, A. Q. (1999). A binaural analog of gap detection. J. Acoust. Soc. Am., 105, 2807-2820.

Amitay, S., Irwin., A., Hawkey, D. J. C., Cowan, J. A., and Moore, D. R. (2006). A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naïve listeners. J. Acoust. Soc. Am., 119, 1616-1625.

Baer, T., and Moore, B. C. J. (1993). Effects of spectral smearing on the intelligibility of sentences in the presence of noise. J. Acoust. Soc. Am., 94, 1229-1241.

Baer, T., and Moore, B. C. J. (1994). Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. J. Acoust. Soc. Am., 95, 2277-2280.

Batteau, D. W. (1967). The role of the pinna in human sound localization. Phil. Trans. R. Soc. London. Series B, 168, 158-180.

Bernstein, L. R., and Trahiotis, C. (1992). Detection of antiphasic sinusoids added to the envelopes of high-frequency bands of noise. Hearing Research, 62, 157-165.

Bernstein, L. R., Trahiotis, C., Akeroyd, M. A. and Hartung, K. (2001). Sensitivity to brief changes of interaural time and interaural intensity. J. Acoust. Soc. Am., 109, 1604-1615.

Bernstein, L. R., Trahiotis, C., and Freyman, R. L. (2006). Binaural detection of 500-Hz tones in broadband and in narrowband masking noise: Effects of signal/masker duration and forward masking fringes. J. Acoust. Soc. Am., 119, 2981-2993.

Blauert, J. (1968). Ein Beitrag zur Tragheit des Richtungshorens in der Horizontalebene. Acustica, 20, 200-225.

Blauert, J. (1972). On the lag of lateralization caused by interaural time and intensity differences. Audiology, 11, 265-270.

Blauert, J., and Lindemann, W. (1986). Spatial mapping of intracranial auditory events for various degrees of interaural coherence. J. Acoust. Soc. Am., 79, 806-813.

Blodgett, H.C., Wilbanks, W. A., and Jeffress, L. A. (1956). Effect of large interaural time differences upon the judgement of sidedness. J. Acoust. Soc. Am., 28, 639-643.

Bregman, A. S. (1990). Auditory Scene Analysis: The perceptual organization of sound. Cambridge, MA: Bradford books, MIT press.

Burns, E., and Colburn, H. S. (1977). Binaural interaction of modulated noise. J. Acoust. Suppl. 1, 62, S97.

Butler, R. A., Humanski, R. A., and Musicant, A. D. (1990). Binaural and monaural localization of sound in two-dimensional space. Perception, 19, 241-256.

Chandler, D. W., and Grantham, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. J. Acoust. Soc. Am., 91, 1624-1636.

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

Culling, J. F., and Colburn, H. S. (2000). Binaural sluggishness in the perception of tone sequences and speech in noise. J. Acoust. Soc. Am., 107, 517-527.

Culling, J. F., and Summerfield, Q. (1998). Measurement of the binaural temporal window using a detection task. J. Acoust. Soc. Am., 103, 3540-3553.

Culling, J. F., and Edmonds, B. A. (**2006**). "Interaural correlation and loudness," in *Hearing – from basic research to applications*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer Verlag, Heidelberg, in press).

Domnitz, R. H. (1973). The interaural time jnd as a simultaneous function of interaural time and interaural amplitude. J. Acoust. Soc. Am., 53, 1549-1552.

Domnitz, R. H., and Colburn, H. S. (1977). Lateral position and interaural discrimination. J. Acoust. Soc. Am., 104, 2412-2425.

Durlach, N. I. (1972). Equalization and cancellation theory. In J. V. Tobias (Ed.), Foundations of Modern Auditory Theory. (Academic, New York).

Durlach, N. I., and Colburn, H. S. (1978). Binaural phenomena. In E. C. Carterette and M. P. Friedman (Eds.) Handbook of perception (pp. 360-466). Academic, New York.

Fletcher, H. (1940). Auditory patterns. Reviews of modern physics, 12, 47-65.

Gabriel, K. J., and Colburn, H. S. (1981). Interaural correlation discrimination: I. Bandwidth and level dependence. J. Acoust. Soc. Am., 69, 1394-1401.

Glasberg, B. R., and Moore, B. C. J. (2000). Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. J. Acoust. Soc. Am., 108, 2318-2328.

Goldberg, J. M., and Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J. Neurophysiol., 22, 613-636.

Grantham, D. W. (1982). Detectability of time-varying interaural correlation in narrow-band noise stimuli. J. Acoust. Soc. Am., 72, 1178-1184.

Grantham, D. W. (1984). Discrimination of dynamic interaural intensity differences. J. Acoust. Soc. Am., 76, 71-76.

Grantham, D. W. (1995). Spatial hearing and related phenomena. In B. C. J. Moore (Ed.), Hearing (pp. 297-347). Academic: London.

Grantham, D. W. and Wightman, F. L. (1978). Detectability of varying interaural temporal differences. J. Acoust. Soc. Am., 63, 511-523.

Grantham, D. W. and Wightman, F. L. (1979). Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J. Acoust. Soc. Am., 65, 1509-1517.

Hall, J. W., Tyler, R. S., and Fernandes, M. A. (1983). Monaural and binaural auditory frequency resolution measured using bandlimited noise and notched-noise masking. J. Acoust. Soc. Am., 73, 894-898.

Hall, J. W. and Fernandes, M. A. (1983). The role of monaural frequency selectivity in binaural analysis. J. Acoust. Soc. Am., 76, 435-439.

Hafter, E. R., and Dye, R. H., Jr. (1983). Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. J. Acoust. Soc. Am., 73, 644-651.

Hafter, E. R., and Buell, T. N. (1990). Restarting the adapted binaural system. J. Acoust. Soc. Am., 88, 806-812.

Harris, J. D. (1972). A florilegium of experiments on directional hearing. Acta Oto-Laryngol. Suppl., 298.

Harris, J. D., and Sergeant, R. L. (1971). Monaural/Binaural minimum audible angle for a moving sound source. J. Speech Hear. Res., 14, 618-629.

Hartmann, W. M., and Johnson, D. (1991). Stream segregation and peripheral channelling. Music percept., 9, 155-184.

van der Heijden, M. L., and Trahiotis, C. (1998). Binaural detection as a function of interaural correlation and bandwidth of masking noise: Implications for estimates of spectral resolution. J. Acoust. Soc. Am., 103, 1609-1614.

Hirsh, I. J. (1948). The influence of interaural phase on interaural summation and inhibition. J. Acoust. Soc. Am., 20, 536-544.

Holube, I., Kinkel, M., and Kollmeier, B. (1998). Binaural and monaural auditory filter bandwidths and time constants in probe tone detection experiments. J. Acoust. Soc. Am., 104, 2412-2425.

Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (1962). Effect of Interaural Correlation on the Precision of Centering a Noise. J. Acoust. Soc. Am., 34, 1122-1123.

Jorris, P. X., van de Sande, B., Recio-Spinoso, A., and van der Heijden, M. (2006). Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. Journal of Neuroscience, 26, 279-289.

Klein, M. A., and Hartmann, W. M. (1981). Binaural edge pitch. J. Acoust. Soc. Am., 70, 51-61.

Klumpp, R. G., and Eady, H. R. (1956). Some measurements of interaural time difference thresholds. J. Acoust. Soc. Am., 28, 859-860.

Kohlrausch, A. (1986). The influence of signal duration, signal frequency and masker duration on binaural masking-level differences. Hear. Res., 23, 267-273.

Kohlrausch, A. (1988). Auditory filter shape derived from binaural masking experiments. J. Acoust. Soc. Am., 84, 573-583.

Kollmeier, B., and Gilkey, R. H. (1990). Binaural forward and backward masking: Evidence for sluggishness in binaural detection. J. Acoust. Soc. Am., 87, 1709-1719.

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. J. Acoust. Soc. Am., 62, 157-167.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am., 49, 467-477.

Lopez-Poveda, E. A., and Meddis, R. (1996). A physical model of sound diffraction and reflections in the human concha. J. Acoust. Soc. Am., 100, 3248.

McFadden, D. (1966). Masking level differences with continuous and with burst masking noise. J. Acoust. Soc. Am., 40, 1414-1419.

Mills, A. W. (1958). On the minimum audible angle. J. Acoust. Soc. Am., 30, 237-246.

Moore, B. C. J. (1989). Space perception. In B. C. J. Moore (Ed.), An Introduction to the Psychology of Hearing (pp. 194-228). London: Academic Press Ltd.

Moore, B. C. J. (1998). Cochlear Hearing Loss. Whurr, London.

Moore, B. C. J., and Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. J. Acoust. Soc. Am., 74, 750-753.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). The shape of the ear's temporal window. J. Acoust. Soc. Am., 83, 1103-1116.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1990). Auditory filter shapes at low centre frequencies. J. Acoust. Soc. Am., 88, 132-140.

Mossop, J. E., and Culling J. F. (1998). Lateralization of large interaural time delays. J. Acoust. Soc. Am., 104, 1574-1579.

Musicant, A. D. and Butler, R. A. (1985). Influence of monaural speech cues on binaural localization. J. Acoust. Soc. Am., 77, 202-208.

Nedler, J. A., and Mead, R. (1965). A simplex method for function minimization. Comput. J., 7, 308.

van de Par, S., and Kohlrausch, A. (1999). Dependence of binaural masking level differences on centre frequency, masker bandwidth, and interaural parameters. J. Acoust. Soc. Am., 106, 1940-1947.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. J. Acoust. Soc. Am., 59, 640-654.

Perrot, D. R., and Musicant, A. D. (1977). Dynamic minimum audible angle: Binaural localization of moving sound sources. J. Acoust. Soc. Am., 62, 1463-1466.

Perret, S., and Noble, W. (1997). The effect of head rotations on vertical plane sound localization. Perception and Psychophysics, 59, 1018-1026.

Perrot, D. R., and Pacheco, S. (1989). Minimum audible angle thresholds for broadband noise as a function of delay between the onset of the lead and lag signals. J. Acoust. Soc. Am., 85, 2669-2672.

Plack, C. J., Gallun, F. J., Hafter, E. R., and Raimond, A. (2006). The detection of increments and decrements is not facilitated by abrupt onsets or offsets. J. Acoust. Soc. Am., 119, 3950-3959.

Plack, C. J., and Moore, B. C. J. (1990). Temporal window shape as a function of frequency and level. J. Acoust. Soc. Am., 87, 2178-2187.

Plomp, R. (1994). Noise, amplification, and compression: considerations of three main issues in hearing aid design. Ear Hear., 15, 2-12.

Pollack, I. (1978). Temporal switching between binaural information sources. J. Acoust. Soc. Am., 63, 550-558.

Robinson, D. E., and Jeffress, L. A. (1963). Effect of varying the interaural noise correlation on the detectability of tonal signals. J. Acoust. Soc. Am., 65, 1947-1952.

Rose, M. M., and Moore, B. C. J. (1997). Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. J. Acoust. Soc. Am., 102, 1768-1778.

Rose, M. M., and Moore, B. C. J. (2000). Effects of frequency and level on auditory stream segregation. J. Acoust. Soc. Am., 108, 1209-1214.

Saberi, K. (1995). Some considerations on the use of adaptive methods for estimating interaural-delay thresholds. J. Acoust. Soc. Am., 98, 1803-1806.

Siveke, I., Ewert, S. D., and Wiegrebe, L. (2006). Perceptual and physiological characteristics of binaural sluggishness. In B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Eds.), Hearing – from basic research to applications. Springer Verlag, Heidelberg (in press).

Sondhi, M. N., and Guttman, N. (1966). Width of the spectrum effective in the binaural release of masking. J. Acoust. Soc. Am., 40, 600-606.

Spitzer, M. W., and Semple, M. N. (1993). Responses of inferior colliculus neurons to time-varying interaural phase disparity: effects of shifting the locus of virtual motion. J. Neurophysiol., 69, 1245-1263.

Spitzer, M. W., and Semple, M. N. (1995). Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. J. Neurophysiol., 73, 1668-1690.

ter Keurs, M., Festen, J. M., and Plomp, R. (1993). Effect of spectral envelope smearing on speech reception. II. J. Acoust. Soc. Am., 93, 1547-1552.

Tobias, J. V., and Zerlin, S. (1959). Lateralization threshold as a function of stimulus duration. J. Acoust. Soc. Am., 31, 1591-1594.

Trahiotis, C., Bernstein, L. R., Buell, T. N., and Spektor, Z. (1990). On the use of adaptive procedures in binaural experiments. J. Acoust. Soc. Am., 87, 1359-1361.

Trahiotis, C., Dolan, T. R. and Miller, T. H. (1972). Effect of 'backward' masker fringe on the detectability of pulsed diotic and dichotic tonal signals. Percept. Psychophys., 12, 335-338.

Viemeister, N. F. (1977). Temporal factors in audition: A systems analysis approach. In E. F. Evans and J. P. Wilson (Eds.), Psychophysics and Physiology of Hearing (pp. 419-428). London, Academic.

Viemeister, N. F., and Wakefield, G. H. (1991). Temporal integration and multiple looks. J. Acoust. Soc. Am., 90, 858-865.

Wagner, H. (1991). A temporal window for lateralization of interaural time differences in barn owls. J. Comp. Physiol. A, 169, 281-289.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. J. Exp. Psychol., 27, 339-368.

Wightman, F. L., and Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. J. Acoust. Soc. Am. 91, 1648-1661.

Wightman, F. L., and Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. J. Acoust. Soc. Am. 105, 2841.

Witton, C., Green, G. R., Rees, A., and Henning, G. B. (2000). Monaural and binaural detection of sinusoidal phase modulation of a 500-Hz tone. J. Acoust. Soc. Am., 108, 1826-1833.

Witton, C., Simpson, M. I. G., Henning, G. B., Rees, A., and Green, G. R. (2003). Detection and direction-discrimination of diotic and dichotic ramp modulations in amplitude and phase. J. Acoust. Soc. Am., 113, 468-477.

Yama, M. (1992). Effects of temporal separation and masker level on binaural analysis in forward masking. J. Acoust. Soc. Am., 91, 327-335.

Yost, W. A. (1985). Prior stimulation and the masking level difference. J. Acoust. Soc. Am., 78, 901-907.

Zurek, P. M. (1980). The precedence effect and its possible role in the avoidance of interaural ambiguities. J. Acoust. Soc. Am., 67, 952-964.

Zurek, P. M. (1980). The precedence effect. In W. A. Yost and G. Gourevitch (Eds.) Directional Hearing. Springer-Verlag, New York.

Zurek, P. M., and Durlach, N. I. (1987). Masker-bandwidth dependence in homophasic and antiphasic tone detection. J. Acoust. Soc. Am., 81, 459-464.

Zwicker, E., (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J. Acoust. Soc. Am., 33, 248.

Zwicker, E., and Feldtkeller, R. (1967). Das Ohr als Nachrichtenempfanger. Stuttgart: Hirzel-Verlag.