SPATIAL HEARING AND TEMPORAL PROCESSING IN OLD AND HEARING-IMPAIRED INDIVIDUALS

A thesis submitted to the University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Medical and Human Sciences

2015

Andrew J King

School of Psychological Sciences

Table of Contents

Word Count: 57,303 words

List of Tables

List of Figures

Abbreviations

3I2A	Three-interval, two-alternative forced choice task where the first interval is the reference, and the latter two intervals are the alternatives
8S2A	Two-interval, two-alternative, forced choice task with four tones per interval (AAAA *versus* ABAB)
ABR	Auditory brainstem response
AM	Amplitude-modulated
AMBBs	Amplitude-modulated binaural beats
AN	Auditory nerve
ANOVA	Analysis of variance
ANCOVA	Analysis of covariance
ASSR	Auditory steady-state response
AT	Absolute threshold
AT250	Absolute threshold at 250 Hz
AT500	Absolute threshold at 500 Hz
BIC	Binaural interaction component
BM	Basilar membrane
BMLD	Binaural masking level difference
CF	Characteristic frequency
CHL	Cochlear hearing loss
CNS	Central nervous system
δ	The variable difference in a stimulus parameter under discrimination
DFT	Discrete Fourier transform
EEG	Electro-encephalography
EFFR16	Envelope-frequency following response at 16 Hz
EFFR27	Envelope-frequency following response at 27 Hz
EFFR115	Envelope-frequency following response at 115 Hz
EFFR145	Envelope-frequency following response at 145 Hz
ENV	Envelope
ENV250	Envelope-IPD discrimination threshold with 250 Hz pure-tone carrier
ENV500	Envelope-IPD discrimination threshold with 500 Hz pure-tone carrier
ERBN	Equivalent rectangular bandwidths of normal-hearing listeners
f	Frequency (Hz)
f_c	Carrier (or centre) frequency
f_m	Modulation frequency
F0DL	Fundamental frequency difference limen
FFR	Frequency following response
FM	Frequency modulation

HI	Hearing-impaired
HRIR	Head-related impulse response
IHC	Inner hair cell
ILD	Interaural level difference
IPD	Interaural phase difference
ITD	Interaural time difference
ITD500	500 Hz Pure tone ITD discrimination
MEG	Magneto- encephalography
MHA	Master hearing aid system
n (or N)	Sample size
NH	Normal-hearing
OHC	Outer hair cell
PTA	Pure-tone audiometry (or audiogram)
PTAHF	mean high-frequency audiometric threshold (from 2 to 8 kHz)
PTALF	mean low-frequency audiometric threshold (from 0.125 to 1.5 kHz)
SD	Standard deviation
SL	Sensation level
SNHL	Sensorineural hearing loss
SNR	Signal-to-noise ratio
SPL	Sound pressure level
SR	Spontaneous rate
SRM	Spatial release from masking
SRT	Speech reception threshold
TFS	Temporal fine structure
TFS250	TFS-IPD discrimination threshold with 250 Hz pure-tone carrier
TFS500	TFS-IPD discrimination threshold with 500 Hz pure-tone carrier
TFS1, $TFS1_{H6}$, $TFS1_{H12}$	Discrimination thresholds for identically band-pass filtered harmonic and inharmonic complexes (constant frequency shift across components). Subscript denotes harmonic number at the pass-band centre frequency
$TFSFFR_{HIGH}$	Temporal-fine-structure-frequency following response (Mean across 433, 472, 510, 537, 564, 578 and 723 Hz)
$TFSFFR_{LOW}$	Temporal-fine-structure-frequency following response (Mean across 242, 291, 307, 323 and 357 Hz)
TMR	Target-to-masker ratio.
$TMR_{50\%}$	Target-to-masker ratio estimated as 50% word recall threshold.
TMTF	Temporal modulation transfer function
TVC	Tone vocoding

Abstract

The University of Manchester

Andrew J King

Doctor of Philosophy

Spatial Hearing and Temporal Processing in Old and Hearing Impaired Individuals

January 2015

Small timing differences occur when sounds reach one ear before the other, creating interaural phase differences (IPDs). The phase-locked activity in the auditory nerve can, at low frequencies, preserve IPDs. IPDs are used for localising and separating sounds from different directions. Chapters 3, 5, and 6 report three studies of the independent effects of age and sensorineural hearing loss on the temporal processing of sound that aids spatial hearing. Chapters 2 and 4 describe two supporting methodological studies. Chapter 2 compared the duration of training required for stable IPD-discrimination thresholds for two stimulus presentation procedures. The procedure requiring the least training was adopted for subsequent studies. Age and hearing loss are related and both may affect sensitivity to IPDs. Chapter 3 demonstrated that hearing loss, regardless of listener age, is related to poorer sensitivity to IPDs in the temporal fine structure (TFS), but not in the temporal envelope. Chapter 3 also showed that age, independent of hearing loss, is related to poorer envelope-IPD sensitivity at low modulation rates, and somewhat poorer TFS-IPD sensitivity. In Chapter 5, listener age and IPD sensitivity were both compared to subcortical neural phase locking measured through the frequency-following response (FFR). Phase coherence in the envelope-FFR at 145 Hz modulation and in the TFS-FFR deteriorated with age, suggesting less precise phase locking in old age. However, age-related changes to IPD sensitivity were not strongly related to age-related changes in FFR phase coherence. IPD sensitivity declines may be predominantly caused by deterioration of binaural processing independent of subcortical phase locking. Chapter 4 showed that electrodes at the mastoids recorded TFS-FFR generated earlier in the auditory pathway than electrodes from the nape of the neck to forehead, which recorded FFR generated later in the brainstem. However, these electrode montages did not reveal different age- or hearing-loss-related FFR deficits in Chapter 5. Chapter 6 determined whether hearing loss affected the ability to use TFS IPDs to achieve better speech perception. On average, old hearing-impaired listeners gained a small, but significant, benefit from a lateral separation of the speech sources. Replacing the TFS with binaurally in-phase sine waves (removing the TFS IPDs) significantly reduced the benefit of lateral separation. How much a listener benefitted from intact TFS IPDs in speech perception was strongly related to the extent of their hearing loss at low frequencies and their monaural processing of TFS, but not to their ability to discriminate IPDs. In general, this thesis shows that low-frequency hearing loss is associated with poor sensitivity to TFS IPDs and the ability to benefit from them when sounds are laterally separated. The thesis also shows that old age can reduce sensitivity to IPDs and weaken subcortical temporal coding. Although only partly related, these effects are likely to cause problems for old individuals in challenging listening environments.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.

Acknowledgements

First and foremost, I would like to thank my supervisors, Kathryn Hopkins and Chris Plack, for the guidance and support they provided me during my studentship; their consideration and feedback on my work has been extraordinary. I could not ask for better mentors. Second, I thank the members of the Eriksholm Research Centre whom I had the pleasure of working with during a glorious Spring in 2013, particularly my supervisor Niels Henrik Pontoppidan and my colleagues Lars Bramsløw, Renskje Hietkamp, Marianna Vatti, Atefeh Hafez, and Ariane Laplante-Lévesque. The support from colleagues at the University of Manchester should also be acknowledged: Keith Wilbraham, Richard Baker, Fred Marmel, Oliver Bones, Ronan McGarrigle, Michael Stone and Garreth Prendergast; I thank you. My thanks also goes to Hannah Guest, Ann-marie McCormack, my brother, my father and my uncle, who all helped proofread and provide feedback on various parts of this thesis.

Most importantly, I thank my parents for knowing, and saying, exactly what I needed to hear to help me through the darker times. I am convinced all my merits stem from their unconditional love and encouragement.

Chapter 1. Introduction, literature review and thesis research overview

This chapter was written about the published research available to the author at the beginning of the studentship (articles published before and in early 2012). This is to demonstrate how the research available at the time informed and directed the research reported in subsequent chapters of this thesis. The subsequent chapters make reference to more recent studies.

1.1 Introduction

Beyond the age of 40, humans can experience a range of hearing difficulties. This includes:

i. diminished audibility of quiet sounds (Davis, 1989; Gordon-Salant *et al.*, 2010),
ii. diminished processing of the sounds that remain audible (e.g., Grose *et al.*, 2006),
iii. development of poorer speech understanding in noisy environments (Working Group on Speech Understanding and Aging, 1988).

These problems accelerate after the age of 50 (Plomp and Mimpen, 1979; Patterson *et al.*, 1982) and prevalence increases. Age-related hearing loss (ARHL) is at least the third most common long-term health problem affecting people aged 65 years and older (Gordon-Salant *et al.*, 2010). As average life expectancy is increasing in many developed regions (DESA, 2007), ARHL is an important area of research.

In noisy environments, sounds arrive at a listener from multiple directions. Listeners use binaural cues to help locate the sound sources relative to the listener's position (Darwin, 2006). Young, normal-hearing (NH) listeners can use this to improve understanding of speech when there are competing sounds from other angles (e.g., Bronkhorst and Plomp, 1988; Bronkhorst, 2000; Behrens *et al.*, 2008; Marrone *et al.*, 2008b). This predominantly relies on precise temporal processing at each ear and binaural integration of the neural signals (e.g., Wightman and Kistler, 1992). Old and some hearing impaired (HI) people benefit less from spatial cues when separating out sounds, and are also poorer at localising sounds (Noble *et al.*, 1994; Abel *et al.*, 2000; Marrone *et al.*, 2008a; Neher *et al.*, 2009). This may be a result of impaired temporal processing.

The main aim of this thesis is to test whether age and hearing loss independently impair the temporal processing needed to localise and spatially separate sounds. This is important because age and hearing impairment are closely related. It is unclear whether the age-related hearing impairments are mainly due to cochlear hearing loss (CHL) accumulated over time (e.g. due to noise exposure) or declines of higher auditory processes in the brain.

The first section of this review describes the temporal coding of sounds through the neural-firing patterns and how this can be measured. The second section describes how this temporal code aids the binaural abilities of the auditory system, and why these are important for locating the lateral position of sounds and for identifying speech in noisy environments. The third section discusses the relationship between threshold elevation (due to cochlear pathology) and the processing of audible sounds, focusing on temporal and spatial processing. The fourth section considers how much temporal processing and spatial hearing deteriorate with age, what the causes may be and how these relate to threshold elevation. The final section details how this thesis addresses some of these questions.

1.2 Temporal processing

1.2.1 **Encoding sound into neural firing patterns**

Once a sound wave has been passed through the external and middle ear to the inner ear, the movement of fluid inside the cochlea generates a travelling wave along the basilar membrane (BM). The BM behaves like a series of overlapping band-pass filters (von Békésy, 1960; Moore, 2007) as the wave travels from the base to the apex of the cochlea (resolving the high and low ends of the audible frequency range, respectively) . This begins the tonotopic analysis of sound in auditory perception.

There are two types of hair cells arranged in rows along the BM: inner hair cells (IHCs) and outer hair cells (OHCs). Each IHC is innervated by approximately 20 to 25 nerve fibres that send signals up the afferent pathway towards the auditory central nervous system (CNS) (Spoendlin, 1970; Liberman, 1982). The OHCs are mostly innervated by a small proportion of afferent fibres, but also by efferent nerve fibres that receive signals from the CNS which control the OHCs (Guinan, 2006). OHCs do not encode sound, but actively change the response of the BM. The OHCs sharpen the filter bandwidth and provide amplification at low to moderate sound levels—this gives the auditory system the non-linear, compressive input-output function (Robles and Ruggero, 2001). This provides NH listeners with audibility across a large dynamic range.

Movement of the BM makes stereocilia on the hair cells move in the endolymph (the fluid surrounding the hair cells), initiating a change in membrane potential in the nerve fibres stemming from the hair cell terminals (Spoendlin and Schrott, 1989). The firing of neurons connected to each IHC can encode the frequency spectrum of sounds in two ways: tonotopically and temporally. In the first case, neurons fire most for travelling waves that peak at the place on the BM where the neuron synapses with the

IHC, giving the neuron a characteristic frequency (CF) (Liberman, 1978; Sachs and Young, 1979). The rate of firing (above their spontaneous rate) as a function of nerve fibre CF provides a "rate-place" code. The temporal pattern of neural firing also encodes the frequency spectrum. In response to low-frequency stimuli, auditory nerve (AN) fibres fire most around the peak depolarisation of the inner hair cell (Davis *et al.*, 1950; Rose *et al.*, 1967; Brugge *et al.*, 1969; Palmer and Russell, 1986), and thus at a particular phase of the stimulus periodicity (Rose *et al.*, 1971). The stimulus frequency limit, above which phase locking does not occur, varies dramatically with species (see Palmer and Russell, 1986). By locking to a specific phase, the firing pattern represents temporal information in the stimulus with great accuracy (Young and Sachs, 1979). This phase locking provides a temporal code. Individual fibres do not always fire every cycle of the stimulus, but the responses from the collective fibres synapsing an IHC combine with great temporal precision into a very accurate code of the stimulus frequency in the cochlear nucleus (Joris *et al.*, 1994a; Joris *et al.*, 1994b).

A rate-place code would be dependent on the sharpness of the filter tuning on the BM. This filtering broadens with increasing input level and reduces tonotopic resolution or frequency selectivity of the filtering (e.g., Moore and Glasberg, 1987b; Baker and Rosen, 2006). Sachs and Young (1979) showed that the rate-place coding of vowel formants in cat AN fibres lost spectral definition at high sound levels. Phase locking, on the other hand, is robust from moderate to high levels above threshold, providing consistent frequency information despite changes in BM filtering with level (Young and Sachs, 1979). However, phase locking appears to break down above a certain high-frequency limit which varies between animals (Tasaki, 1954; Kiang *et al.*, 1965). This limit appears to be set by the low-pass-filtering characteristics of the IHC membrane potential (Palmer and Russell, 1986), but possibly also by time limits in signal transmission at synapses (Anderson *et al.*, 1971). However, behavioural tests (Moore and Sek, 2009b) and comparisons of performance modelled with rate-place and temporal encoding (Heinz *et al.*, 2001) suggest humans may be able to encode the temporal properties of sounds up to 10 kHz.

In response to complex sounds, the band-limited filtering on the BM produces narrowband signals with two types of temporal information. For stimulation at a given place on the BM, the temporal fine structure (TFS) is the oscillation close to the auditory filter-band centre frequency (f_c). Slower fluctuations in amplitude are superimposed on the TFS. This modulation is called the temporal envelope. A rate-place code can represent temporal changes in sound by changes in firing rate over time. This could communicate the envelope of the output of each filter-band, but the TFS would be lost. Phase locking however, can preserve both TFS and envelope information in the timing of the neural firing. When a target sound has a different TFS to masker sounds, changes in the TFS as the masker fluctuates may help the auditory system detect the target (e.g., Moore and Glasberg, 1987a; Füllgrabe *et al.*, 2006; Lorenzi *et al.*, 2006). Moore and Glasberg (1987a) found detection thresholds for a pure tone were lower in a fluctuating masker than a non-fluctuating masker, but this masking release disappeared above 5 kHz (as does phase locking). If the change in place of peak excitation (from masker f_c to target f_c) in the masker dips was the cue used for masking release then masking release should not diminish above 4-5 kHz (Moore and Glasberg, 1987a). Moore and Glasberg (1987a) argued that TFS information is used for identifying a target sinusoid in a fluctuating masking sound of a different f_c by 'listening in the dips' of the masker. Differences in envelope modulation depth may also cue the presence of the target signal, but masking release occurs independently of modulation depth and modulation depth discrimination thresholds are not acute enough to explain masking release (Moore and Glasberg, 1987a). This suggests it is unlikely envelope cues are used for masking release with fluctuating maskers.

Temporal coding is thought to have benefits for supra-threshold auditory perception in noisy environments (Sachs *et al.*, 1983). The ability of listeners to identify target speech in a fluctuating background noise also depends on their access to TFS cues. When the TFS is replaced with noise (whilst the envelope cues are left relatively unchanged) performance is worse than when the speech is unprocessed (Hopkins and Moore, 2009). Lorenzi *et al.* (2006) found that young HI listeners understood syllables less well in steady noise than in 100% amplitude modulated (8 Hz) noise. They showed that this difference correlated with the ability to identify the same syllables without noise when they were processed to preserve TFS cues and disrupt envelope cues. This suggests that although envelope information is sufficient to understand speech in quiet (Shannon *et al.*, 1995), TFS information is required in the presence of modulated maskers (Hopkins and Moore, 2009). However, the assumption that use of TFS and envelope information can be measured in isolation through extraction, manipulation and recombination is debated (Kates, 2011; Shamma and Lorenzi, 2013). It is possible that TFS and envelope information can be reconstructed from each other in the auditory system.

From the research detailed above, it is clear that the temporal code provided by phase locking is useful for identifying and separating sounds in noisy and fluctuating backgrounds, particularly at high sound levels.

1.2.2 **Measuring phase locking in humans**

One non-invasive method of measuring neural phase locking in humans is through electroencephalography (EEG). EEG measures the electrical potentials at the scalp created across a large number of neurons as they fire (Burkard *et al.*, 2007). It can be used to physiologically measure a listener's neural response to sounds (or auditory-evoked potentials). Evoked responses are usually determined from a large sample of repeated stimulus presentations by averaging. With increasing repetitions of the stimulus, the responses to the stimulus add constructively whereas the noise in the response cancels out. At the scalp, auditory brainstem responses are very weak, but the collective phase-locked neural response to a periodic stimulus can be observed in the average EEG after thousands of presentations (Moushegian *et al.*, 1973). By taking the Fourier transform of one such average waveform, a peak in the magnitude spectrum can be observed at the frequency of the acoustic signal (e.g., Worden and Marsh, 1968; Galbraith *et al.*, 2000). This shows that there is periodicity in the neural firing that follows the frequency of the stimuli; hence this response is known as the frequency following response (FFR). With complex stimuli, peaks in the spectrum can be found at the component frequencies (which constitute the TFS), at envelope periodicities (such as the frequency difference in a tone pair) and at harmonics related to the stimulus (Krishnan, 2007). A procedure developed by Goblick and Pfeiffer (1969) can be used to emphasise the FFR reflecting phase locking to either the envelope or the TFS (Huis in't Veld *et al.*, 1977). On each trial, the complex stimulus is sequentially presented twice, the second time with opposite polarity to the first. Summation of the two presentations emphasises the response following the envelope whilst the response to the TFS mostly cancels out; subtraction of the two presentations emphasises the response following the TFS, whilst response to the envelope mostly cancels out (Aiken and Picton, 2008; Gockel *et al.*, 2011). This procedure allows the researcher to differentially study the neural representation of these two temporal aspects of sound from the same data.

FFR measured at the scalp appears to be generated at multiple sources. FFR is often measured in a vertical electrode montage orientation: from the midline of the forehead at the hairline, or from the vertex of the head, to the midline of the back of the neck, either at the hairline or at the seventh cervical vertebra. FFR recorded this way is assumed to originate in the rostral brainstem (either inferior colliculus or lateral lemniscus) (e.g., Gerken *et al.*, 1975; see Krishnan, 2007). However, Batra *et al*. (1986) found the FFR oscillated in magnitude as a function of stimulus frequency. They inferred this was a result of at least two sources that were a fixed distance apart. This distance leads to responses cancelling out or constructively summing, depending on response frequency. Furthermore, the phase spectrum broadly resembled two linear functions of frequency: one at low frequencies with a steep function slope and the other at higher frequencies with a shallow function slope (Batra *et al.*, 1986). Steep-sloped functions suggest a later latency of response than shallow-sloped functions (see chapter 4).

The nature of the FFR is dependent on the orientation of the recording electrodes. Stillman *et al*. (1978) found two peaks per cycle in FFRs to pure-tone stimuli. The second peak lagged behind the first by 1.3 ms, regardless of frequency (and cycle periodicity). The leading peaks were more prominent than the lagging peaks when FFR was recorded horizontally (from mastoid to vertex), but less prominent than the lagging peaks when FFR was recorded vertically. Stillman *et al*. (1978) took this as evidence that an earlier FFR generator (creating the leading peaks in the waveform) was better measured horizontally and a later FFR generator (creating the lagging peaks) was better measured vertically. Galbraith *et al*. (2000) found that measuring the FFR from the ear canal provided positive signal-to-noise ratios (SNRs) for the FFR up to 950 Hz, whereas the FFR measured from the top of the head only had positive SNRs below 536 Hz. Galbraith *et al*. (2000) suggest that the FFR measured at the ear canal originated in the AN, whilst the FFR from the top of the head originated from the brain stem, because the AN can phase lock to higher frequencies than later brainstem structures.

FFR is a useful measure of neural phase locking to sounds. It may be possible to measure FFR from different subcortical structures using different electrode montages. It is worth keeping in mind that because EEG measures the extracellular activity from large numbers of neurons, a lack of FFR may be caused by a destructive phase relationship between the phase-locked firing of adjacent neurons rather than a lack of phase locking *per se*.

1.3 Binaural processing, localisation and spatial separation

When a sound source lies off the sagittal plane, the sound is greater in level at the nearer ear than the further ear (*i.e.* an interaural level difference, or ILD), and it reaches the nearer ear earlier than the farther ear (*i.e.* an interaural time difference, or ITD). Below about 1600 Hz, half the wavelength of a sine wave (in air) is larger than the distance between the ears, so the sound arrives at the further ear with a phase delay between 0 and 180°. For higher frequencies, the phase delay could be greater than 180°, and from the interaural phase difference (IPD) alone, it would be unclear which ear received the sound first. Above 1600 Hz, the head casts an acoustic shadow, creating a strong ILD cue. The duplex theory (Rayleigh, 1907) purports that IPDs are used below 1600 Hz frequency, and ILDs are used for sounds above this frequency. Henning (1974) revised the duplex theory slightly, showing that an IPD in a 3900 Hz tone modulated at 300 Hz rate was as easy to detect as an IPD in a 300-Hz pure tone. Furthermore, if the 3900 Hz tone carrier was changed to a different frequency in one ear, but the modulation remained 300 Hz, the envelope-IPD threshold did not change substantially (Henning, 1974). IPD cues dictate perceived sound source localisation when the sound contains envelope or TFS information below 1300 Hz, and ILDs direct localisation only when the sound does not contain such information (Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002). This shows that for complex stimuli, time-based cues for localisation can still be used as long as some periodicity is in the unambiguous frequency range, whether it is TFS or envelope.

Discrimination of azimuth between two sound sources is more acute towards the midline (0°azimuth; *e.g.* directly ahead) than when the sources are 90° to the left or right (Mills, 1958). Likewise, as the starting phase of binaural sinusoids is increased (up to 180°), the threshold discrimination of an IPD increases (Yost, 1974). Middlebrooks and Green (1990) showed that envelope ITDs increase monotonically with increasing azimuth of a sound source, at a rate of roughly 7 to 8 μs per angle degree, although it is questionable how salient envelope-ITDs in high-frequency sounds are for localisation (Musicant and Butler, 1985).

Colburn and Esquissaud (1976) suggested that the same binaural neural mechanism processes interaural TFS differences and interaural envelope differences. Bernstein and Trahiotis (2002) tested this by comparing envelope-IPD thresholds for high-frequency carriers with low frequency, half-wave rectified envelopes to IPD thresholds for low-frequency pure tones. Unlike sinusoidal amplitude modulation, which would produce different outputs from the auditory nerve for envelope and TFS, half-wave rectified envelopes should produce envelope responses similar to the TFS responses to pure tones (Van de Par and Kohlrausch, 1997). Bernstein and Trahiotis (2002) found that IPD thresholds for half-wave rectified envelopes and low-frequency pure tones were comparable, supporting Colburn and Esquissaud's (1976) hypothesis.

Interaural cues are also thought to be used to aid perception (such as following a conversation) in challenging listening environments where the competing sounds are arriving from different spatial locations and can therefore be distinguished by their ITDs and ILDs. This phenomenon is often measured by spatial release from masking (SRM). SRM is defined as the improvement in detection threshold or intelligibility when the signal and masker noise come from separate positions around listener, compared to when they come from the same spatial position.. Using headphones, many early studies of binaural hearing showed the contribution of IPDs to simulated SRM by simply inverting the polarity of the target signal in one ear (so it has a π radian IPD), in the presence of diotic noise. Detectability is poor if both noise and target signal are both presented diotically, but detectability improves substantially if the signal is inverted at one ear (Hirsh, 1948; Licklider, 1948; Schubert and Schultz, 1962 ; Levitt and Rabiner, 1967). The reduction (improvement) in the masked threshold is called the binaural masking level difference (BMLD). Young NH listeners can achieve a BMLD of 15 dB (Hirsh, 1948). This demonstrates that the binaural processing of phase information is very advantageous to unmasking sounds, even without ILDs.

To determine the benefits of spatial hearing most commonly occurring in day-to-day life, many researchers have studied the benefit of SRM to speech intelligibility in noise (Hirsh, 1950; Bronkhorst and Plomp, 1988) and in competing talkers (e.g., Bronkhorst, 2000; Freyman *et al.*, 2001). Young, NH listeners can achieve around 14 dB SRM, but it is dependent on the type of stimuli (e.g. predictability of target speech, type of masker). It is also dependent on how large is the azimuth of separation between the target and masker.

When sound sources are spatially separated, it is possible for listeners to take advantage of the higher SNR at the ear nearer the target source (the better-ear effect). However, there is an additional speech intelligibility benefit of listening with both ears that implicates the use of interaural differences (Hawley *et al.*, 1999; Hawley *et al.*, 2004). Culling *et al*. (2004) used head-related transfer functions to simulate sound sources at different angles around the listener. Using headphones allowed the researchers to present either the ITD or ILD individually, or together. They found that when only ILDs were available for spatially-separated speech, SRM only occurred when the better-ear effect was possible (*i.e.* when the maskers were not evenly distributed around the listener). The ITD-only and combined-cues conditions produced similar patterns of results; SRM occurred for all conditions with the maskers spatially separated from the target, regardless of the potential for the better-ear effect. Interestingly, although localisation and SRM appear to benefit greatly from ITDs or IPDs, the binaural unmasking in SRM does not appear to be driven by localisation (Culling *et al.*, 2004). For example, speech SRM was good regardless of whether the ITDs were consistent across high- and low-frequency bands or not, as long as the ITDs were opposite for the target and masker in each frequency band (Edmonds and Culling, 2005).

Whilst the better ear effect may help listening in noisy environments, the interaural cues give an advantage in certain conditions. The evidence suggests IPDs in particular are useful for this, as well as for locating sound sources.

1.4 Hearing loss

Hearing loss is usually considered as the extent to which the absolute threshold of a sound (the quietest level at which it can be detected by the listener) is elevated. Cochlear hearing loss (CHL) is usually attributed to damage to the hair cells and the connected nerve fibres in the cochlea (Gleeson and Felix, 1987). The effects of damage to the different types of hair cells have been studied in the responses of cats' auditory nerves (Liberman *et al.*, 1986). Liberman *et al*. (1986) measured neural responses after acoustic trauma and ototoxic drug exposure and traced the fibres to the IHCs onto which they synapsed. The condition of the IHCs and neighbouring OHCs differed with neural response. Where the IHCs had been damaged, but the OHCs less so, a 40 dB loss in sensitivity was observed, but neural responses were still tuned finely to the CF. Damage to the OHCs produced much broader filtering; the lower frequency shoulder was as sensitive as normal, but the lack of a sharp peak to the filter corresponded to a 40 to 50 dB loss (Liberman *et al.*, 1986).

Schrott *et al.* (1989) showed that a loss of 70% of IHCs in mice led to only moderate threshold elevation, as long as the loss was evenly spread and the OHCs were not damaged. However, a lack of substantial threshold elevation does not preclude impairments to supra-threshold hearing. Kujawa and Liberman (2009) found that after 1 hour of high-level noise exposure, thresholds were elevated in mice, but returned to normal after 8 weeks. However, approximately 50% of the afferent nerve fibres no longer innervated the IHCs in the basal turn of the cochlea. This deafferentation has been suggested as a reason why some people cannot detect signals in noise (Kujawa and Liberman, 2009), or why they perform poorly at tasks requiring sensitivity to TFS (Hopkins and Moore, 2011). A lack of afferent nerve fibres connecting to the IHCs (Kujawa and Liberman, 2009) or number of IHCs themselves (Buss *et al.*, 2004) could limit the clarity of phase locking and therefore of accurate TFS representation.

1.4.1 **Supra-threshold changes**

A great amount of research has explored the difference between NH listeners and HI listeners in discriminating between audible sounds (see, for example, Moore, 2007). Two aspects of supra-threshold processing have received particular attention: frequency selectivity and temporal processing. Frequency selectivity refers to the bandwidths of the overlapping cochlear filters and how well the ear can resolve complex sounds. Temporal processing involves the ability to sense changes in sounds over time, including gaps and fluctuations. Such features may be conveyed by phase locking, but also by changes in rate-place neural firing over time.

Studies have investigated how frequency selectivity and temporal processing relate to speech perception (Dreschler and Plomp, 1980; Tyler *et al.*, 1982; Dreschler, 1983; Dreschler and Plomp, 1985). Reduced frequency selectivity and temporal processing may explain why HI listeners struggle to perceive audible sounds accurately, such as comprehending speech in noisy backgrounds (Plomp, 1978). Whilst threshold elevation renders the quiet parts of speech inaudible and thus reduces the articulation the listener perceives (Zurek and Delhorne, 1987), at least amplification by hearing aids may restore received articulation (Moore, 2007). Supra-threshold deficits, on the other hand, cannot be corrected for by amplification (Plomp, 1978).

In CHL, OHCs are usually more damaged than IHCs, leading to notable frequency selectivity reductions. The active mechanism provided by the OHCs is lost, producing broader auditory filters. Auditory filter bandwidth can be estimated by determining pure-tone thresholds in noise with band-stop notches either symmetrically or asymmetrically centred at the pure tone frequency (e.g., Patterson, 1976; Moore and Glasberg, 1983; 1987b; Glasberg and Moore, 1990). As the notch bandwidth increases from zero, the threshold decreases (improves), but above a certain bandwidth the threshold ceases to decrease. This bandwidth is taken as the auditory filter bandwidth. There is consistent evidence that reduced frequency selectivity is related to poorer speech perception (Patterson *et al.*, 1982; Strelcyk and Dau, 2009; Hopkins and Moore, 2011).

Irwin *et al*. (1981) studied the minimum detectable gap in broadband noise by NH and HI listeners. Thresholds for NH listeners varied widely (7 to 35 ms) at 30 dB sound pressure level (SPL), but converged, after decreasing rapidly as noise level increased, on an asymptote of approximately 4 ms by 50 dB SPL. Listeners with conductive hearing losses showed a similar steep decrease in threshold with increasing level, but displaced to approximately 40 to 50 dB higher levels (converging at 4 ms by 100 dB SPL). Listeners with CHL showed much shallower gap-detection-threshold functions of noise level, decreasing gradually from around 14 to 21 ms at 30 dB SPL to around 10 ms on average by 90 dB, never reaching the acuity of the NH listeners' thresholds. These data suggest that whilst conductive hearing loss may not affect temporal resolution (normal gap-detection thresholds can be attained by simple audibility compensation), sensorineural hearing loss appears to fundamentally change (for the most part, deteriorate) temporal resolution. Whilst temporal resolution may be important for perceiving gaps and stops in sounds (such as between syllables and words in speech), it is not necessarily related to the phase locking (or TFS processing) capabilities of the auditory system.

1.4.2 Changes to phase locking and TFS processing

It is not clear whether listeners with CHL lack fidelity in phase locking (Harrison and Evans, 1979; Woolf *et al.*, 1981; Miller *et al.*, 1997). It is possible that nerve fibres do not lose their phase locking accuracy with CHL, but the loss in number of nerve fibres (Spoendlin, 1970; Felder and Schrott-Fischer, 1995) may reduce the strength of the phase locked response in the neural population, as phase locking depends on multiple fibres (Joris and Smith, 2008). Alternatively, the TFS at the output of the BM may be disrupted by the abnormal phase response of the BM due to the loss of the active mechanism provided by the OHCs (Carney *et al.*, 2002). Thirdly, Hopkins and Moore (2011) suggested poor frequency selectivity would result in more complex TFS in each auditory filter channel because it would encompass more components. The information in adjacent channels would correlate more (Heinz, 2010), which could make differentiating between the components of two simultaneous sounds more difficult. It is not clear from these studies whether the decline in phase locking is due to peripheral damage such as CHL or due to changes to the central auditory system.

Old HI listeners are less able than young NH listeners to discriminate between sounds that only differ in TFS, even when they have normal thresholds in the frequency region encompassing the stimuli (Hopkins and Moore, 2011). HI listeners also appear to benefit less from TFS in speech (Hopkins *et al.*, 2008), even when the manipulation of TFS is restricted to frequency regions where they have normal thresholds (Lorenzi *et al.*, 2006; Lorenzi *et al.*, 2009). Lorenzi *et al.* (2006) processed consonant sounds using the Hilbert transform to separate the TFS and envelope. The performance of NH listeners and HI listeners with flat, moderate hearing loss in identifying these consonants was examined when only the TFS or only the envelope information was presented. NH listeners performed well at the task in both these conditions, but the HI listeners only performed well when the envelope was available, not when only the TFS was available. Lorenzi *et al.* (2009) used similar processing, but with nonsense syllables that were low-pass filtered (at 1.5 kHz), restricting the TFS and envelope information to low-frequencies. Lorenzi *et al.* (2009) found that HI listeners with mid- or high-frequency hearing loss, but normal low-frequency hearing, performed much worse than NH listeners when only the TFS cues were available. The HI and NH listeners performed similarly when only envelope cues were available. These results suggest that, as well as elevated thresholds at (usually) high frequencies, hearing loss may impede the listener’s ability to benefit from the TFS of sounds at low frequencies. The majority of the HI listeners in the studies described above were over

50, whilst the NH listeners were usually younger than 30, so there is a possible confound of age.

Damage to the hair cells and nerve fibres in the cochlear may result in substantial deficits to processing clearly audible sounds, through reduced frequency selectivity and temporal processing. These deficits may not be related to the extent of threshold elevation, but often occur in people with hearing loss.

1.4.3 **Spatial problems**

Spatial hearing and localisation of sound is an important aspect of the disability felt by hearing impaired listeners (Kramer *et al.*, 1998; Gatehouse and Noble, 2004) and is worth consideration in audiological rehabilitation (Byrne and Noble, 1998). Noble *et al.* (1994) found that HI listeners made more localisation errors than NH listeners. Lorenzi *et al.* (1999a) found that HI listeners were poorer than NH listeners at localisation of clicks in noise, particularly when the noise source was 90° left or right of the listener. The deficits of some HI listeners could be modelled in NH listeners by low-pass filtering, suppressing ILDs cues for localisation (Lorenzi *et al.*, 1999b), but the deficits of other HI listeners could not be modelled this way (Lorenzi *et al.*, 1999a). It is possible the latter HI listeners have ITD processing deficits.

Smoski and Trahiotis (1986) and Hawkins and Wightman (1980) found that HI listeners performed more poorly at ITD discrimination than NH listeners at moderate-to-high SPLs. When the stimuli were presented at equal sensation levels (SLs) for NH and HI listeners, Smoski and Trahiotis (1986) found the listener groups performed similarly. Hawkins and Wightman (1980), on the other hand, found HI listeners still performed worse than NH listeners. Both these studies tested HI listeners and NH listeners that were comparable in age, so it is unlikely any age-related temporal processing deficits confounded the results. However, both studies only tested a few listeners, so it is unclear how broadly the results can be generalised to the heterogeneous HI population. The ITDs were present in both TFS and envelope, so it was unclear whether the deficits were specific to TFS processing or not. Lacher-Fougère and Demany (2005) measured TFS-IPD and envelope-IPD thresholds separately using amplitude-modulated (AM) stimuli. They found that TFS-IPD thresholds were poorer for HI listeners compared to NH listeners, but the difference was less marked for the envelope-IPD thresholds. This suggests HI listeners may have a deficit in temporal processing of TFS separate to envelope processing at some stage, but it is not conclusive because the HI listeners' age ranged between 42 and 68 years, whilst the NH listeners' age ranged between 24 and 45 years.

HI listeners achieve less SRM than NH listeners (Duquesnoy, 1983; Marrone *et al.*, 2008a). Whilst the HI listeners reported by Duquesnoy (1983) were substantially older than the NH listeners, Marrone *et al.* (2008a) tested SRM for young NH, young HI, old NH and old HI listeners. In non-reverberant conditions, Marrone *et al.* (2008a) found young NH listeners received the greatest SRM (11 dB), followed by old NH listeners (8 dB), then young HI listeners (5dB), and old HI listeners had the worst SRM (3 dB). SRM was worse for all groups in reverberant conditions, but this affected NH listeners more than HI listeners. Gelfand *et al.* (1988) tested SRTs and SRM for young and old NH listeners and old HI listeners. Whilst the older NH listeners had higher (worse) SRTs than the younger NH listeners, they achieved similar SRM. However, the old HI listeners had higher SRTs than any NH group and achieved less SRM.

Hall *et al.* (1984) suggested that poor BMLD performance for HI listeners may be due to poor coding of the TFS. Monaural measures of intensity, frequency and temporal resolution did not relate well to BMLDs at 500 Hz, but 500 Hz pure-tone ITD discrimination did. The fact that HI listeners' BMLDs were poorer in wide-band noise than narrow-band noise suggests problems in across-frequency temporal processing may exist for these listeners (Hall *et al.*, 1984). Goverts and Houtgast (2010) showed that SRTs in speech-shaped noise were higher for HI listeners than NH listeners, but binaural unmasking due to inverting the speech signal in one ear (like BMLDs) was normal for 17 out of 25 HI listeners. Goverts and Houtgast (2010) tried to determine in which domains—phase, frequency, time, or intensity—the HI listeners had processing deficits. This was done by measuring sensitivity of SRTs and binaural unmasking to distortions of the stimuli in each domain. The 8 HI listeners with poorer-than-normal binaural unmasking improvements were least sensitive to phase and time distortions, which implies the coding of these domains is deficient in these HI listeners. This supports the idea that binaural unmasking takes advantage of IPD information. Without binaural information, NH and HI listeners' SRTs were sensitive to time distortions, but HI listeners' SRTs were less sensitive than normal to phase, intensity, and especially frequency distortions (Goverts and Houtgast, 2010). However, the distortion–sensitivity approach is limited by the interdependency of distortions in the frequency, phase and time domains, so concluding only one domain is affected by hearing loss would be presumptuous. Monaural and binaural processing appear to be affected by hearing loss differently, with binaural unmasking depending on the nature of the loss (Goverts and Houtgast, 2010). Binaural unmasking appears to be mostly affected by temporal deficits, presumably deficits in processing ITDs or IPDs (Hall *et al.*, 1984; Goverts and Houtgast, 2010).

Hearing loss appears to affect the ability to use spatial auditory cues in some cases, but not all. It may be in part due to asymmetrical thresholds and monaural damage that affects ILDs, but some HI listeners clearly have an extra, and possibly independent, binaural temporal processing deficit. Together, the monaural and binaural deficits have a detrimental impact on perceptually separating out different talkers in noisy, reverberant conditions.

1.5 Age

The association between aging and particular forms of hearing loss (particularly high-frequency sensorineural hearing loss) has been known for many decades (Schuknecht, 1955; Hinchcliffe, 1959; Schuknecht, 1964). In a population survey, Davis (1989) found that the prevalence of self-reported hearing difficulty, both in noise and in quiet, increases monotonically with age. Additionally, the prevalence of threshold elevation approximately doubled with each increasing ten-year age band. Above 2 kHz, threshold elevation becomes progressively more pronounced with increasing age and low frequency thresholds also increase beyond the age of about 60 (Morrell *et al.*, 1996). A longitudinal study of pure-tone thresholds showed how the change in threshold per year increases with increasing frequency, and also with increasing age at baseline (Lee *et al.*, 2005). Importantly, the rate of change in threshold did not significantly differ between those with a self-reported history of noise exposure and those without. Therefore, Lee *et al.* (2005) could not conclude that ARHL results from noise-induced damage. However, noise-induced damage accumulated throughout life may cause supra-threshold deficits in old age (Kujawa and Liberman, 2006).

1.5.1 Supra-threshold changes to hearing related to age

Sommers *et al*. (2011) studied how comprehension of passages of speech changes across lifespan. The passages ranged from 2 to 10 minutes long, presented at 62 dB SPL in quiet. They tested 433 listeners from 20 to 89 years old with hearing thresholds within the range for which no impairments are reported by that age group (Morrell *et al.*, 1996). Audiometric sensitivity diminished progressively across the age groups, but speech comprehension did not follow the same progression. Instead, comprehension accuracy remained stable across listener age until 65 years, then performance dropped markedly. Sommers *et al.* (2011) argue this may be because hearing loss at low frequencies only becomes prominent in the 70-79 and 80-89 age groups, or possibly because middle-aged listeners can compensate for sensitivity loss with semantic context (e.g, Pichora-Fuller *et al.*, 1995). Comprehension required understanding the passages holistically and remembering details in order to answer questions regarding the content, rather than the ability to simply recall the words. It is possible cognitive factors impaired the oldest listeners ability to retain and use semantic context, as the ability to remember spoken sentences decreases as human age (Gilchrist *et al.*, 2008). As the passages were presented in quiet, it is unlikely that performance related specifically to the ability to use the TFS, which is more useful for unmasking speech from modulating sounds (e.g., Moore and Glasberg, 1987a; Lorenzi *et al.*, 2006; Hopkins and Moore, 2009).

Understanding speech in noisy environments becomes more difficult with age. This difficulty is related to threshold elevation in ARHL (see review by Turner *et al.*, 1996), but is unlikely to be solely due to threshold elevation (Frisina and Frisina, 1997; Souza *et al.*, 2007). Even with similar audiometric thresholds, large individual differences are observed in listeners' abilities to comprehending speech in noisy backgrounds (Plomp and Mimpen, 1979). Age-related difficulties in hearing may be reflected in performance on various psychoacoustic tasks that measure dimensions of hearing such as temporal resolution (Grose *et al.*, 2006), frequency selectivity (Patterson *et al.*, 1982) and localisation of sounds (Eddins and Hall, 2010).

As well as threshold elevation, damage to the OHCs with age leads to reduced frequency selectivity (Patterson *et al.*, 1982; Sommers and Humes, 1993; Hopkins and Moore, 2011). Patterson *et al*. (1982) found the derived filters broadened progressively with age, which Sommers and Humes (1993) suggested was likely related to the listeners' hearing loss. Increased prevalence of damaged OHCs with age is consistent with these data. Age and frequency selectivity were not significantly correlated for old NH listeners (Sommers and Humes, 1993). However, Patterson *et al*. (1982) found that filter bandwidths broadened as a linear function of age, whereas audiometric thresholds remained relatively stable until 55 years, then increased rapidly with age.

Ageing also leads to changes in the temporal processing resolution of the auditory system. Older listeners have larger gap-detection thresholds than young listeners, despite closely matched audiometric thresholds within the frequency region encompassing low-pass noise stimuli (Snell, 1997). Using pure-tone bursts, Moore *et al*. (1992) also found some, but not all, older listeners had large gap-detection thresholds. Strouse *et al*. (1998) studied both monaural and binaural temporal processing in young and elderly listeners that were closely matched on audiometric thresholds. Monaural processing acuity was determined using gap-detection thresholds. The acuity of binaural temporal processing was assessed through discrimination thresholds of ITDs in click trains. Elderly listeners performed worse on both measures, particularly when the SL was low. Whilst gap detection and ITD discrimination were correlated in young listeners, they were not in old listeners, suggesting that age affects temporal resolution differently in monaural and binaural processing. Neither measure correlated with performance on the task of discriminating consonant-vowel syllables (e.g. /ba/ from /pa/), despite older listeners also performing worse on speech perception tasks overall. Strouse *et al*. (1998) argue that their findings suggest temporal processing deficits of elderly listeners cannot be wholly attributed to peripheral hearing loss.

Peripheral damage (such as to OHCs) may distort sounds for old people, making speech perception more difficult. However, other complications, such as temporal processing and cognitive capacity deficits, that are not necessarily a consequence of peripheral damage, may also cause problems for old people, particularly in noisy backgrounds.

1.5.2 **Localisation and age**

Localisation appears to decline with age from 40 years old, even when hearing thresholds are taken into account (Abel *et al.*, 2000). Abel *et al*. (2000) found localisation accuracy deteriorated most with age for noise bands centred at 500 Hz, and noted this is within the frequency range where the ITD cues are considered most useful. Age has not been associated with deterioration in ILD sensitivity (Eddins and Hall, 2010). For example, Herman *et al*. (1977) found that old listeners required greater ITDs than young listeners to correctly lateralise a click to the left or right. Similar ILDs were needed for the two age groups to correctly lateralise clicks. Furthermore, Babkoff *et al*. (2002) found that pure-tone and click lateralisation as a function of ILD did not change significantly with age. Equivalent lateralisation functions of ITDs became shallower with age; changes in ITD produced 1.3 times smaller changes in perceived lateralisation at 71 years than at 25 years old.

People can suppress perception of acoustic reflections, which may otherwise mislead localisation, in favour of localisation cues in the direct wave-front. This is the precedence effect (Wallach *et al.*, 1949). Cranford *et al*. (1993) studied the ability of young-NH, young-HI, old-NH and old-HI listeners to localise clicks from loudspeakers equidistant from each ear. The clicks came in pairs—one from each loudspeaker—with inter-speaker delays between 0.1 and 8 ms. For delays below 0.7 ms, age and hearing loss both impaired correct identification of the leading speaker, although age had the greater effect. For longer delays, there was little difference in performance of the four groups, suggesting age and hearing loss did not affect the precedence effect. At 0.3 to 0.5 ms delays, the young HI listeners performed better than the elderly NH listeners, despite having hearing losses of 30 dB. This may be because the clicks were presented at fixed SLs. Presentation stimuli in SL reduces differences between NH and HI listeners in ITD discrimination compared to fixed SPL presentation (Smoski and Trahiotis, 1986). Cranford *et al.* (1993) claimed the systematic pattern of localisation errors in all four groups suggested a sensory problem rather than a cognitive problem (which could produce a more random pattern of errors). Also, if the age-related deficits had cognitive causes, they may be expected to also impair the precedence effect (Cranford *et al.*, 1993).

Ross *et al*. (2007) studied the threshold of detecting a 180° change in IPD in the TFS of an AM tone as a function of the carrier frequency and listener age. Young adults could detect the IPD switch up to around 1.2 kHz, middle aged and elderly listeners could do so only up to around 700 Hz and 600 Hz respectively. The middle aged and elderly listeners' thresholds were much more variable, some performing no better than chance. Ross *et al*. (2007) also found that the cortical EEG response to the

IPD switch diminished at similar carrier frequencies with age as the behavioural thresholds. Grose & Mamo (2010) repeated the behavioural paradigm used by Ross *et al*. (2007) and found similar frequency-limit thresholds. Grose and Mamo (2010) also measured TFS-IPD discrimination thresholds at fixed frequencies from the same sample of NH listeners. Young listeners were most sensitive to IPDs at 750 Hz and became progressively less sensitive above this frequency. Middle-aged and older listeners performed best at 500 Hz, also performing progressively worse at frequencies above and below this. These results show age-related declines in TFS-IPD sensitivity depend on carrier frequency.

In old age auditory localisation becomes more difficult. This appears to be predominantly due to temporal processing deficits, particularly TFS processing, rather than binaural level processing deficits. The precedence effect does not seem to be as affected by age as the sub-millisecond processing required for using ITDs or IPDs. The interaction between age and carrier frequency suggests age affects processing as a time constraint rather than a phase constraint.

1.5.3 **Aetiologies of age-related auditory deficits**

Cochlear and neural causes of hearing loss are arguably the most common causes associated with age (Schneider, 1997). High-frequency ARHL has been attributed mainly to OHC stereocilia damage, particularly at the basal end (Wright *et al.*, 1987; Soucek and Michaels, 1990). Wright *et al*. (1987) showed that normal IHCs also decline in number with age, at a shallower rate than OHCs. The number of nerve fibres in human temporal bones declines steadily with increasing age of the person at death (Makary *et al.*, 2011). A decline in the number of functional IHCs or afferent AN fibres could deteriorate temporal coding.

The lateral wall of the cochlea, the stria vascularis, deteriorates with age. Less potassium is 'pumped' into the endolymph (Schmiedt, 1996). This results in a relatively flat loss across frequency (Schmiedt, 2010). Potassium is required to depolarize the hair cells when the tip-links on the stereocilia open. The stria vascularis can be thought of as the power supply of the cochlea. Schmeidt (1996) showed that increasing age was correlated with decreasing electric potential in the endolymph in gerbils raised in a quiet environment. Schmeidt (2010) argued that this change can explain hearing loss where OHC loss is minimal. A weaker electric potential may affect the processing of sounds above threshold. Schmiedt (2010) reviewed evidence showing that frequency selectivity is not affected by a weaker endolymph potential, but it is possible that temporal processing is affected. Less strong depolarisation of IHCs may reduce the chance of AN fibres firing.

A third possible cause of age-related temporal deficits is a change to the transmission speeds of potentials along the neurons, which would desynchronise signals in different neurons (Pichora-Fuller *et al.*, 2007). Fast transmission of an action potential along axons is possible because supporting cells wrap sections of the axons in myelin sheaths (Peters, 2002). Ageing changes myelin in several ways including ballooning and hardening, and splitting and enclosing other bodies. Such changes reduce conduction velocity to varying degrees in different neurons, desynchronising potentials that were supposed to terminate synchronously (Peters, 2002). Phase locking and the binaural processes that rely on precise timing information may be hindered as a result. Peters and Sethares (2003) showed that re-myelination occurs in primates, but because these new sections of myelin are shorter than original sections it is unlikely the original conduction speeds are restored.

Pichora-Fuller *et al*. (2007) found that distortion of speech via temporal jitter of the low-frequency part of the signal reduced word identification for young NH listeners. Jittering distorts the TFS, in a manner representative of neural de-synchronisation, with negligible alteration of the temporal envelope or long-term spectrum (Pichora-Fuller *et al.*, 2007). Crucially, when the low-frequency part of the signal was spectrally-distorted (instead of temporally), the same deficits in identification were not observed. This suggests that word identification was not reduced due to any un-intended spectral distortion arising from temporal jitter. Word identification was reduced to similar accuracy to that shown by old, NH listeners (Pichora-Fuller *et al.*, 1995). Anderson *et al*. (2012) found physiological results that supported the de-synchronisation hypothesis for difficulties understanding speech in noise in old age.

Cranford *et al*. (1990) found Multiple Sclerosis (MS) patients were poorer than normal at identifying the leading loudspeaker for inter-speaker delays in clicks below 0.7 ms. MS is a degenerative disease where small regions of neurons, at multiple sites in the nervous system, are demyelinated, and as a result, nerve conduction velocities are reduced (Bear *et al.*, 2007). Most of the MS group reported by Cranford *et al*.

(1990) had normal audiograms and were broadly age-matched to the control group. Performance on the task was negatively correlated with the length of time since the patient had been diagnosed. This suggests that demyelination may deteriorate the neural synchronisation required for localisation and lateralisation.

A fourth possible cause of poor temporal processing in old age is a decline in neural inhibition. The main neurotransmitter responsible for signalling inhibitory responses in neurons is GABA. GABA appears to become depleted in old age (Caspary *et al.*, 1995), which may be the cause of imbalanced excitation and inhibition of envelope coding in the inferior colliculus (Walton *et al.*, 2002). Also, older neurons fire with less accuracy cycle-by-cycle; rapid-onset type neurons have greater latency and a longer recovery period in old age (see Frisina and Walton, 2006). Depleted GABA quantities, or reduced GABA receptor binding, is likely to be responsible for age-related changes in neural temporal acuity (Milbrandt *et al.*, 1994); blocking GABA receptors by an agonist shows similar effects to ageing on phase locking (Koch and Grothe, 1998; Klug *et al.*, 2002). Schatteman *et al*. (2008) suggested that the poorer temporal resolution exhibited behaviourally by older listeners (e.g., Strouse *et al.*, 1998) may be partially due to weaker inhibition in the neural coding of amplitude modulations. It is plausible that this would also affect the ability to take advantage of the sub-millisecond ITDs when localising sound. Neural inhibition is necessary to maintain the precise phase locking in the signals from each ear required for localisation (Brand *et al.*, 2002).

There are multiple potential neurological causes of temporal processing deficits that occur after middle age. It is likely that these are not mutually exclusive. Whether the deficit is due to a weaker endolymph potential, fewer IHCs or AN fibres to encode the signal, jitter introduced into the signal by myelin degradation, poorer inhibitory control of temporal acuity, or a combination of these factors, it is likely to impact negatively on the fine temporal precision required for IPD or ITD sensitivity.

1.6 Conclusion

In conclusion, there is evidence that ageing affects not only absolute thresholds, but also temporal processing of supra-threshold sounds including the binaural temporal processing used in spatial hearing. Hearing loss (*i.e.* elevated absolute thresholds) also affects supra-threshold processing, including frequency selectivity and TFS processing. Both the elderly and the hearing impaired perform worse than young NH listeners on a variety of tasks, particularly ones that rely on the ability to use TFS information. The worst performance is often seen in listeners who are both elderly and hearing impaired. However, few studies have shown the effects of hearing loss on TFS processing are independent of effects of ageing. Furthermore, whilst some HI listeners have poorer spatial hearing than normal, this may be due to asymmetrical hearing loss, rather than temporal processing (Bronkhorst and Plomp, 1989). The extent to which deficits in temporal processing, specifically of TFS, may be responsible for impaired spatial hearing is uncertain, but likely to be significant (e.g., Neher *et al.*, 2011).

The cause of insensitivity to TFS is still not clear. The causes may be one, or a combination, of the following:

i. The phase-locked representation of TFS may be impaired by insufficient functional IHCs or afferent nerve fibres to provide an accurate temporal code.

ii. Age-related neurological changes (such as demyelination) might disrupt the transmission or processing of TFS information.

iii. Changes in neurotransmitters or neural networking may affect the integration of signals that allows the auditory system to take advantage of TFS information.

iv. A change in the tuning of the BM and the phase response due to losing the active response of the OHCs could affect the TFS (although this seems unlikely considering recent evidence, see Strelcyk and Dau, 2009; Hopkins and Moore, 2011).

It is important to determine the extent of TFS deficits in people with CHL, and the implications it has for everyday listening demands, because this will help audiologists to manage patient expectations and guide hearing-aid designers in prioritising the deficits for which devices should compensate.

1.7 Thesis Research Overview

In the following chapters the experiments that were designed to meet the aims of this thesis are described in the format of scientific journal articles. Each chapter contains a single paper. Because of this chapter format, some concepts and literature reviews are repeated in sections (such as the introductions at the start of each chapter).

The papers themselves (published, submitted, or intended for submission) have multiple authors; the first author is always the author of this thesis and the contributions to the research by each author are detailed in the following paragraphs.

The first study (chapter 2) was designed to determine the usefulness of a two-alternative, forced-choice discrimination task to test IPD discrimination thresholds quickly and reliably. This was to test an appropriate task for measuring the IPD thresholds of a large number of listeners. This was expected to be necessary for the second study (chapter 3) where the effects of age and of hearing loss on IPD discrimination were assessed independently via partial correlation. A large sample was needed to adequately represent a large age-range of listeners with a range of hearing loss severities. The third study (chapter 4) tested whether vertically- and horizontally-oriented EEG electrode montages measured FFR with different latencies, in order to make inferences about phase-locking at earlier and later neural stages of the sub-cortical auditory pathway without invasive methods. In the fourth study (chapter 5) vertically- and horizontally-oriented EEG electrode -montages were used to measure later- and earlier-generated FFR in a group of listeners similar to that in the second study. IPD discrimination was also measured. The FFR was used to determine whether the age- and hearing-loss-related effects on IPD discrimination were explained by differences in phase-locking fidelity earlier and later in the auditory system.

In the four studies reported in chapters 2 to 5, the primary author (King) conducted the experiments, analysed the results and drafted the manuscripts. Authors Hopkins and Plack supervised King, advised on study design, analysis and interpretation of the results, and edited the manuscripts.

Chapter 6 explores the use of behind-the-ear hearing aids to process sounds from an acoustic free-field to experimentally manipulate interaural differences, in speech stimuli, in as close to real time as possible. This was expected to reveal how IPDs are used in everyday spatial hearing better than psychoacoustic tests with simple stimuli. This research was conducted at the Eriksholm Research Centre, with the primary author (King) being responsible for the study design, carrying out the psychoacoustic discrimination tests, data analysis and writing the manuscript. King contributed to some of the programmatic setup. Hopkins also contributed to the study design and Plack advised on analysis and manuscript edits. Pontoppidan supervised King as well as overseeing the calibration and verification of the experimental setups. Bramsløw programmed the speech test and advised on the setup of the discrimination tests. Vatti set up and calibrated the master hearing aid system. Heitkamp and Hafez collected the speech test data (in Danish) and performed audiological screening of participants. Heitkamp also contributed to study design.

Chapter 2. Differences in short-term training for interaural phase difference discrimination between two different forced-choice paradigms

Andrew King, Kathryn Hopkins and Christopher J. Plack

School of Psychological Sciences, The University of Manchester, Manchester, M13 9PL

This chapter was published as a letter to the editor in the Journal of the Acoustical Society of America:

King, A., Hopkins, K., and Plack, C. J. (2013). Differences in short-term training for interaural phase difference discrimination between two different forced-choice paradigms. *The Journal of the Acoustical Society of America*, *134*, 2635-2638.

DOI: http://dx.doi.org/10.1121/1.4819116

2.1 Abstract

Improvement in interaural phase difference (IPD) discrimination over two to three hours was compared for two two-alternative forced-choice paradigms: a three-interval paradigm, in which the IPD was in interval two or three, and a paradigm with two intervals of four stimuli in which the IPD was in the second and fourth stimuli of one interval (AAAA vs. ABAB). The difference in performance between the beginning and end of the testing period was smaller for the two-interval paradigm, supporting the use of this paradigm for fast measurement of discrimination thresholds without the need for a long period of training.

2.2 Introduction

Performance on psychophysical discrimination tasks often improves with increasing number of trials (Wright and Fitzgerald, 2001; Hawkey *et al.*, 2004; Amitay *et al.*, 2006), due to learning attributes of the stimuli (perceptual learning) and/or procedural learning. Learning may reduce the accuracy of estimates from repeated measurements, and different learning rates for different conditions, listeners, or tasks can introduce bias. Training participants until performance reaches asymptote can minimize these problems, but this can take thousands of trials (Hafter and Carrier, 1970). Long training periods are not an option if the available time to test each listener is short, for example, in clinical assessment. Hence, it is important to find a procedure that minimizes the change in performance over time.

Hopkins and Moore (2010b) and Moore and Sęk (2009a) described a two-interval discrimination paradigm (8S2A) for which they reported minimal training effects. One interval (chosen at random) contained four non-target stimuli (AAAA). The other interval contained alternating non-target and target stimuli (ABAB). Hopkins and Moore (2010b) found no significant improvement in pure-tone, interaural phase difference (IPD) discrimination over the 16 adaptive tracks for each of the three frequencies that were tested. Moore and Sęk (2009a) found no significant difference in discrimination of harmonic and frequency-shifted band-pass filtered complex tones over two, two-hour training sessions for 10 trained and 10 untrained listeners. However, neither study compared the 8S2A paradigm with any other presentation paradigm, so it is not clear whether the paradigm, or some other aspect of the experimental procedure, produces stable performance.

Here the effect of training on IPD discrimination performance for the 8S2A paradigm was compared to that for the three-interval, two-alternative (3I2A) paradigm described by Lacher-Fougère and Demany (2005).

2.3 Method

2.3.1 **Stimuli**

Three types of stimuli were used: pure tones (PT IPD), amplitude modulated (AM) tones with the IPD in the temporal fine structure (TFS IPD), and AM tones with an IPD in the temporal envelope (Envelope IPD). The AM tones were specified by equation 2.1:

$$s(t) = \sin(2\pi f_c t + \varphi_c) \cdot [1 + \sin(2\pi f_m t + \varphi_m)], \tag{2.1}$$

where f_c is the carrier frequency, f_m is the modulation frequency, φ_c is the carrier phase, φ_m is the modulation phase and t is the sample index. Differences in φ_c and φ_m between ears (referred to here as δ) produced IPDs in the TFS (TFS IPD) and envelope (Envelope IPD), respectively. IPDs were created by introducing a positive starting phase in one ear and zero starting phase in the other. The frequency of the pure tones and f_c for the AM tones was 500 Hz. For the AM tones f_m was 20 Hz. Stimuli were presented at a level of 80 dB sound pressure level (SPL), to allow for comparison with previous studies conducted in the laboratory using the same test. The stimulus duration was 500 ms, including 50-ms raised-cosine onset and offset ramps, which were synchronous across ears. The inter-stimulus silent interval was 500 ms for both paradigms. For the 8S2A paradigm the four stimuli in each interval were separated by 50 ms of silence. Stimuli were created in MATLAB (Mathworks) at a sample rate of 48 kHz and output via a Creative E-MU 0202 USB 24-bit soundcard and Sennheiser HD 650 circum-aural headphones within a double-walled listening booth.

2.3.2 **Procedure**

For the 3I2A paradigm, listeners indicated whether the second or third interval differed from the first interval (a diotic reference stimulus) by pressing ‘2’ or ‘3’ on a computer keyboard. One of the last two intervals (chosen at random) had an IPD of δ; the other was identical to the diotic reference. For the 8S2A paradigm, each interval contained four stimuli. In one interval (chosen at random) the stimuli were all diotic (AAAA). In the other interval, the first and third stimuli in the sequence were also diotic, but the second and fourth contained an IPD of δ (ABAB). Listeners indicated whether the ABAB interval was first or second by pressing ‘1’ or ‘2’ on the keyboard. An onscreen light indicated each interval and other lights provided feedback. At the beginning of each run, δ was set to 180°. A two-down, one-up adaptive method was used to track 70.7% correct on the psychometric function. The initial step size was a factor of 1.25^2. After four reversals, the step size was reduced to 1.25 for a further 10 reversals. The geometric mean of the values of δ at the last 10 reversals was taken as the threshold δ.

Thirty-six normal-hearing listeners (18-35 years, mean=23) were tested (audiometric thresholds <=20 dB HL between 0.25 and 8 kHz, and <10 dB difference in thresholds between ears at 500 Hz). Listeners were randomly allocated into six groups of six, with each group allocated a different combination of paradigm and stimulus. Twenty-two runs were completed over a total period of approximately two hours with the 3I2A paradigm and three hours with the 8S2A paradigm. The majority of listeners tested with the 8S2A paradigm completed the experiment over two sessions whilst the listeners tested with the 3I2A paradigm were more likely to complete the experiment in one session.

2.4 Results

Figure 2.1 shows the geometric mean IPD discrimination thresholds for each group as a function of run number. For analysis, the geometric mean of the first four values (pre training) was compared with the geometric mean of the last four values (post training) for each listener. The listener-group geometric means for these measures are shown in Figure 2.2. All analyses were performed on the log-transformed data values to satisfy the assumption of normality. A mixed model ANOVA was performed with training (pre and post) as the within-subjects factor, and paradigm (3I2A and 8S2A) and training stimulus (envelope IPD, TFS-IPD and PT-IPD) as between-subjects factors. Thresholds were lower following training [$F(1,30)=10.0$, $p=0.004$], and listeners tested using the 8S2A paradigm had lower mean thresholds than listeners tested using the 3I2A paradigm [$F(1,30)=5.3$, $p=0.03$]. There was no significant effect of stimulus.

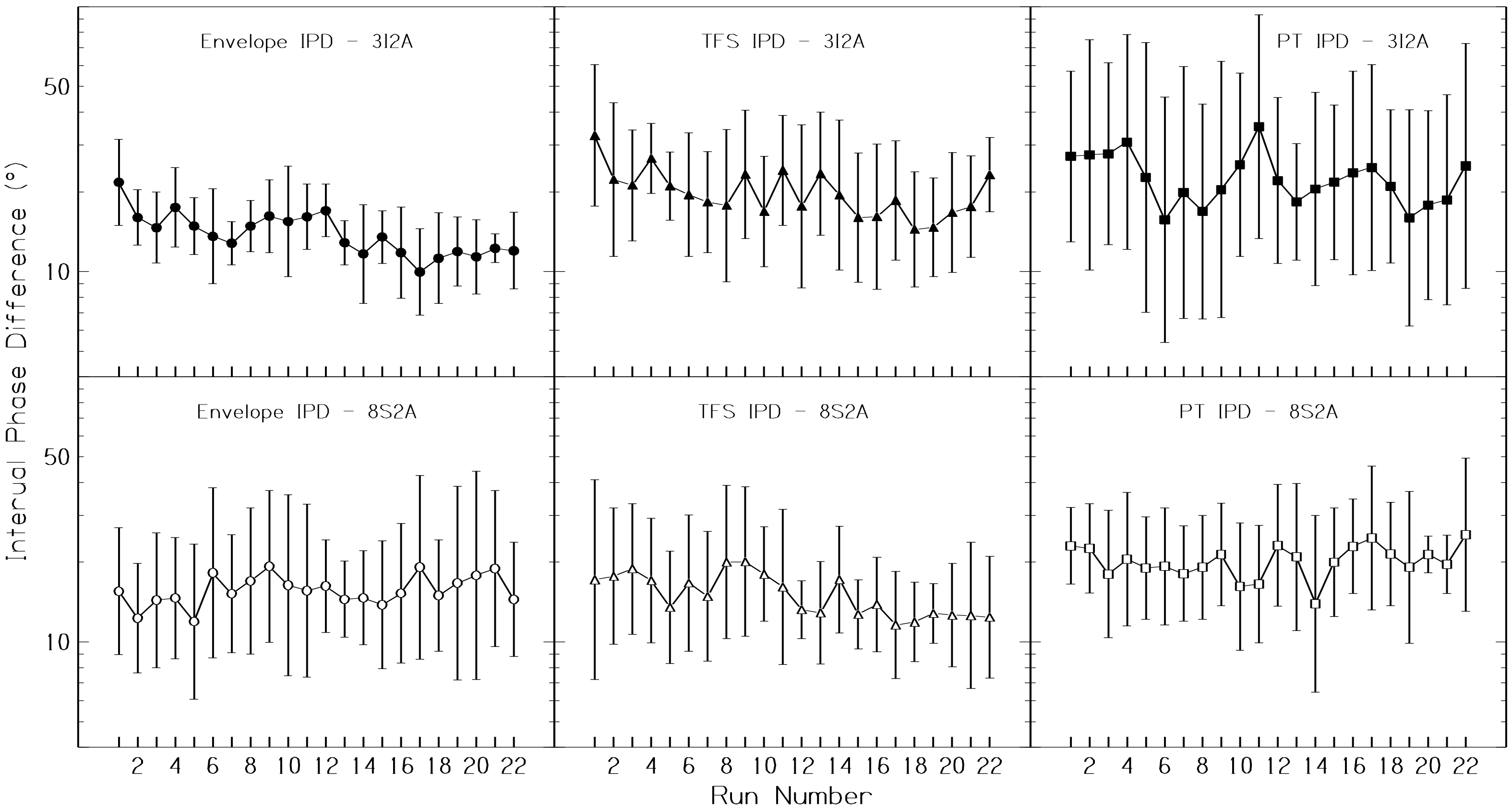

Figure 2.1: The geometric mean IPD discrimination thresholds for each group of six listeners as a function of run number. Error bars show 95% confidence intervals.

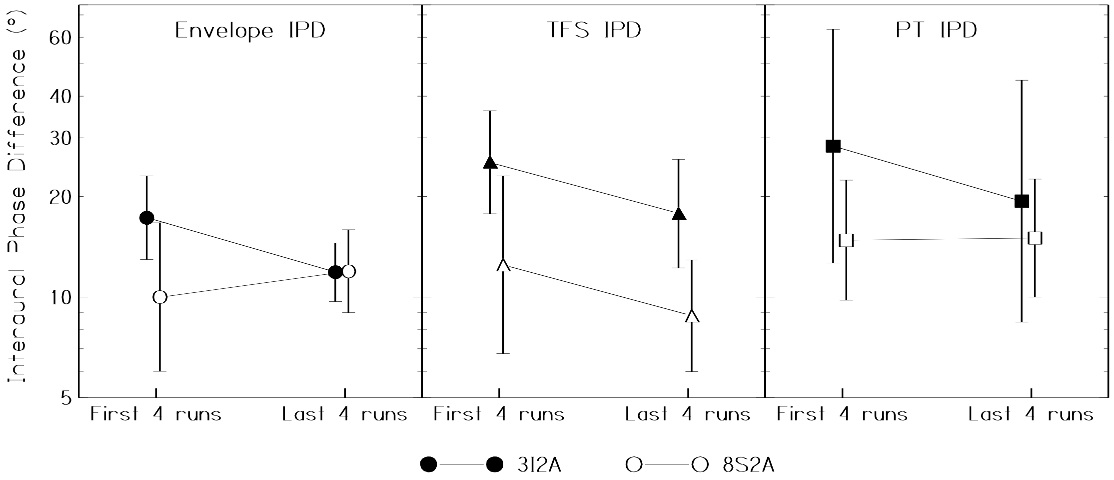

Figure 2.2: The geometric mean IPD discrimination thresholds obtained from the first four adaptive runs and the last four adaptive runs completed in each training condition, averaged across listeners. Data are offset on the abscissa for visual clarity. Error bars show 95% confidence intervals.

The interaction between training and paradigm was significant [*F*(1,30)=5.6, *p*=0.02]. To examine the effect of training for the two paradigms separately, thresholds were collapsed across stimuli and paired t-tests were performed for each paradigm. Thresholds for the two paradigms are shown in the left panel of Figure 2.3 by symbols joined by solid lines. Listeners in the 8S2A paradigm groups improved only slightly (pre=12.3°, SD=1.9°; post=11.6°, SD=1.9°) and the improvement was not significant. Listeners in the 3I2A paradigm groups showed a greater improvement (pre=23.2°, SD=2.0°; post=16.0°, SD=2.0°), which was significant [*t*(34)=3.7, *p*=0.002]. None of the other interactions in the ANOVA were significant (*p*>0.05).

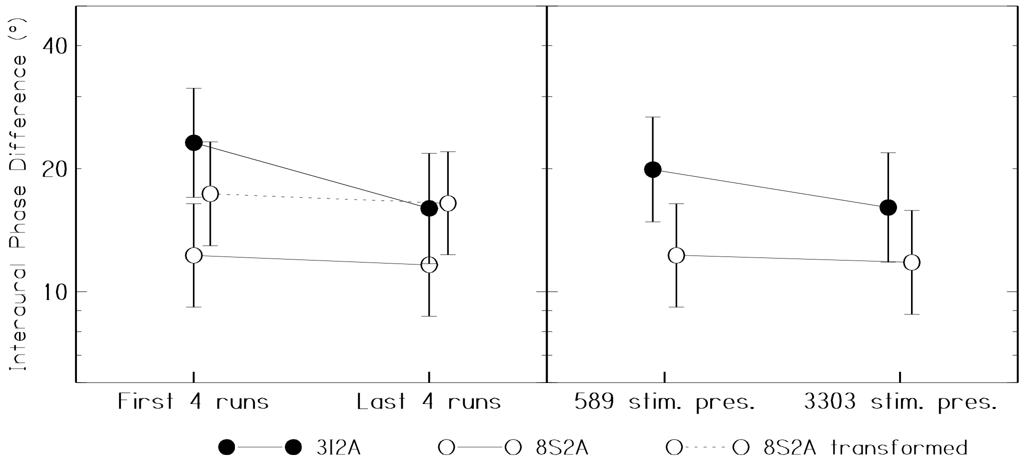

Figure 2.3: In the left panel the symbols joined by solid lines show the geometric means of the first four runs and the last four runs in the two paradigm conditions, averaged across stimuli and listeners. The symbols joined by the dashed line show transformed 8S2A data. The right panel shows the geometric means of the four runs up to and including the run that exceeded 589 and 3303 stimulus presentations, averaged across stimuli and listeners. Data are offset on the abscissa for clarity. Error bars show 95% confidence intervals.

In the first ANOVA, thresholds for the 8S2A paradigm were significantly lower than for the 3I2A paradigm. Both paradigms were two alternative forced-choice tasks, but the 8S2A paradigm had twice as many observations as the 3I2A paradigm that contributed to the discrimination decision (*i.e.* with *a priori* probabilities of being either reference or target). d-prime (d´) increases by the square root of the number of times a trial is repeated before a forced-choice decision is made (Swets *et al.*, 1959) , and this might explain why lower thresholds were observed for the 8S2A paradigm. Assuming independent internal noise and equal variance in the distributions of reference and target stimuli (Green and Swets, 1974), d´ would be expected to be $\sqrt{2}$ times greater for the 8S2A paradigm than for the 3I2A paradigm. d´ increases linearly with IPD (Hafter and Carrier, 1972). Hence, to compensate for this effect, the 8S2A IPD thresholds (calculated over four runs) were multiplied by $\sqrt{2}$ before being log-transformed. The dashed line in Figure 2.3 shows the mean transformed δ. Notice that pre-training performance for the 8S2A paradigm was similar to post-training performance for the 3I2A paradigm. The mixed model ANOVA was repeated using these transformed 8S2A thresholds. The main effect of paradigm was not significant. The main effects of training and stimulus, and the interaction effects were unchanged.

The 3I2A paradigm had only three stimuli per trial and the 8S2A paradigm had eight. To take account of this difference, non-transformed mean thresholds were compared after approximately equal numbers of stimulus presentations. Pre training was defined as the runs up to and including the run in which listeners had heard 589 stimulus presentations (the mean number presented in the first four runs of the 8S2A paradigm) and post training was defined as the four runs up to and including 3303 stimulus presentations (the smallest total number of stimuli heard by all listeners). The right panel of Figure 2.3 shows these data for the two paradigms. The first mixed-model ANOVA was repeated using these threshold estimates. The main effect of training was significant [F(1,30)=11.06, p=0.002]. However, there were no significant main effects of paradigm or of stimulus. The interaction between training and paradigm was significant [F(1,30)=5.09, p=0.032]. All other interactions were non-significant.

The difference between thresholds at the end of session one and at the start of session two (Mean=−1.0°, SD=1.9°), for those who completed the study in two sessions, was not significantly greater than the difference between the 11th and 12th runs (Mean=−0.8°, SD=2.0°) for those who completed the study in one session (t=−0.76, p=0.45).

Observed power was calculated, in G-Power 3.1.3 (Kiel, Germany), to be 0.98, based on the mean partial η^2 of all the effects tested with the first ANOVA as an effect size, 36 listeners and an α of 0.05.

2.5 Discussion

The PT-IPD thresholds collected using the 8S2A paradigm were 25% lower than the thresholds collected by Hopkins and Moore (2010b) at 50 dB SL (the most similar condition to the PT-IPD condition tested here). The higher SLs in the current study may explain the lower IPD thresholds, but Hopkins and Moore (2010b) report no evidence of a stimulus level effect between 30 and 50 dB SL. The Envelope- and TFS-IPD thresholds collected using the 3I2A paradigm were approximately 50% and 470% higher, respectively, than the thresholds reported by Lacher-Fougère and Demany (2005) using similar stimuli. The current study and Hopkins and Moore (2010b) used inexperienced listeners, but Lacher-Fougère and Demany (2005) do not report listener experience. Therefore, previous experience may have been a contributing factor to the differences in IPD thresholds. Lacher-Fougère and Demany (2005) tested listeners over four, 1 h sessions, whereas the current study and Hopkins and Moore (2010b) only used two sessions. This may affect training, but in the current study there were no significant differences in thresholds between listeners who completed the study in two sessions and those who completed it in one session.

The smaller effect of training for the 8S2A paradigm than the 3I2A paradigm supports the use of the 8S2A paradigm for testing IPD discrimination thresholds (and potentially thresholds for other discrimination tasks) if stable performance is desired over a two- or three-hour time frame. Stable performance in the 8S2A paradigm may have occurred for two reasons. One explanation is that the task was very easy to learn, so either no training was needed to reach asymptote, or performance reached its asymptote within the first four runs. The alternative explanation is that insufficient training was provided for substantial improvement to occur during the course of the experiment. When adjusted for the greater total target duration, the 8S2A thresholds pre training were similar to the 3I2A thresholds post training. This suggests that performance asymptotes quickly for the 8S2A paradigm. The lack of training effect for the 8S2A paradigm is consistent with Moore and Sęk (2009a) and Hopkins and Moore (2010b).

In conclusion, IPD thresholds may not be comparable across the two paradigms without accounting for the different numbers of stimulus presentations per trial. However, the 8S2A paradigm appears better than the 3I2A paradigm for avoiding training effects and the potential variability and bias they may produce over the course of several hours of testing. Hence, the 8S2A paradigm may be more appropriate for fast assessment of discrimination abilities.

2.6 Acknowledgments

This work was supported by Oticon A/S and the Medical Research Council (UK) (ref G1001609).

Chapter 3. The effects of age and hearing loss on interaural phase difference discrimination

Andrew King, Kathryn Hopkins, and Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester, UK

Portions of this chapter were presented at the British Society of Audiology Conference, Nottingham, UK, September 2012. This chapter was published as an original research article in the Journal of the Acoustical Society of America:

King, A., Hopkins, K., and Plack, C. J. (2014). The effects of age and hearing loss on interaural phase difference discrimination[a)]. *The Journal of the Acoustical Society of America*, *135*, 342-351.

DOI: http://dx.doi.org/10.1121/1.4838995

3.1 Abstract

The discrimination of interaural phase differences (IPDs) requires accurate binaural temporal processing, and has been used as a measure of sensitivity to temporal envelope and temporal fine structure (TFS). Previous studies found that TFS-IPD discrimination declined with age and with sensorineural hearing loss (SNHL), but age and SNHL have often been confounded. The aim of this study was to determine the independent contributions of age and SNHL to TFS and envelope IPD discrimination by using a sample of adults with a wide range of ages and SNHL. A two-interval, two-alternative forced-choice procedure was used to measure IPD discrimination thresholds for 20-Hz amplitude-modulated tones with carrier frequencies of 250 or 500 Hz when the IPD was in either the stimulus envelope or TFS. There were positive correlations between absolute thresholds and TFS-IPD thresholds, but not envelope-IPD thresholds, when age was accounted for. This supports the idea that SNHL affects TFS processing independently to age. Age was positively correlated with envelope-IPD thresholds at both carrier frequencies and TFS-IPD thresholds at 500 Hz, when absolute thresholds were accounted for. These results suggest that age negatively affects the binaural processing of envelope and TFS at some frequencies independently of SNHL.

3.2 Introduction

The auditory system can discriminate interaural time differences (ITDs) in the arrival of sounds (Klump and Eady, 1956) or interaural phase differences (IPDs) if the sounds are periodic and on-going (Zwislocki and Feldman, 1956). These cues are used for lateralisation and localisation (Wightman and Kistler, 1992). Discrimination of ITDs or IPDs has been used as a way of measuring temporal coding ability, because the coding of these cues relies on the accurate synchronisation of neural activity to the stimulus waveform (Jeffress, 1948). For low frequency sounds, auditory nerve fibres are most likely to fire at a particular phase of basilar membrane (BM) motion (Tasaki, 1954; Palmer and Russell, 1986), a phenomenon known as phase locking (Rose *et al.*, 1967). Phase locking codes time intervals between corresponding peaks in the pass-band filtered output from the BM, which represent the temporal fine structure (TFS) of the sound. TFS coding is thought to contribute to accurate pitch discrimination (Moore *et al.*, 2006a), speech perception (Young and Sachs, 1979) and perceptual segregation of target sounds, such as speech, from complex background sounds (Hopkins and Moore, 2009; Moore, 2012).

Sensorineural hearing loss (SNHL) is associated with poorer performance on tasks that are thought to provide behavioural measures of TFS coding (e.g., Buss *et al.*, 2004; Lacher-Fougère and Demany, 2005; Hopkins and Moore, 2007; Strelcyk and Dau, 2009; Hopkins and Moore, 2011). Early studies (Hawkins and Wightman, 1980; Buus *et al.*, 1984; Smoski and Trahiotis, 1986) found that listeners with SNHL were poorer at lateralisation based on ITDs than listeners with normal hearing (NH). However, these ITDs were implemented by delaying the whole waveform to one ear, so deficits may have arisen due to impaired coding of either TFS or slower fluctuations in amplitude caused by the interaction of TFS components (commonly referred to as the envelope) or both. Later research investigated sensitivity to envelope and TFS IPDs separately using amplitude modulated (AM) tones (Lacher-Fougère and Demany, 2005) and sensitivity to TFS-IPDs exclusively using pure tones (Hopkins and Moore, 2011).

Both Lacher-Fougère and Demany (2005) and Hopkins and Moore (2011) reported better TFS-IPD sensitivity for NH listeners than for those with SNHL. Hopkins and Moore (2011) found that the TFS-IPD thresholds of SNHL listeners were between 1.5 and two times those of NH listeners, whilst Lacher-Fougère and Demany

(2005) found a 6.5- to 19.7-fold deficit for SNHL listeners. Lacher-Fougère and Demany (2005) found that envelope-IPD thresholds were also greater for SNHL listeners than for NH listeners, but only by 2.9 to 4-fold. Lacher-Fougère and Demany (2005) interpreted the larger deficit in TFS-IPD thresholds than envelope-IPD thresholds as evidence that SNHL specifically affects TFS processing. Lacher-Fougère and Demany (2005) used the same sound pressure level (SPL) for all listeners, so the sensation levels (SL) of the stimuli would be lower for the SNHL listeners than the NH listeners. Buus *et al.* (1984) and Smoski and Trahiotis (1986) showed envelope ITD discrimination was affected by SL whilst pure-tone ITD discrimination was not, so Lacher-Fougère and Demany (2005) suggested the differing SL may have affected the SNHL listeners' envelope-IPD thresholds more than their TFS-IPD thresholds.

The deficit in TFS-IPD sensitivity for SNHL listeners reported by Lacher-Fougère and Demany (2005) and Hopkins and Moore (2011) may be partly explained by the higher mean ages in the SNHL listener groups than the NH listener groups. The age ranges were 24 to 45 years for NH listeners and 42 to 68 years for SNHL listeners in the study of Lacher-Fougère and Demany (2005), and 20 to 35 years for NH listeners and 29 to 82 (mean=62.8) for SNHL listeners in the study of Hopkins and Moore (2011). Age is associated with a decrease in the highest carrier-tone frequency (f_c) at which a 180° IPD in the TFS of binaurally presented AM tones is detectable by listeners with minimal hearing loss (Ross *et al.*, 2007; Grose and Mamo, 2010). Ross *et al.* (2007) suggested that their results were due to a loss of neural synchrony with age, which would degrade the precision of temporal coding. Consistent with this idea, low-frequency pure-tone IPD thresholds increase with age (Grose and Mamo, 2010; Moore *et al.*, 2012a; Moore *et al.*, 2012b) and performance on other measures of temporal coding also declines with age (e.g. Strouse *et al.*, 1998; Purcell *et al.*, 2004). In order to assess the effect of age on temporal coding, Hopkins and Moore (2011) included a second sample of NH listeners, with a similar mean age to the SNHL listeners. The age-matched NH group did not perform significantly differently to the SNHL listeners. Whilst this showed that age can affect TFS IPD discrimination, it was not possible to assess the independent effects of age and hearing loss on IPD discrimination as age and hearing loss were highly correlated.

Hawkins and Wightman (1980) and Smoski and Trahiotis (1986) found poorer ITD sensitivity for SNHL and NH listener groups that were comparable in age.

Hawkins and Wightman (1980) used NH and SNHL listeners with mean ages of 25 and 27 years, respectively, and Smoski and Trahiotis (1986) used NH and SNHL listeners with mean ages of 24 and 36 years respectively. However, Hawkins and Wightman (1980) only used three NH listeners and eight SNHL listeners and Smoski and Trahiotis (1986) only used two NH listeners and four SNHL listeners. There appears to be at least some non-age-related effect of SNHL on binaural TFS processing.

This paper reports the TFS- and envelope-IPD thresholds for listeners across a wide age range with normal hearing up to moderate SNHL. Partial correlations were used to assess the effects of age and SNHL independently by removing the variability associated with one variable when assessing the other. AM tones were used, like Lacher-Fougère and Demany (2005), to measure TFS and envelope IPD thresholds separately. However, equal SL across listeners was used rather than a fixed SPL to avoid level affecting IPD thresholds (Buus *et al.*, 1984; Smoski and Trahiotis, 1986).

3.3 Methodology

3.3.1 **Listeners**

Forty-six listeners were tested. Their ages ranged from 18 to 83 years and they had either normal hearing or SNHL as confirmed by air- and bone-conduction pure-tone audiometry (AC- and BC-PTA respectively), tested in accordance with the British Society of Audiology (2011) recommended procedure. Listeners with suspected conductive hearing loss, or asymmetry between ears greater than 15 dB below 1 kHz, were excluded. Table 3.1 lists the listeners by ascending age, with each listener's AC-PTA averaged from 2 to 8 kHz in dB hearing level (HL) ($\mathrm{PTA}_{\mathrm{HF}}$) given also. $\mathrm{PTA}_{\mathrm{HF}}$ was used to estimate the influence of high-frequency hearing loss on IPD sensitivity. There was a significant positive correlation between age and $\mathrm{PTA}_{\mathrm{HF}}$ (r=0.439, p=0.002).

Table 3.1: The 46 listeners listed by their age. Absolute thresholds at 250 and 500 Hz (AT250 and AT500, respectively) and mean audiometric threshold between 2 and 8 kHz (PTA_{HF}) are given for left (L) and right (R) ears. Preference for equal SL or equal SPL across ears in IPD stimulus presentation and which ear contained the positive starting phase in IPD discrimination are given for f_c=250 Hz and f_c=500 Hz.

Listener	Age (years)	AT250 (dB SPL)		AT500 (dB SPL)		PTAHF (dB HL)		IPD stimulus presentation		Leading Ear	
		L	R	L	R	L	R	250 Hz	500 Hz	250 Hz	500 Hz
1	18	44.0	41.8	41.1	35.1	46.7	43.3	SPL	SPL	R	R
2	20	15.9	13.3	6.1	10.9	6.7	8.8	SL	SPL	R	R
3	21	29.3	31.8	36.4	39.0	65.0	73.3	SL	SL	L	L
4	21	21.0	32.0	15.6	23.4	8.3	11.7	SL	SPL	L	L
5	21	20.0	17.5	10.4	12.1	-1.7	3.3	SL	SPL	R	R
6	22	19.6	14.6	16.6	9.1	5.0	5.0	SL	SL	R	R
7	23	10.0	14.1	5.8	1.0	8.3	6.7	SPL	SPL	L	R
8	23	16.6	9.9	5.4	3.5	10.0	5.0	SPL	SPL	R	R
9	25	25.4	25.5	28.0	28.6	30.0	23.3	SL	SL	L	L
10	25	14.9	13.5	12.1	10.8	0.0	1.7	SL	SPL	R	R
11	26	5.4	5.8	-1.3	2.6	6.7	5.0	SL	SL	L	L
12	27	6.8	9.6	0.4	3.3	20.0	10.0	SPL	SL	L	L
13	27	45.0	43.1	41.1	36.1	53.3	48.3	SPL	SPL	R	R
14	28	3.6	7.8	-1.4	3.1	3.3	41.3	SL	SL	L	L
15	28	18.6	18.4	14.0	6.9	16.7	5.0	SL	SL	R	R
16	31	36.4	39.4	55.4	55.4	63.3	65.0	SPL	SPL	R	R
17	31	7.1	9.0	6.3	5.3	3.3	1.7	SPL	SPL	L	L
18	38	73.0	71.1	68.3	63.9	46.7	45.0	SL	SPL	R	R
19	40	16.8	24.0	15.4	19.8	15.0	23.3	SPL	SPL	L	L
20	43	36.6	38.6	45.4	47.6	78.3	73.3	SPL	SL	L	L
21	45	13.4	11.9	1.9	1.8	-3.3	5.0	SL	SL	R	R
22	46	40.3	39.4	44.1	41.6	60.0	56.7	SL	SPL	R	R
23	48	39.4	41.5	40.3	37.3	65.0	66.7	SPL	SL	R	R
24	48	12.9	8.6	8.3	1.8	18.3	10.0	SL	SL	R	R
25	52	12.6	24.8	12.6	17.1	33.8	35.0	SL	SL	L	L

Listener	Age (years)	AT250 (dB SPL)		AT500 (dB SPL)		PTAHF (dB HL)		IPD stimulus presentation		Leading Ear	
		L	R	L	R	L	R	250 Hz	500 Hz	250 Hz	500 Hz
26	52	17.9	16.1	13.6	13.4	5.0	8.3	SPL	SL	R	R
27	56	21.6	21.4	13.8	17.9	33.8	28.3	SPL	SL	L	L
28	65	21.5	27.0	19.8	20.8	40.0	18.3	SPL	SPL	L	L
29	65	17.4	18.4	4.5	5.9	57.5	57.5	SL	SL	L	L
30	66	20.1	23.0	20.6	29.9	23.3	31.7	SL	SL	L	L
31	66	20.5	29.8	20.6	22.1	27.5	26.7	SPL	SL	L	L
32	67	12.6	15.3	15.1	14.6	45.0	25.0	SPL	SL	L	R
33	67	63.0	52.4	59.4	51.1	66.7	71.7	SL	SL	R	R
34	68	13.4	12.4	11.4	18.9	45.0	38.3	SL	SPL	R	R
35	69	21.4	21.1	11.6	11.9	36.7	21.7	SL	SL	R	L
36	69	16.0	12.1	12.9	8.3	53.8	27.5	SL	SL	R	R
37	71	7.0	10.3	-0.6	8.4	36.7	31.7	SPL	SPL	L	L
38	72	20.1	21.5	24.0	29.4	47.5	50.0	SPL	SPL	L	R
39	73	27.3	23.6	22.9	18.0	36.3	36.7	SPL	SL	R	R
40	74	14.6	17.1	15.6	17.8	36.7	38.3	SPL	SPL	L	L
41	76	20.4	22.9	20.5	22.9	13.3	16.7	SL	SL	L	L
42	80	14.8	16.5	10.8	10.1	35.0	28.8	SL	SL	L	R
43	81	35.5	36.8	35.8	33.6	61.7	66.7	SL	SL	L	R
44	82	18.5	23.6	8.9	14.6	56.7	53.3	SL	SL	L	L
45	82	62.5	62.0	56.6	56.3	71.7	75.0	SPL	SPL	L	L
46	83	20.1	17.5	19.5	10.4	58.8	55.0	SPL	SPL	R	R

3.3.2 **Absolute thresholds**

Absolute thresholds (ATs) in dB SPL were measured in order to set the level for the IPD discrimination task at 30 dB SL for each listener.

3.3.2.1 Stimuli

The stimuli were pure tones with a 200-ms steady state duration and 20-ms raised-cosine onset and offset ramps. Frequencies of 250 and 500 Hz were used, which corresponded to the f_c of the stimuli used in the IPD sensitivity test (see section 3.3.4). ATs were determined separately for each ear.

3.3.2.2 Procedure

A three-interval, three-alternative forced-choice task was used with a two-down, one-up adaptive procedure. The step size was 4 dB until three turn points occurred and decreased to 2 dB for a subsequent eight turn points. The threshold corresponding to 71% correct (Levitt, 1971) was estimated as the arithmetic mean of the stimulus level at the last eight turn points. Two runs were completed for each ear at each frequency and the final threshold was taken to be the mean of the thresholds from these two runs. These mean thresholds are given in Table 3.1. Listener's age and average AT over both f_c 250 and 500 Hz were not significantly correlated (r=0.076, p=0.615). The average AT is plotted against age in Figure 3.1.

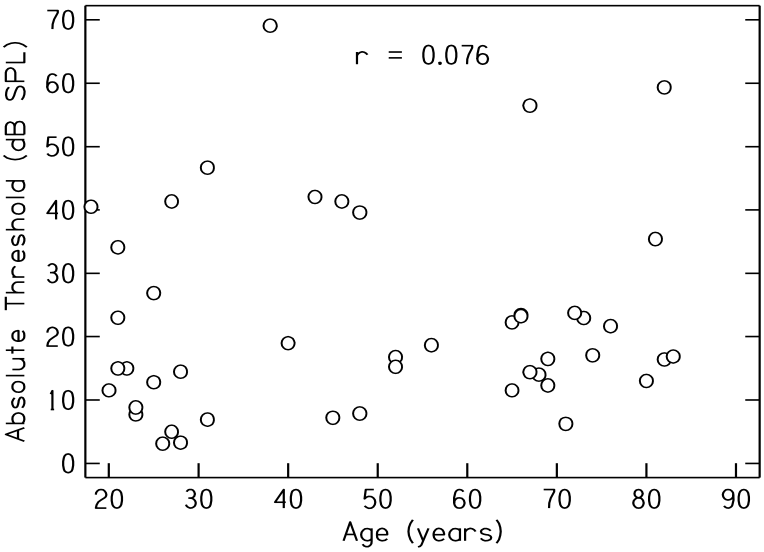

Figure 3.1: Listeners' ATs (averaged across 250 and 500 Hz and ears) as a function of their age.

3.3.3 **Setting AM tone presentation level**

Hopkins and Moore (2010b) showed that pure-tone IPD discrimination was independent of level for levels of 30 dB SL or greater. However, for some listeners, presenting the AM tones at 30 dB SL in each ear resulted in a strongly left or right lateralized sound image. To obtain a stimulus level that resulted in a sound image positioned roughly in the centre of the listener's head, participants were asked to compare AM tones at 30 dB SL at each ear (Equal SL) and at 30 dB above the average of the left and right AT at the f_c (Equal SPL). Whichever version of the stimulus the listener reported as sounding more centred between the ears was used for the IPD sensitivity test described in section 3.3.4. The AM tones were played for 4 s to the listener for both level settings in a random order. Table 3.1 lists, at each f_c, whether Equal SL or Equal SPL was used for each listener.

3.3.4 **IPD sensitivity test**

3.3.4.1 Stimuli

Sensitivity to IPDs was measured using AM tones. Carrier tones of f_c=250 Hz and 500 Hz were amplitude modulated at 20 Hz. The IPD was created in either the TFS or the envelope by introducing a positive starting phase (δ°) in the signal to one ear and a zero starting phase to the other ear. TFS and envelope IPDs are shown schematically in panels A and B of Figure 3.2, respectively. Thus, there were four conditions: non-zero IPDs in the TFS at f_c=250 Hz (TFS250), in the TFS at f_c=500 Hz (TFS500), in the envelope at f_c=250 Hz (ENV250) and in the envelope at f_c=500 Hz (ENV500).

3.3.4.2 Procedure

IPD discrimination thresholds were measured four times for each condition using a procedure based on that described by Hopkins and Moore (2010b). A two-interval, two-alternative forced-choice task was used, with each interval comprising four 500-ms tone bursts (which included 50-ms raised-cosine onset and offset ramps that were synchronous across ears). The tone bursts were separated by 20-ms of silence within each interval and 500-ms of silence between the two intervals. In one interval the four tones all had a zero IPD (AAAA), whilst in the other interval the second and fourth tones had a non-zero IPD (ABAB). The two intervals were randomly ordered, and listeners were instructed to pick the alternating interval. Panel C of Figure 3.2 shows a schematic example of this when the ABAB interval is second. Listeners were advised to focus on lateral position alternation, but that they were free to use any perceptual cue to perform the task. Feedback was given by lights on a screen.

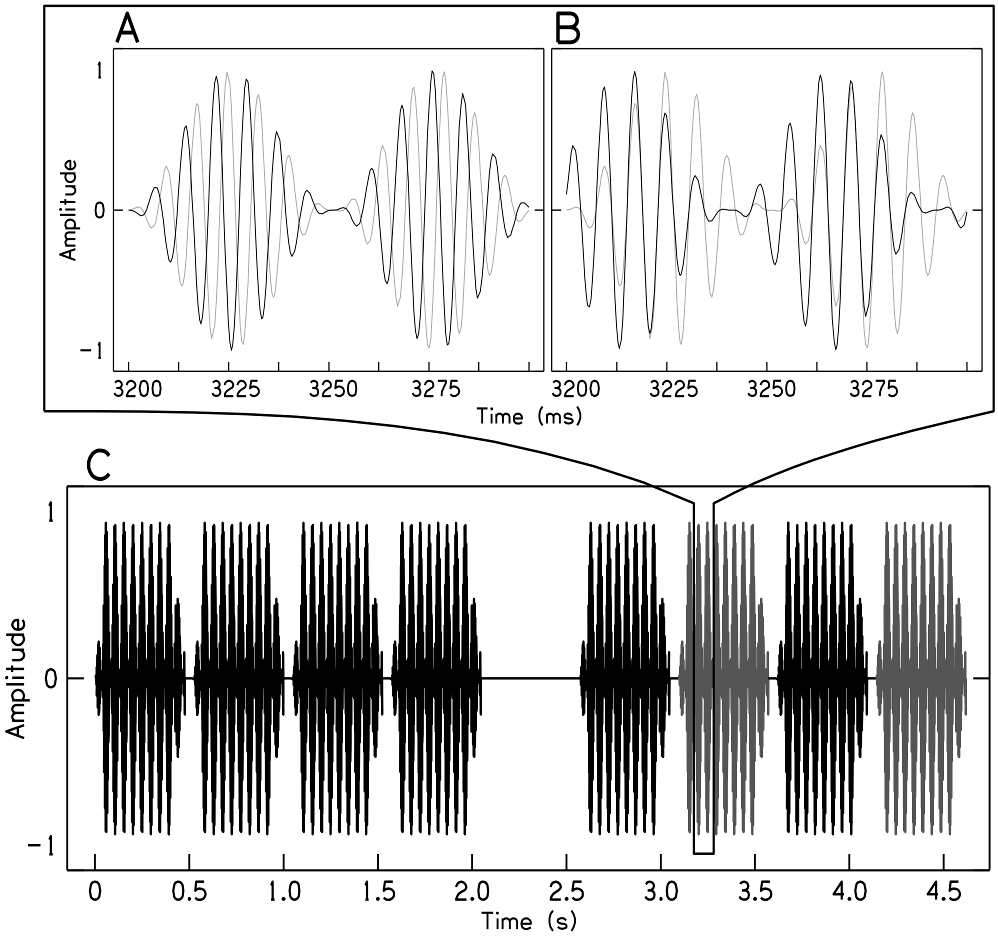

Figure 3.2: A schematic diagram of the stimuli and the presentation paradigm. Panels A and B are magnified from Panel C to give an indication of relative time scales. Panel A shows the AM tones for each ear; one (in grey) had a starting phase of 90° in the TFS whilst the other started at 0°. Presented binaurally, these tones resulted in an IPD in the TFS only—note the synchronous envelopes. Panel B shows the reverse: the grey tone had a starting phase of 90° in the envelope only—note the synchronous zero crossings of the TFS. Panel C shows the presentation paradigm with the target interval (ABAB) second. The grey tone bursts contained the IPD whilst black tone bursts were diotic.

The target IPD (δ°) was initially set to 180° and could not exceed this value. A geometric adaptive two-down, one-up procedure was used. The step size factor was 1.25^2 until three turn points occurred and 1.25 for eight subsequent turn points. The geometric mean of δ at the last eight turn points was taken as the IPD discrimination threshold. As δ was restricted to 180°, this algorithm would estimate a threshold even when performance was purely driven by chance. Therefore, if a listener failed to detect a δ of 180° at any point after the initial three turn points, the adaptive track stopped and 40 further trials with a fixed δ of 180° were presented. This happened 51 times out of 736 runs in total (15 out of 46 listeners). In these cases, a value of d' was calculated from the percent correct score (Hacker and Ratcliff, 1979). The relation between IPD threshold in degrees and d' has been shown to be linear (Hafter and Carrier, 1972), so an extrapolated threshold $\delta°$ was derived from the measured d' and the d' for 71% correct (0.78) by Eq. 3.1:

$$\delta(extrapolated) = \frac{(0.78 \times 180°)}{d'(40\ trials\ with\ \delta = 180°)} \tag{3.1}$$

All audio stimuli for absolute and IPD threshold measurement were created in MATLAB 7.6 (The MathWorks, 2008). Sounds were converted from digital to analog at a sample rate of 48 kHz and a 24-bit depth and amplified using a Creative E-MU 0202 USB soundcard. Sounds were played over Sennheiser HD 650 circum-aural headphones. Listeners sat in a double-walled listening booth, and made responses via a computer keyboard. Audiometric thresholds were measured using VIASYS GSI-Arrow and Kamplex AC30 audiometers coupled to TDH39P supra-aural headphones and Radioear B-71 bone vibrators.

3.4 Results

The geometric mean of the four repeated measurements of threshold discrimination was taken as the listener’s threshold in each of the four IPD conditions. IPD thresholds are plotted as a function of AT in Figure 3.3 and as a function of age in Figure 3.4. Some thresholds are plotted as upwards pointing arrows at 312°. These reflect performance below 62.5% correct in the constant stimuli method, which cannot be assumed (with 95% confidence) to be above chance. Although d' would not, in reality, continue to increase for δ>180°, extrapolated thresholds below 312° probably indicate some ability to detect IPDs. Extrapolated thresholds were limited to 312° for analysis, but cases where extrapolated thresholds exceeded this value should be interpreted as indicating an inability to discriminate the IPDs in those conditions.

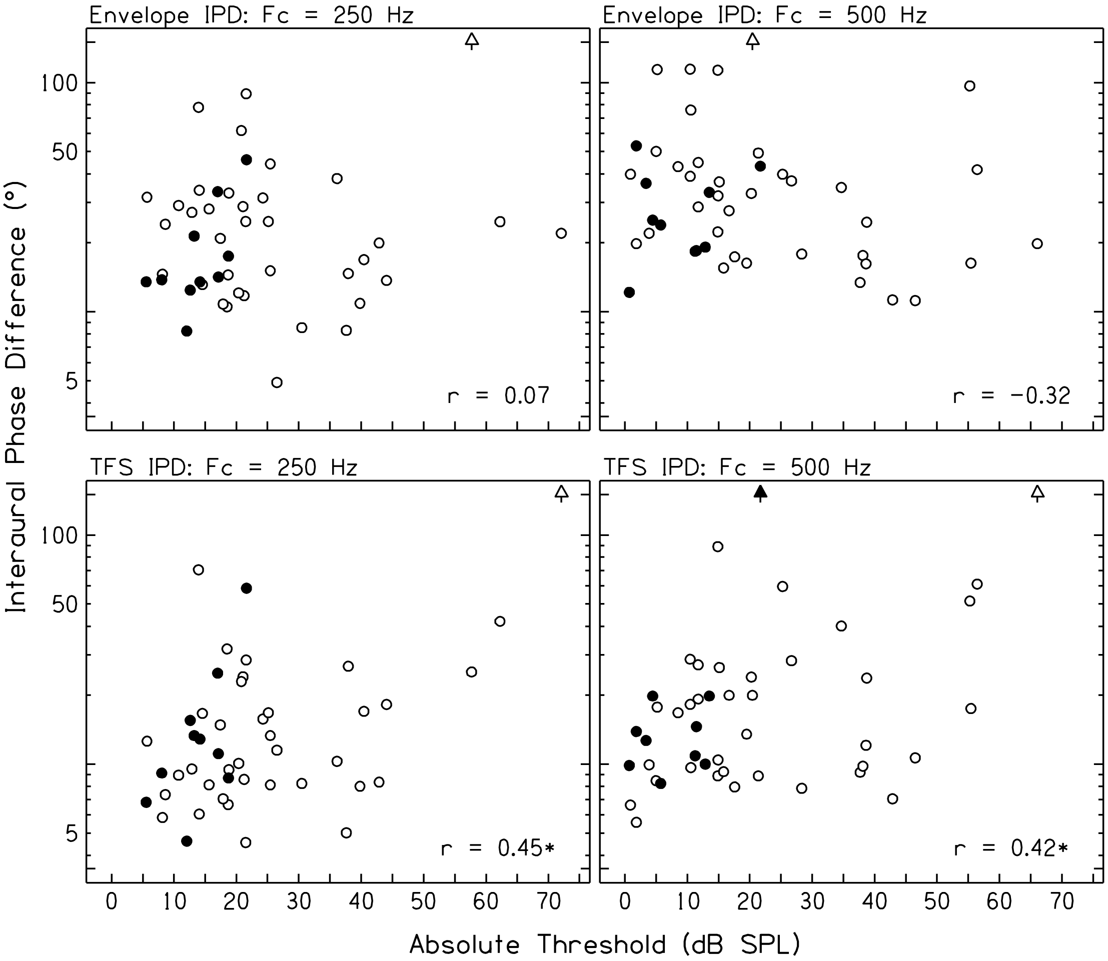

Figure 3.3: IPD thresholds as a function of AT averaged across ears (dB SPL). Clinically normal hearing listeners' thresholds are plotted with filled symbols and clinically hearing impaired listeners' thresholds (at any audiometric frequency tested) with open symbols. Upward pointing arrow symbols indicate a case where no IPD threshold could be measured. The top two panels show envelope IPD thresholds and the bottom two show TFS IPD thresholds. Left and right panels show thresholds for f_c=250-Hz and f_c =500-Hz, respectively. Correlation coefficients with age partialled out are given in each panel. Correlations significant (α=0.05) after Holm-Bonferroni correction are shown by asterisks.

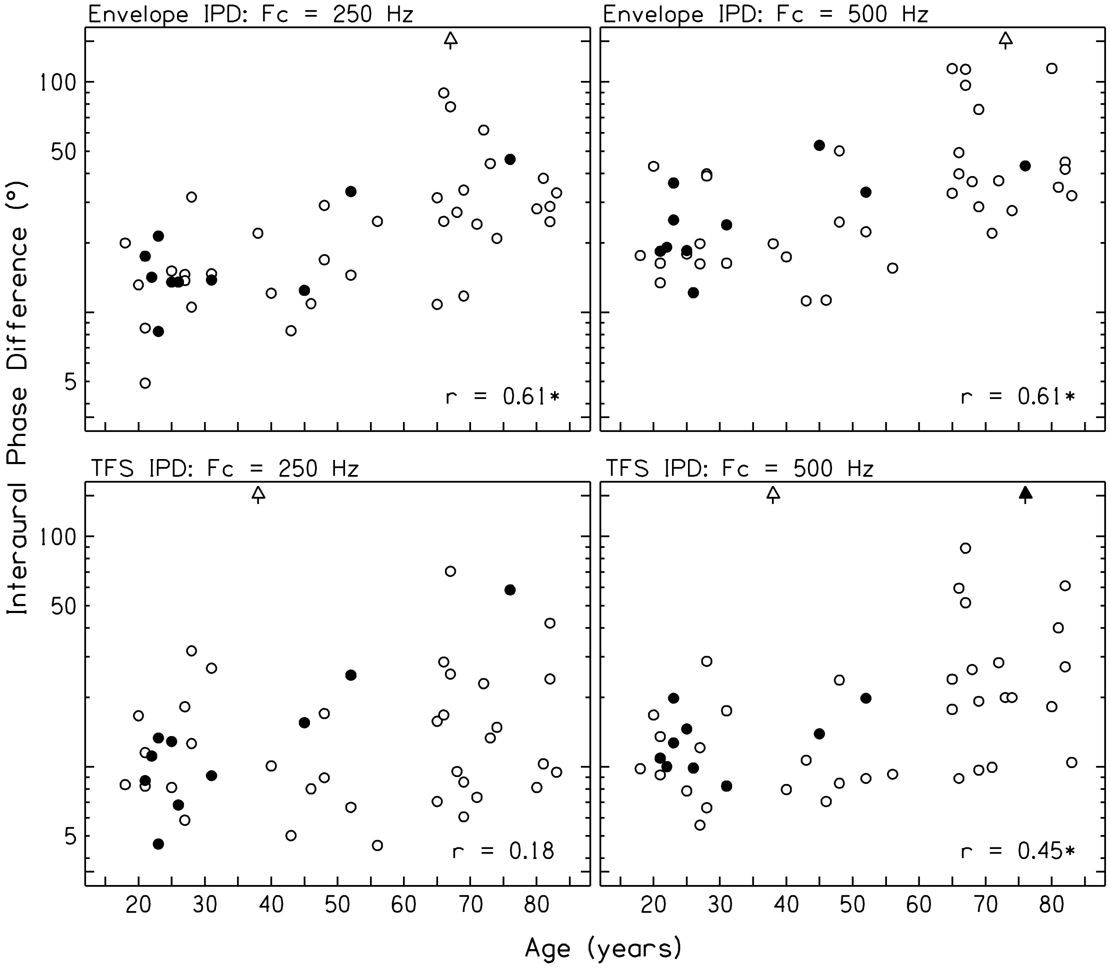

Figure 3.4: IPD thresholds as a function of age. Clinically normal-hearing listeners' thresholds are plotted with filled symbols and clinically hearing impaired listeners' thresholds with open symbols. Upward pointing arrow symbols indicate a case where no IPD threshold could be measured. Panels are in the same order as in Fig. 3. Correlation coefficients with AT partialled out are given in each panel. Correlations significant (α= 0.05) after Holm-Bonferroni correction are shown by asterisks.

IPD thresholds were log-transformed before statistical analysis as this resulted in thresholds that were more normally distributed. Pearson's product–moment correlations (r) were calculated between the IPD thresholds and the listeners' ages with ATs at f_c partialled out, and between the IPD thresholds and the ATs at f_c with age partialled out. Finally, correlations were calculated between IPD thresholds and PTA_{HF}. These correlations and partial correlations are given in Table 3.2. A sequentially rejective Bonferroni correction (Holm, 1979) was applied to the alpha criterion for each correlation to account for the increased familywise error rate due to testing the significance of 14 correlations (the twelve correlations described above and the correlations between age and PTA_{HF} and between age and AT). The Holm–

Bonferroni correction involves comparing the *p* values, from least to most likely given the null hypothesis, to $\alpha/(n+1-k)$ where *n* equals the number of familywise tests and *k* equals the index of the tests from least to most likely given the null hypothesis. Only one hypothesis test was affected by this correction. The correlation between ENV500 and AT was significant (p=0.035) before correction, but not after correction of α to 0.0063.

Table 3.2: The Pearson product–moment correlation coefficients (*r*) and the probability values (*p*) for the correlation or partial correlation between listeners' age, ATs and PTAHF and each of the four IPD conditions. Asterisks indicate significant correlations after Holm-Bonferroni correction.

IPD	Age (AT partialled out)		AT at f_c (age partialled out)		$\mathrm{PTA}_{\mathrm{HF}}$	
	r	*p*	*r*	*p*	*r*	*p*
TFS250	0.183	0.228	0.448	0.002*	0.102	0.500
TFS500	0.452	0.002*	0.415	0.005*	0.221	0.141
ENV250	0.613	<0.001*	0.063	0.679	0.202	0.179
ENV500	0.608	<0.001*	−0.315	0.035	0.039	0.796

With AT partialled out, age was significantly positively correlated with TFS500, ENV250 and ENV500, but not with TFS250. With age partialled out, AT was significantly positively correlated with TFS250 and TFS500. In contrast, the partial correlations between envelope-IPD thresholds and ATs (controlling for age) were weak and not significant after correction for multiple comparisons. No significant correlations were found between thresholds for the four IPD conditions and $\mathrm{PTA}_{\mathrm{HF}}$.

In order to determine whether TFS and envelope processing were affected differently by either age or AT, some of the partial correlations were compared to see whether they were significantly different from each other using Fisher's *r* to *z*-score transform. The difference between the *z* scores was divided by the standard error of the difference between the two *z* scores and evaluated against the *t* distribution with n_1+n_2-4 degrees of freedom (where n_1 and n_2 equal the sample sizes in the two correlations). The significance of the difference between the age and TFS250 correlation and the age and TFS500 correlation was also tested using this technique. These comparisons were calculated in the software package Statistica 10 (StatSoft Inc., 2011). Again, a sequentially rejective Bonferroni correction (Holm, 1979) was applied to account for the five comparisons made. The results of these comparisons are given in Table 3.3.

Table 3.3 The probability values (*p*) for the tests of the difference between pairs of partial correlations (after Fisher's *z* transformation). Each *p* value refers to the comparison between the correlation on the same line and the correlation on the line above. Asterisks indicate significant correlations after Holm-Bonferroni correction.

Partial Correlation	*r*	*p*
Age–TFS250 (controlled for AT)	0.183	
Age–ENV250 (controlled for AT)	0.613	0.016
Age–TFS500 (controlled for AT)	0.452	
Age–ENV500 (controlled for AT)	0.608	0.314
AT–TFS250 (controlled for Age)	0.448	
AT– ENV250 (controlled for Age)	0.063	0.055
AT–TFS500 (controlled for Age)	0.415	
AT– ENV500 (controlled for Age)	−0.315	0.001*
Age–TFS500 (controlled for AT)	0.452	
Age– TFS250 (controlled for AT)	0.183	0.165

The difference between the partial correlation between AT and TFS500 and the partial correlation between AT and ENV500 was significant (p<0.001), but the difference between the partial correlations between TFS250 and ENV250 and AT was not significant. The partial correlations between age and TFS500 and between age and ENV500 were not significantly different from each other. The partial correlations between age and TFS250 and between age and ENV250 were significantly different from each other before correction for multiple comparisons, but not after. The difference between the correlation between TFS500 and age and the correlation between TFS250 and age was also not significant.

Observed power for the correlations was 0.92, based on the mean coefficient of determination (r^2) of the partial correlations as the effect size, 46 listeners and an α of

0.05. The tests of differences between *z*-transformed correlation coefficients were calculated to have a mean power of 0.67, based on Cohen's *q* (difference between z-transformed correlation coefficients; Cohen, 1992), 46 listeners and an α of 0.05. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

3.5 Discussion

The results suggest that age and SNHL have negative, but independent, effects on TFS-IPD discrimination. Age was also associated with poorer envelope-IPD discrimination; in contrast, poorer envelope-IPD discrimination was not associated with increasing AT; instead, ENV500 performance may have improved very slightly with increasing AT, although this may be a type I error, as the correlation was not significant following correction for multiple comparisons.

3.5.1 SNHL and IPD sensitivity

The significant positive correlation between AT and TFS-IPD thresholds supports the idea that SNHL involves a reduction in the quality of, or ability to use, phase-locked information related to TFS (Lacher-Fougère and Demany, 2005; Hopkins and Moore, 2011). The present findings suggest deficits in TFS processing with elevated ATs are independent of age-related changes in TFS processing. There are numerous reasons why this relationship may occur (Moore, 2008; Hopkins and Moore, 2011):

(1) A reduction in the number of auditory nerve fibres can occur after damage to innervating the cochlea (Schuknecht and Woellner, 1955; Spoendlin, 1970), which may lead to reduced phase-locked information and consequently a degraded neural signal (Lopez-Poveda and Barrios, 2013).

(2) An abnormal phase response of the BM (Ruggero, 1994). An abnormal phase response may occur with loss of the nonlinear gain mechanism brought about by damage to the outer hair cells. This could affect comparisons of phase information across adjacent points along the BM that may be used to encode TFS information (Carney *et al.*, 2002).

(3) Changes to the central auditory system, such as a loss of inhibition, might disrupt the decoding of TFS (Moore, 2008).

It has also been suggested that the apparent TFS coding deficit for listeners with SNHL may arise because of the poorer frequency selectivity often associated with cochlear hearing loss. However the current study used stimuli containing only components that would fall within the equivalent rectangular bandwidth of a single normal auditory filter (Glasberg and Moore, 1990). Therefore, it seems unlikely that the relation between raised ATs and poor TFS-IPD thresholds can be explained by poor frequency selectivity.

Whilst Hopkins and Moore (2011), Lacher-Fougère and Demany (2005) and Moore *et al.* (2012a) found listeners with SNHL were poorer than NH listeners at TFS-IPD discrimination, they found no strong evidence of correlation between AT and TFS-IPD threshold at test frequencies of 500 Hz and lower. Hopkins and Moore (2011) and Moore *et al.* (2012a) used pure tones, which, unlike the AM tones used in the current study, would not provide conflicting interaural envelope and TFS cues of a zero and non-zero IPD respectively. The conflicting cues may impede TFS-IPD discrimination by those with elevated ATs as envelope cues may become dominant over TFS cues with noise-induced hearing loss (Kale and Heinz, 2010). However, Lacher-Fougère and Demany (2005) also did not find a correlation between AT and TFS500 thresholds with stimuli that were similar to those used in the current study (except in level). Rather than due to stimuli, the inconsistency in observed relationship between AT and TFS-IPD thresholds may be due to differences in the nature of the SNHL of the listeners, the extent to which age and AT were correlated or the sample sizes in the different studies.

Reduced sensitivity to TFS IPDs may have important consequences for people with SNHL when they are listening in noisy backgrounds. Firstly, Moore and Glasberg (1987a) showed that TFS information is useful for separating target sounds from fluctuating background noises. This appears to extend to more complex hearing abilities such as the intelligibility of speech in background noise (e.g., Füllgrabe *et al.*, 2006; Lorenzi *et al.*, 2006; Hopkins and Moore, 2009; Moore, 2012). Secondly, the TFS insensitivity observed with SNHL in the current study was demonstrated with IPDs, which are thought to be important for separating sounds from different azimuths (Bronkhorst and Plomp, 1988). Consistent with this, the benefit to speech intelligibility of separating target and masker sentences in azimuth declines with SNHL (Neher *et al.*, 2009) and with TFS-IPD thresholds (Neher *et al.*, 2011). However, Neher *et al.* (2012) found that this correlation was no longer significant when age was accounted for, suggesting that the benefit of azimuthal speech separation and TFS-IPD sensitivity may be affected by a common, age-related cause (Neher *et al.*, 2012).

Lacher-Fougère and Demany (2005) found a deficit in envelope-IPD thresholds for SNHL listeners and a positive correlation between AT and envelope-IPD thresholds. The current study found no deficit in envelope-IPD thresholds associated with ATs. Whilst no significant correlation was found between AT and ENV250, the negative correlation between AT and ENV500 suggests a trend towards *better* ENV500 performance with increasing hearing loss. SNHL listeners may perform better than NH listeners at equal SL because of loudness recruitment, which would effectively magnify perceived envelope fluctuations making them more salient (Moore *et al.*, 1996). However, the negative correlation was not significant when corrected for multiple comparisons. The deficit found by Lacher-Fougère and Demany (2005) may be due to the confounding of SNHL and age within the listener groups. The listeners with SNHL in the study of Lacher-Fougère and Demany (2005) might have had poorer envelope IPD thresholds than the NH group because they had a higher mean age. The current study found positive correlations between envelope-IPD thresholds and age, but because listeners with a range of ages and ATs were tested, it was possible to assess the effect of AT at the f_c on envelope-IPD thresholds independently of the effect of age, avoiding this confound.

Lacher-Fougère and Demany (2005) did not view the deficit in envelope IPD thresholds with SNHL as a deficit in envelope processing *per se*, but as a result of NH listeners experiencing a higher SL than the SNHL listeners because stimuli were played at a fixed level (75 dB SPL) for both groups. Lacher-Fougère and Demany (2005) included a control experiment which supported this idea: with 35 dB SPL stimuli, or reduced SL due to the presence of white noise low-pass filtered at 1250 Hz, NH listeners performed worse at envelope-IPD discrimination, but not at TFS-IPD discrimination.

The different relationships of TFS500 and ENV500 sensitivity with AT may be explained by a shift in the balance of TFS and envelope coding in ears with SNHL. Kale and Heinz (2010) provided physiological evidence that, rather than showing an absolute reduction in the precision of phase locking, individual nerve fibres phase lock more to stimulus envelope than to stimulus TFS after mild-to-moderate noise-induced hearing loss. The behavioural results in the present study may reflect a change in nerve fibres' phase locking from predominately following the TFS to predominantly following the envelope. Kale and Heinz (2010) suggested that improved envelope coding may not necessarily benefit SNHL listeners in fluctuating background noise as it may magnify the fluctuations perceptually. Increased fluctuation makes gaps in narrow-band noise more difficult to detect (Glasberg and Moore, 1992) and reduces speech intelligibility in fluctuating noise background noise (Moore and Glasberg, 1993).

3.5.2 **Age and IPD sensitivity**

Moderate positive correlations with AT partialled out were found between age and ENV250, ENV500 and TFS500, but TFS250 was not significantly correlated with age. The correlations of ENV250, ENV500 and TFS500 with age suggest that aging leads to a more general loss of temporal acuity than SNHL, whereas SNHL appears to result in loss of TFS sensitivity specifically. This general loss of temporal acuity with age may stem from changes in processing speed or accuracy in more central parts of the auditory system; parts where the coding of both the envelope and TFS coded signals are vulnerable. This is consistent with the suggestion by He *et al.* (2008) of a general age-related decline in synchronisation of neural responses to both TFS and envelope. He *et al.* (2008) based this suggestion on age-related changes in AM detection as a function of modulation frequency and carrier frequency. Using electrophysiological measures, Ruggles *et al.* (2012) found that the strength of phase-locking to the envelope of the /dah/ syllable was poorer for middle-aged listeners compared to young adults, providing further evidence that age is associated with a decline in the fidelity of temporal coding. Previous research shows that an age-related decrease in the highest modulation rate to which a listener is sensitive (Purcell *et al.*, 2004), and age-related changes in AM detection, and auditory steady-state responses of the brainstem phase-locking to the envelope, are typically less pronounced at modulation rates below 40 Hz (Leigh-Paffenroth and Fowler, 2006; Grose *et al.*, 2009b). However, the current study found age-related changes in envelope-IPD thresholds at 20 Hz, suggesting envelope coding can be affected by age even at low modulation rates.

Aging has been associated with reduced temporal resolution as measured by gap detection (Schneider *et al.*, 1994; Strouse *et al.*, 1998) and modulation detection (He *et al.*, 2008), as well as interaural phase discrimination (Grose and Mamo, 2010) and lateralisation (Strouse *et al.*, 1998). Aging causes a complex collection of changes in the physiology of mammals, changes that are likely to result in a wide range of deficits in hearing. There are several likely age-related causes of degraded acuity of TFS and envelope coding:

(1) Degeneration of cochlear synapses and peripheral axons of spiral ganglion cells (Makary *et al.*, 2011; Sergeyenko *et al.*, 2013), which would lead to less phase-locked information over which to aggregate a temporal code.

(2) Imbalances in excitatory and inhibitory neural mechanisms may change envelope coding in the inferior colliculus (Walton *et al.*, 2002).

(3) Reduced synchrony in the transmission of phase-locked signals, which could weaken the strength of the phase locking (Clinard *et al.*, 2010; Marmel *et al.*, 2013).

However, other age-related changes, such as the functioning of neurotransmitter GABA in the inferior colliculus (Caspary *et al.*, 1995), could affect the auditory system in a variety of ways and interact with the other physiological phenomena listed above.

3.5.3 **Age-related high-frequency SNHL and TFS processing**

Age-related hearing loss is characterized by high-frequency rather than low-frequency hearing loss (Morrell *et al.*, 1996; Dubno *et al.*, 2013). Whilst $\mathrm{PTA}_{\mathrm{HF}}$ was significantly positively correlated with age, it was not correlated significantly with any of the IPD thresholds. This is consistent with the findings of Moore *et al.* (2012a), who studied pure-tone IPD discrimination by elderly listeners with minimal hearing loss at low frequencies, but a range of hearing loss severities at higher frequencies. They found that high-frequency loss was only weakly correlated with pure-tone IPD discrimination and this correlation was not significant once the effect of age (which was strongly correlated with pure-tone IPD discrimination) was partialled out. This result contrasts with the results of Smoski and Trahiotis (1986) showing above-normal IPD discrimination thresholds for 500 Hz pure tones for listeners with moderate to severe high frequency hearing loss, but thresholds below 20 dB HL at 500 Hz. However, Smoski and Trahiotis (1986) only tested four listeners with this profile of hearing loss.

3.5.4 **Conclusions**

The results suggest both SNHL and age have independent relationships with IPD discrimination:

(1) The sensitivity to IPDs in the TFS of AM tones deteriorated with increasing low frequency SNHL.

(2) The correlations between envelope-IPD thresholds and SNHL were weak and non-significant.

(3) Both TFS- and envelope-IPD thresholds increased with age. Temporal processing may deteriorate in the auditory system such that both TFS and envelope processing are affected.

3.6 Acknowledgments

We thank the associate editor, Laurent Demany, Brian C. J. Moore and two anonymous reviewers for the helpful comments on a previous version of this paper. This work was supported by Oticon A/S and the UK Medical Research Council (grant ref: G1001609).

Chapter 4. Differential Group Delay of the Frequency Following Response Measured Vertically and Horizontally

Andrew King, Kathryn Hopkins, and Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom

This chapter was prepared for publication as an original research article in the Journal of the Association for Research in Otolaryngology.

4.1 Abstract

The frequency following response (FFR) arises from the sustained neural activity of a population of neurons that are phase locked to periodic acoustic stimuli. Determining the generating source of the FFR non-invasively may be useful for understanding the nature and function of phase locking in the auditory pathway. The current study compared the FFR recorded with a horizontally aligned (mastoid-to-mastoid) electrode montage and a vertically aligned (forehead-to-neck) electrode montage. Stimuli were five amplitude-modulated tones centred at 576 Hz, each with a different modulation rate, resulting in different side-band frequencies across stimulus conditions. Changes in response phase across stimulating frequency (group delay) were used to determine the latency of the FFR measured by the two montages. For FFR reflecting phase locking to the side-band frequencies, the horizontal montage had a shorter group delay than the vertical montage, suggesting that it measured an earlier generation source within the auditory pathway than the vertical montage. For FFR reflecting phase locking to the modulation rates, no significant difference in group delay was found between montages. However, it is possible that multiple sources of FFR (and the cochlear microphonic) were recorded by each montage, causing interactions across frequency that complicate interpretations of the group delay.

4.2 Introduction

The frequency following response (FFR) reflects the electrical neural response to periodic acoustic stimuli. Its name derives from the synchronized nature of the response to the periodicities in the stimulus (Worden and Marsh, 1968; Moushegian *et al.*, 1973; Glaser *et al.*, 1976; Stillman *et al.*, 1976). FFR relies upon consistent phase locking of neural firing to the filtered output of the basilar membrane (Tasaki, 1954; Rose *et al.*, 1967; Smith *et al.*, 1975; Palmer and Russell, 1986). Phase locking is important for a precise temporal code integral to pitch perception (e.g., Cariani and Delgutte, 1996; Krishnan and Plack, 2011), localisation (Rose *et al.*, 1966), and speech perception (Møller, 1999; Krishnan *et al.*, 2004; 2005) particularly in noisy environments and adverse listening conditions (Sachs *et al.*, 1983).

In animal models, phase locking has been measured directly in neurons throughout the subcortical afferent auditory pathway, from the auditory nerve (Tasaki, 1954) and cochlear nucleus (Galambos and Davis, 1943) up to at least the inferior colliculus (e.g., Rose *et al.*, 1966; Smith *et al.*, 1975). In humans, phase locking can be measured indirectly, through the FFR. The combined extracellular far-field electric field potential created by the responses of many neurons can be measured at the scalp using electroencephalography (Jewett and Williston, 1971). As the FFR is measured remotely, it is difficult to determine which structure in the auditory pathway is generating the response. Early reports assumed the FFR emanates from a single source (Gerken *et al.*, 1975; Smith *et al.*, 1975). However, later studies showed that multiple sources are measurable under certain recording conditions (Stillman *et al.*, 1978; Gardi *et al.*, 1979; Galbraith, 1994; Galbraith *et al.*, 2000; Galbraith *et al.*, 2001).

Identifying FFR generation sites could help explain changes in phase locking at different stages of the auditory pathway. This could help explain individual differences in hearing ability or help with diagnosis and treatment of sensorineural hearing impairments. One way of determining generation site is by pairing the latency of FFR with the latency of compound action potentials measured intra-cranially on various structures in the auditory pathway. This has been done during surgical operations for determining which structures contribute to the peaks in the first 10 ms of the auditory brainstem responses (ABRs) to transient (e.g. click) stimuli. Waves I and II are associated with the auditory nerve and later waves are associated with various nuclei in the brainstem (Møller, 2007). However, the parallel and crossing connections between nuclei make simple, one-to-one relations between waves and anatomical structures difficult beyond the cochlear nucleus. Unlike the ABR, the FFR is sustained, so the number of waves is not related to the generation sites, but to the number of periods in the stimulus. The FFR onset latency may reveal the first generation site, but measuring the latency of the onset amplitude of the FFR to ongoing tone bursts is difficult to define precisely because the onset is gradual (Krishnan, 2007).

One method of differentiating FFR sources is to use differently oriented electrode montages (Stillman *et al.*, 1978; Galbraith, 1994; Galbraith *et al.*, 2000; Galbraith *et al.*, 2001). Vertical and horizontal montage orientations have been shown to differentially emphasize later and earlier waves of the click ABR respectively (Picton *et al.*, 1974; Scherg and Von Cramon, 1985). It has been hypothesized that the same differentiation can be made for FFR. Stillman *et al*. (1978) attempted to determine FFR latency for pure tone stimuli (167 to 500 Hz) by visually comparing FFR waveforms recorded from electrode montages in vertical and horizontal orientations (vertex to earlobe and earlobe to earlobe, respectively). Below 500 Hz, both orientations recorded complex waveforms with two peaks per stimulus period. The second peak (smaller in the horizontal montage, but larger in the vertical montage) trailed the first peak by 1.7 ms, regardless of frequency or level. The time between the second peak and the first peak of the next period was dependent on frequency (*e.g.* 3.3 ms for 200 Hz, 1.8 ms for 286 Hz). The earlier peaks were greater in amplitude than the later peaks in the horizontal orientation and *vice versa* in the vertical orientation. Scherg and Brinkmann (1979) found similar latencies to Stillman *et al*. (1978) with similar electrode montages. Furthermore, they found the earlier FFR was only recorded at the ipsilateral earlobe, not the contralateral earlobe, demonstrating that the generator of the early FFR is before the level of binaural integration (Krishnan, 2007). Whilst Stillman *et al*. (1978) showed that the shorter latency FFR was not the cochlear microphonic (CM)—the non-neural electrical activity that matches the input generated by the cochlear hair cells (Terkildsen *et al.*, 1974)—other studies (Sohmer and Pratt, 1977; Davis and Britt, 1984) have suggested that very short latency (around 1 ms) FFR-like responses are most likely to be the CM. One limitation of these studies is that subjective judgment was used to determine onset latency; this method is not easily defined or repeated.

Galbraith and colleagues have tried other, more objective methods to determine generation sites of the FFR with horizontal and vertical electrode montages. Galbraith (1994) used the latency at which the cross-correlation between stimulus and response was maximal to determine response latency. He found that the horizontally recorded FFR had a cross-correlation latency equivalent to the latency from the auditory nerve, but only with pure-tone stimuli; FFRs to the missing fundamentals of complex stimuli were not present in the horizontal record. The vertical record had a latency equivalent to that for the lateral lemniscus, and both missing fundamental and pure tones were represented in the FFR. Galbraith *et al.* (2000) examined FFR amplitude across frequency for pure tone stimuli. The horizontal montage between external auditory canals picked up FFR tuned to higher frequencies than the vertical channel between inion and vertex. Galbraith *et al*. (2000) inferred from the differences in spectra that the horizontal montage favoured a peripheral generation site and the vertical favoured a central generation site, because the auditory nerve can phase lock to higher frequencies than the rostral brainstem and midbrain (e.g., Palmer and Russell, 1986; Winter and Palmer, 1990; Liu *et al.*, 2006). However, spectral content is not a direct measure of generation site, and cross-correlation may be strongly determined by onset characteristics of the response. Cross-correlation is also limited to latencies less than the stimulus periodicity (restricting it to low-frequency stimuli).

Arguably a more reliable measure of the latency of FFR generation is through the use of group delay. Group delay is the change in the unwrapped phase angle of the response to continuous tones as a function of frequency. For a given delay, higher frequency sinusoids will have a larger unwrapped phase angle than lower frequency sinusoids. Provided the elapsed time is less than the reciprocal of the difference in frequency between consecutive sinusoids, their phases will be less than one cycle apart, and so the phase response can be unambiguously unwrapped to determine the elapsed time.

Batra *et al*. (1986) measured the FFR to pure tones across a range of audio frequencies. They plotted the phase of the Fourier transform of each pure tone FFR against its frequency to determine the group delay. At low frequencies, there appeared to be a steeper change in phase with frequency (larger group delay) than at higher frequencies; for tones below about 300 Hz, group delay was about 8 ms. Batra *et al*. (1986) contended that the response at about 1000 Hz—with a group delay of less than 1 ms—was the CM. Batra *et al*. (1986) suggested that the change in group delay from low to high frequency was descriptive of different generation sites. Batra *et al*. (1986) also found oscillations in spectral magnitude across frequency, which may be evidence of multiple FFR sources interacting destructively or constructively in the recorded response, and is dependent on stimulation frequency and latency between the sources.

The current study aimed to determine if an FFR generated at a peripheral site (auditory nerve, either at the cochlear or the cochlear nucleus) could be measured distinctly from an FFR generated in the brainstem at a more central site (*e.g.* lateral lemniscus or inferior colliculus). Specifically, can a horizontally-aligned electrode montage record earlier generated FFR (latencies around 2-3 ms) than a vertically aligned electrode montage (latencies around 5-6 ms)? Amplitude-modulated tones were used so that group delay over the modulation rates could determine the latency of the responses to the envelope, and group delay over the tone components could determine the latency of the responses to the fine structure.

4.3 Materials and methods

4.3.1 Listeners

Twelve male and eleven female adult listeners (18 to 31 years, mean = 23years) with audiometric thresholds below 25 dB hearing level at 0.25, 0.5, 1 2 4 and 8 kHz were recruited. All procedures of the study were approved by the Research Ethics Committee at the University of Manchester.

4.3.2 Stimuli

Stimuli consisted of three, equal-amplitude pure tones summed together. The highest and lowest tone frequencies were equidistant from the centre tone frequency, thus creating an amplitude-modulated tone with a modulation rate (f_m) equal to the frequency spacing. Five frequency spacings were used, all with the same centre tone frequency (see Table 4.1). All stimuli were presented at a root mean square average of 85 dB SPL. This sound level was used because at low sound levels the FFR is difficult to record and FFR amplitude increases monotonically with stimulus level until about 65 dB above threshold (Marsh *et al.*, 1975; Krishnan, 2002). In the current study, 85 dB SPL was expected to be about 65 to 75 dB above threshold for most listeners. For each trial, the modulated tones were presented for 140 ms in the positive starting polarity, then after a 120 ms silent interval, presented again in the negative starting polarity, followed by another silent interval of 170 ms before the next presentation pair. Each 140 ms tone included 20 ms raised-cosine onset and offset ramps. When the response to the negative polarity stimulus is subtracted from the response to the positive polarity stimulus the response following the fine structure adds constructively and the response following the envelope mostly cancels out (Goblick and Pfeiffer, 1969). This was used to quantify the TFS FFR. On the other hand, when the responses to the two stimuli of opposing polarities are summed the response following the fine structure mostly cancels out, whereas the response following the envelope adds constructively (see Aiken and Picton, 2008). This was used to quantify the envelope FFR. These two manipulations are shown in Figure 4.1 for one listener in the condition with f_m=115 Hz.

Table 4.1: The frequency components (in Hz) of the five stimuli over which group delay was calculated. Each row corresponds to one stimulus.

Modulation	*Lower side-tone*	*Centre Frequency*	*Upper side-tone*
85	491	576	661
100	476	576	676
115	461	576	691
130	446	576	706
145	431	576	721

The five stimuli were tested separately in blocks of 1600 trials. Responses were stored for analysis as 16 sub-averages (each an average of 100 trials). Only sub-averages were stored for analysis. The stimuli were created in MATLAB (The Mathworks, 2011) and presented using Tucker Davis Technologies (TDT) SigGen and BioSigRP software. Stimuli were converted to analogue signals by a TDT RP2.1 processor and transduced to acoustic waves outside the listening booth by ER30 earphones (Etymotic Research, Illinois) to minimize stimulus artefact. The transducers were connected to the listener by 6 m of tubing terminating in the listener’s right external auditory canal through a foam earplug. The listener’s left ear was plugged with a foam plug. Pilot tests showed that no stimulus artefact was recorded when the stimulus was presented with the tubing detached from the earplug and sealed with tape.

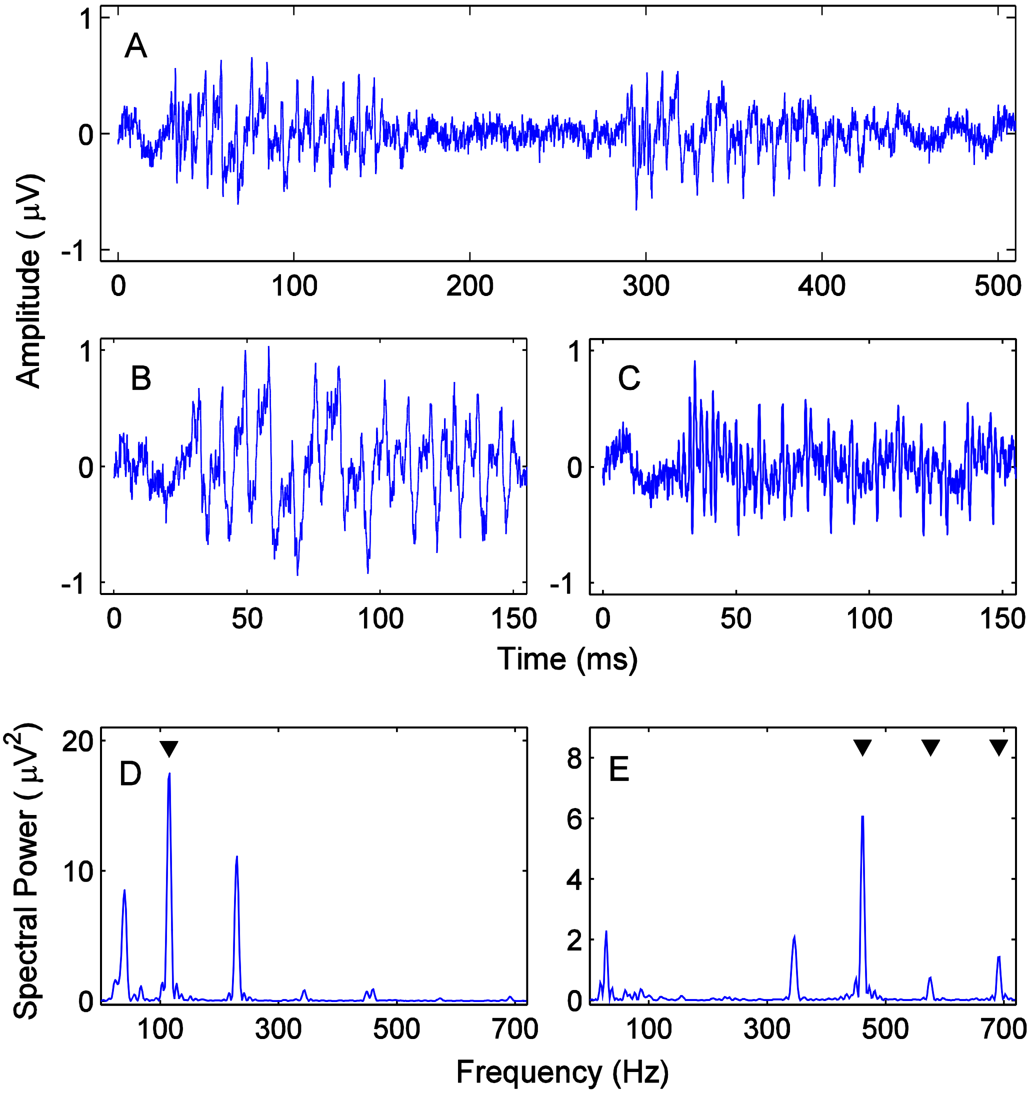

Figure 4.1: An example of a listener's FFR to the alternating polarity stimuli for one condition (f_m=115 Hz). A) The mean FFR waveform for both polarities in sequence. B) Addition of the responses to the two polarities. C) Subtraction of the second response from the first. D) The fast Fourier transform power spectrum of the addition waveform with a downward arrow denoting f_m. E) The power spectrum of the subtraction waveform with downward arrows denoting the three component frequencies.

The electrical field potential was recorded by two montages of gold-plated passive electrodes. The horizontal montage recorded at the ipsilateral mastoid (referenced to the contralateral mastoid) and the vertical montage recorded at the seventh cervical vertebra (referenced to the forehead hairline on the sagittal line). Both montages shared a common ground electrode on the listener’s brow. The electrodes were wired into a TDT RA16LI-D head-stage linked to a TDT RA4PA pre-amplifier and analogue-to-digital convertor. The digital signal was sent via fibre optics to a TDT RA16 Medusa Base Station for processing. The base station and RP2.1 processor were linked for clock synchronisation and both communicated with the BioSigRP software via optical fibres. No filtering was applied to the recordings. Individual trials with a peak amplitude exceeding ±60 μV were rejected from the sub-averages. Listeners lay in a reclining chair and were asked to relax as much as possible and try to sleep during the recordings. Listener wakefulness was not recorded.

4.3.3 **Analysis**

Recordings were exported to text files, read and analysed by MATLAB scripts. Records were divided into the horizontal and vertical montages. For the envelope FFR, the magnitude of the discrete Fourier transform (DFT) at the modulation rate was calculated from the mean added responses for each stimulus condition. For the TFS FFR, magnitude of the DFT at the lower side-tone and upper side-tone frequencies was calculated from the mean from the subtraction waveform.

A statistical criterion based on the signal-to-noise ratio (SNR) was used to determine the presence or absence of a response to the stimulus. FFR was accepted as present if the magnitude for the DFT at the signal frequency was greater than the mean magnitude at frequencies surrounding it (noise frequencies) by a factor of 2.57 standard deviations (SDs) of the magnitude at the noise frequencies. A factor of 2.57 SDs means a 0.01 probability of a fluctuation in the noise being accepted as a response to the signal. Noise frequencies were selected at a resolution of 4 Hz, from 9 to 37 Hz above and below the signal frequency. The listeners’ mean strength of FFR magnitude at each frequency, above the noise floor, is shown in Figure 4.2. The deviation in FFR magnitude shows that it was not always sufficiently above the noise floor to pass the SNR criterion. The phase of the DFT was used only when the FFR magnitude passed this criterion. The phase of the FFR was unwrapped for each frequency region (modulation rates, lower side-tones, upper side-tones). Group delay for a frequency region was taken only if FFR at three or more frequencies in that frequency region passed the SNR criterion.

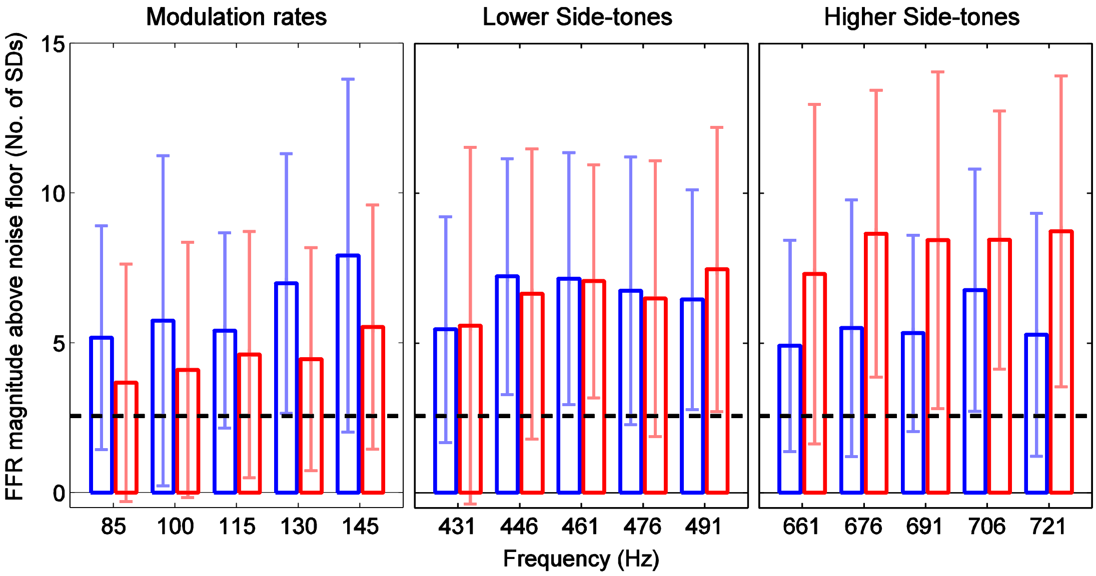

Figure 4.2: FFR magnitude at each modulation rate (left panel), lower side-tone (middle panel) and higher side-tone (right panel) above the mean magnitude of the noise floor as a ratio of the SD of the magnitude of the noise floor. The columns represent the mean across listeners and the error bars indicate the SD across listeners. Blue columns indicate the vertical montage and red columns and bars indicate the horizontal montage. 2.57 SDs from the noise floor (the SNR criterion) is given by the black dashed lines.

4.3.4 **Phase unwrapping**

Because a group delay fit could be made with phase values at a minimum of three out of five frequency points in any given frequency region meeting the SNR criterion, gaps of 30 or 45 Hz between consecutive data points occasionally existed. Without any gaps (a frequency spacing of 15 Hz), sequential unwrapping was unambiguous for group delays under 33.3 ms, but if phase values were missing at one or two consecutive frequency points (frequency spacing of 30 and 45 Hz respectively), the maximum group delay for which unambiguous unwrapping is possible dropped to 16.7 and 11.1 ms, respectively. As such group delays were within the test range, sequential unwrapping was not used. Instead, all possible unwrapping possibilities that could produce group delays between 0 ms and 20 ms were calculated. The unwrapping that had the best linear fit was selected and the slope of that fit was taken as the group delay only if the fit was good (the sum and squared residual error was less than 0.05). Figure 4.3 shows an example of this method with model data. This method was employed because sequentially unwrapping the phase within tolerances of

±π sometimes led to poor linear fits or linear-fit slopes suggesting physiologically unreasonable latencies.

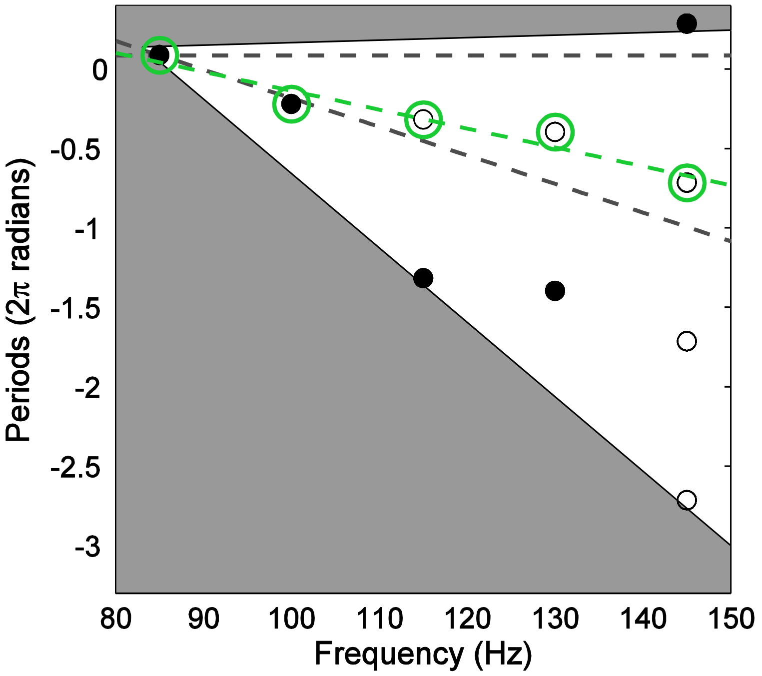

Figure 4.3: A model example of phase values of the DFT at each of the modulation frequencies (filled black circles). The grey area indicates the limits of the unwrapping (only phase values within the non-grey area were considered by the unwrapping algorithm). The open circles indicate unwrapping possibilities considered by the unwrapping algorithm in this example. The green circles indicate the unwrapping option that provides the best linear fit (green dashed line) within the boundaries of the fits allowed by the algorithm (black dashed lines).

4.4 Results

For each listener, up to six group delays were calculated. In each of three frequency regions, two group delays were calculated; one for the horizontal montage and one for the vertical montage. The three frequency regions were the Modulation frequency region (far-left column of Table 4.1 and left panel of Figure 4.2), the Lower Side-Tone frequency region (middle-left column of Table 4.1 and middle panel of Figure 4.2), and the Higher Side-Tone frequency region (far-right column of Table 4.1 and right panel of Figure 4.2). Due to the imposition of an FFR SNR criterion and group delay slope and fit criteria, not all listeners' data provided the full six group delay values.

Table 4.2 shows the number of listeners for whom FFR passed the SNR criterion at each frequency point and the number of listeners for whom group delay passed the slope and fit criteria for each frequency region and montage. Twenty-five percent of the group delay data were missing, with little difference in the amount of missing data between frequency regions, but more missing group delays from the horizontal montage (33%) than from the vertical montage (16%). Analysis of missing data was not pursued further. Of the group delays that were passed the criteria, the mean group delay for each montage in each frequency region is plotted in Figure 4.4. Figure 4.4 shows that the horizontal montage generally measured shorter group delays in the FFR than the vertical montage, most notably for the Lower Side-Tone frequency region (and to a lesser extent the Higher Side-Tone frequency region). However, the difference in group delay between the montages was only slight in the Modulation frequency region.

To avoid excluding the 17 listeners with missing group delay data from further analysis, a linear mixed-effects model was used (‘xtmixed’ in STATA 13; StataCorp LP, Texas). This model allows unbalanced amounts of data across factor levels, but assumes that the data are missing independent of the observed or missing data values. The effects of montage and frequency region, and the interaction between them, on group delay were analysed as fixed-effects factors. A random-effect term was included to account for possible within-subject clustering effects (*i.e.* intrinsic listener effects). An adjusted degrees-of-freedom likelihood-ratio test indicated no statistical difference between the mixed-effects model with and without the random-effects term [$\chi^2(1) = 1.7e^{-13}$, $p=1.00$]. There was a statistically significant interaction between the effects of montage and frequency region on group delay [$\chi^2(2)=6.17$, $p<0.05$]. In the presence of interaction terms, there was also a highly statistically significant effect of montage on group delay [$\chi^2(1)=18.81$, $p<0.001$]; on average, it is estimated that group delay was 2.13 ms shorter with the horizontal montage than the vertical montage.

Table 4.2: The number of FFR, recorded by each montage, that passed the criterion for on-frequency magnitude being sufficiently above noise floor; and the number of calculable group delays in each frequency region that passed the criterion for acceptable group delay.

Frequency region	Modulation					Lower side-tone					Upper side-tone				
Frequency (Hz)	85	100	115	130	145	431	446	461	476	491	661	676	691	706	721
Vertical Montage	18	17	19	21	21	17	21	19	21	20	16	19	18	20	17
Group Delay	20					20					18				
Horizontal Montage	15	15	18	15	18	13	16	19	15	19	17	21	20	22	21
Group Delay	14					15					17				

The difference in group delay between the montages at each frequency region showed that modulation FFR was not significantly different across montages (0.84 ms, $p>0.05$), but the horizontal montage measured FFR with a significantly shorter group delay than the vertical montage for the lower (3.77 ms, $p<0.001$) and higher (1.75 ms, $p<0.05$) side-tones. This can be seen in Figure 4.4.

Comparing the modulation-FFR group delay with the side-tone TFS-FFR group delay indicated modulation-FFR group delay was significantly longer than the higher side-tone FFR group delay in both montages (3.16 and 4.08 ms longer in vertical and horizontal montage respectively, $p<0.001$). However, modulation-FFR group delay was only significantly longer than lower-side-tone FFR group delay in the horizontal montage (4.25 ms, $p<0.001$), not the vertical montage (1.31 ms, $p>0.05$), whilst the lower and higher side-tone FFR group delay only differed significantly in the vertical montage (1.85 ms, $p<0.05$), not the horizontal montage (0.17 ms, $p>0.05$).

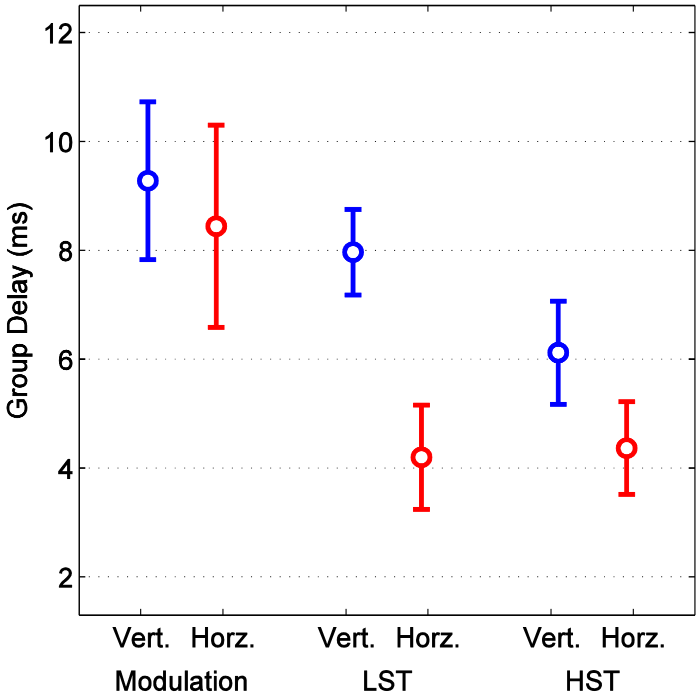

Figure 4.4: The mean group delays (across listeners) for FFR to the modulation rates (Mod), lower side-tones (LST), and higher side-tones (HST) for the vertical (blue) and horizontal (red) montages. Error bars indicate confidence intervals of 95%.

4.5 Discussion

For most listeners, group delays for the lower and higher side-tones of the TFS FFR recorded with the horizontal montage were around 4 ms for the horizontal montage. Group delays of about 3 to 4 ms are consistent with latencies of waves II or III in click-ABR literature (Møller and Jannetta, 1981; 1983), which are suggested to originate from the rostral end of the auditory nerve or the cochlear nucleus. On the other hand, the vertical montage recorded significantly longer higher side-tone group delays of around 6 ms, and lower side-tone group delays of about 8 ms. These delays are broadly consistent with the latency of waves V or later, originating from the inferior colliculus (Møller and Jannetta, 1982; Møller *et al.*, 1994). Both montages recorded modulation (envelope) FFR with group delays around 8 to 9 ms suggesting sources in the more rostral brainstem, or midbrain. The difference between montages in side-tone TFS-FFR group delay supports previous claims that horizontal and vertical montages can record FFR from distinct earlier and later sources respectively (Stillman *et al.*, 1978; Galbraith, 1994).

It is possible that the envelope-FFR group delays did not differ significantly between montages because there is only a single source (or one that dominates others). The mean FFR magnitude above the noise floor plotted in Figure 4.2 for the modulation rates are greater in the vertical montage than the horizontal montage, which may suggest sources of envelope FFR were closer to the vertical montage than the horizontal montage. A more rostral source of envelope FFR was suggested by Galbraith (1994) based on a lack of, or inconsistent, FFR to the missing fundamental of a complex tone on horizontal montage recording. Whilst the current study found evidence of envelope following with a horizontal montage, it is possible that the envelope is not well represented in the FFR until later in the brainstem and that the horizontal montage picked this up remotely. Likewise, it is possible that the vertical montage group delay at the upper side-tone frequencies was more influenced by phase-locked activity from the auditory nerve or cochlear nucleus, because the FFR magnitude was greater in the horizontal montage than the vertical montage. Galbraith *et al*. (2000) also found FFR was not very strong in the upper brainstem and midbrain at higher frequencies.

4.5.1 **Multiple neural, and microphonic, sources per montage**

A linear fit was used to determine the group delay in the current study; this assumes that a single latency value can be ascribed to the FFR recorded at each montage. This is problematic because multiple sources with different latencies may confound the group delay estimate as they are combined into the response recorded at the scalp. The group delays calculated by the method described in the current study would not be able to disentangle multiple group delays within a single montage. As Scherg and Von Cramon (1985) demonstrate, potentials recorded at the scalp are a composite of dipole sources within the brain. A dipole model of click ABR associated wave I with the auditory nerve, but suggested later waves were a more complex combination of dipoles from multiple structures including the lateral lemniscus, trapezoid body and inferior colliculus (Scherg and Von Cramon, 1985). The dipole model, as used by Scherg and Von Cramon (1985), could not be used to determine latencies of FFR generators, because it uses tri-phasic transient waves to model the click ABR, whereas FFR is cyclic.

If two sources of FFR with differing latencies contributed equally (in amplitude) to the response, it would have a phase corresponding to the difference of the phases of the source FFRs. For example, the addition of a sine wave with phase 45° and a sine wave with phase 135° results in a sine wave with phase 90° and a √2 times larger amplitude. However, this only holds for equal amplitude waves and a 90° phase difference. For differing amplitudes the resultant wave's phase is weighted towards the phase of the wave with the larger amplitude. Increasing the phase difference decreases the resultant wave's amplitude until the two waves cancel completely at a phase difference of 180°.

In terms of the current study, the extent to which multiple sources confound the group delay estimate is dependent on whether the addition or subtraction waveform is used, and possibly on which montage is considered. For example, the CM may have considerable influence on the group delay derived from the subtraction waveforms, but not the addition waveforms. Because the CM is not half-wave rectified like the neural response is, addition will cancel out the CM and subtraction will enhance the CM (Picton *et al.*, 1974; Sohmer and Pratt, 1977). Therefore, whilst one can assume the addition waveforms contain little influence of the CM on the envelope FFR group delay, the CM may have had substantial influence on the group delays derived from the subtraction waveforms (the TFS FFR). Very short group delays may indicate the data represent the CM rather than a neural response. Stillman *et al.* (1978), Sohmer and Pratt (1977) and Davis and Britt (1984) found evidence of the CM with latencies around 1 ms. In the current study, group delay from the subtraction waveforms was generally longer than 3 ms, but there were two cases that may strongly represent CM (1.4 and 0.6 ms). The extent to which CM influenced each listener's subtraction waveform FFR is unclear. It would depend on the relative strength and latency of the CM and the FFR.

4.5.2 **Implications for clinical diagnostics**

Group delay may be a useful method of determining response latency of FFR in situations where FFR onset is gradual or unclear, or where stimulus–response cross-correlation is not appropriate. However, the need for a high signal to noise ratio in the FFR, and FFR to multiple stimuli (in separate test blocks), means that group delay may prove impractical and costly in diagnostic or clinical use due to long data acquisition times. ABR latencies have been used to provide insight into hearing difficulties related to age, hearing loss and retro-cochlear disorders (Rowe III, 1978; Elberling and Parbo, 1987; Jerger and Johnson, 1988). FFR strength and group delay change with age (Clinard *et al.*, 2010; Ruggles *et al.*, 2012; Marmel *et al.*, 2013), and group delay may be a useful tool for determining changes in FFR latency with age and hearing loss, but decreases in FFR strength mean extra time and care may be needed to ensure a sufficient SNR in the FFR to accurately determine group delay in older and hearing impaired listeners. For example, it is likely that older adults will require a larger number of trials per condition than young adults to achieve satisfactory FFR strength to meet SNR criteria.

The individual differences in group delay were large (ranging about 12 ms for modulation rates and about 7 ms for side-tones). It is unlikely this was due to anatomical differences in the auditory pathway structure alone. The differences in group delay were likely to be due a complex interaction between position and orientation of FFR generators, electrode positions and the strength of FFR from multiple generators. Based on group delay, it would be difficult to confidently say where, anatomically, the FFR is generated. It may be useful, in future studies, to correlate individual differences in FFR group delay with individual differences in the latencies of wave peaks of the click-evoked ABR.

4.6 Summary

In summary, the results presented here support the assertion that a horizontal electrode montage records FFR from an earlier stage of the auditory pathway than a vertical electrode montage. However, this was the case for TFS-FFR, but not envelope-FFR; there was no evidence that envelope-FFR recorded by the two montages represents activity at different stages of the pathway. The results are consistent with previous reports of vertical and horizontal montages recording activity from different generators, and suggest that group delay can provide a measure of latency for these generators.

Chapter 5. The Effects of Age and Hearing Loss on Neural Synchrony and Interaural Phase Difference Discrimination

Andrew King, Kathryn Hopkins , and Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom.

This chapter was prepared for publication as an original research article in the Journal of the Association for Research in Otolaryngology.

5.1 Abstract

Hearing difficulties occurring in middle-age and later may be partly due to reduced neural synchrony to the temporal characteristics of sounds. Poor neural synchrony may affect sensitivity to interaural phase differences (IPDs), which are used to locate and separate sounds from different azimuths. This study aimed to determine whether changes in IPD sensitivity associated with hearing loss and aging are related to neural synchrony as recorded in the electrophysiological frequency-following response (FFR). Human listeners (N=37) varied in age (18-83 yr) and absolute threshold (−1 to 59 dB SPL at 250 and 500 Hz). They discriminated IPDs in either the temporal fine structure (TFS) or envelope of 250 or 500 Hz pure tones amplitude modulated at 20 Hz. In a second session the FFRs to four different modulated tones were measured simultaneously (tones presented dichotically, two to each ear). Pure tones of 307, 537, 357 and 578 Hz were amplitude modulated at 16, 27, 115 and 145 Hz, respectively. With absolute threshold partialled out, age correlated with poorer envelope-IPD discrimination and TFS-IPD discrimination at 500 Hz. FFR phase coherence to the TFS components, and to the 145 Hz modulation, deteriorated with increasing age irrespective of absolute threshold. Correlations between TFS-IPD thresholds and TFS-FFR, and between envelope-IPD thresholds and 145 Hz envelope-FFR, did not remain significant after age was partialled out. On the other hand, IPD thresholds still correlated with age when FFR phase coherence was partialled out. Absolute threshold was not correlated with either IPD discrimination or FFR phase coherence. Overall, the results suggest that, while the FFR may measure aspects of neural coding relevant to IPD sensitivity, other age-related factors contribute to the performance deficits.

5.2 Introduction

5.2.1 Phase locking and binaural hearing

One of the most remarkable properties of the auditory system is that neural activity follows periodicities in sound with firing intervals tightly distributed around multiple integers of the period (e.g., Galambos and Davis, 1943; Tasaki, 1954). In other words, the neural response is “phase locked” to the stimulus (Rose *et al.*, 1967). Phase locking provides a code of temporal fine structure (TFS) in each cochlear filter-band (Cariani and Delgutte, 1996). Phase locking may also represent the envelope—the slower amplitude modulations superimposed upon the TFS by multiple components interacting within the filter band—at least at early stages of the auditory neural pathway (for review, see Joris *et al.*, 2004).

Phase locking at low frequencies in each ear provides essential input for binaural stages of the auditory system that are sensitive to interaural time differences in sounds. When a sound arrives at one ear tens to hundreds of microseconds before the other ear there will be an interaural phase difference (IPD) that is retained in the phase-locked responses from each cochlea (Jeffress, 1948; Rose *et al.*, 1966). IPDs are one of the most important cues for the perception of auditory space on the horizontal plane, for example, in localisation (Wightman and Kistler, 1992), separation of sounds from different directions (Bronkhorst, 2000; Behrens *et al.*, 2008; Marrone *et al.*, 2008b), or the detection of tones or speech in noise with different IPDs (Hirsh, 1948; Licklider, 1948; Webster, 1951; Hafter and Carrier, 1970).

5.2.2 Deficits due to aging

Compared to young adults, older humans are less accurate at localising sounds (Abel *et al.*, 2000; Dobreva *et al.*, 2011), discriminating between sounds from different angles (Briley and Summerfield, 2014; Freigang *et al.*, 2014), and detection of signals with IPDs in noise without IPDs (Pichora-Fuller and Schneider, 1992; Grose *et al.*, 1994). Old age is also associated with poor speech perception in noisy backgrounds where spatial cues can be used to benefit speech perception (Duquesnoy, 1983; Gelfand *et al.*, 1988; Gatehouse and Noble, 2004; Helfer and Freyman, 2008; Marrone *et al.*, 2008a; Marrone *et al.*, 2008c). Age-related declines in sensitivity to IPDs in the TFS of tones emerge as early as middle age (40 to 60 years), both in terms of smallest IPD discriminable from zero (Grose and Mamo, 2010; Hopkins and Moore, 2011; Moore *et al.*, 2012a; Moore *et al.*, 2012b; Neher *et al.*, 2012; Füllgrabe, 2013) and highest carrier frequency in which an IPD can be detected (Ross *et al.*, 2007; Grose and Mamo, 2010; Neher *et al.*, 2011). Deficits in IPD sensitivity have been correlated with deficits in speech reception with speech maskers laterally separated from the target speech (Neher *et al.*, 2011); both these deficits may be due to aging effects on the auditory system (Neher *et al.*, 2012).

One hypothesis is that IPD sensitivity declines with age because the phase locking that encodes TFS (and possibly envelope) is unreliable or poorly transmitted to the binaural processor of the auditory system. This 'input-deficit' hypothesis is indirectly supported by evidence of degraded TFS and envelope sensitivity after middle age. For example, there are age-related declines in discrimination of harmonic and inharmonic complexes that differ primarily in TFS (often referred to as the TFS1 task; Hopkins and Moore, 2011; Moore *et al.*, 2012b; Füllgrabe, 2013), and detection of amplitude modulation as a function of modulation frequency (Purcell *et al.*, 2004; Leigh-Paffenroth and Fowler, 2006; He *et al.*, 2008; Grose *et al.*, 2009b). However, Hopkins and Moore (2011) and Moore *et al*. (2012b) found that TFS1 scores and IPD discrimination were only moderately correlated and had different relations to hearing loss; TFS1 score correlated with absolute threshold whereas IPD discrimination did not. From this, Moore *et al*. (2012b) suggested that age-related declines in IPD discrimination do not reflect peripheral processing.

Nevertheless, age-related declines in IPD discrimination may be caused by supra-threshold changes in peripheral processing (unrelated to absolute threshold) before binaural processing. Strelcyk and Dau (2009) found a strong relation between monaural low-rate frequency modulation detection (a behavioural measure thought to reflect TFS processing; e.g., Moore and Sek, 1995; 1996) and IPD-based lateralisation measured at the same carrier frequency, even after audiometric thresholds were partialled out. From this, Strelcyk and Dau (2009) inferred that IPD discrimination is limited by the same peripheral TFS sensitivity as low-rate frequency modulation detection. Furthermore, there is evidence that auditory nerve fibres lose their connections to the inner hair cells with age and noise trauma (Kujawa and Liberman, 2009; Lin *et al.*, 2011; Makary *et al.*, 2011; Sergeyenko *et al.*, 2013). This deafferentation is selective to fibres effective at moderate to high sound levels, not fibres effective around threshold levels (Schmiedt *et al.*, 1996; Furman *et al.*, 2013).

This may explain why considerable damage to the auditory nerve may not affect absolute thresholds greatly (e.g., Schuknecht and Woellner, 1955), but fewer connected fibres may reduce the fidelity of the phase locking required for IPD discrimination, particularly at high sound levels.

5.2.3 Phase locking measured through the FFR and the effects of age

The link between performance on a behavioural task and phase locking in the auditory system is indirect, and based on assumptions. However, it is also possible to measure neural synchrony, at least on a neural group population level, through the scalp-recorded frequency-following response (FFR). The FFR is an auditory evoked potential generated by many neurons firing in synchrony with a stimulus (Worden and Marsh, 1968). This objective measurement of phase locking may be helpful in identifying the underlying mechanisms of behavioural performance (Du *et al.*, 2011; Bharadwaj *et al.*, 2014). Changes in the FFR have been associated with age (Clinard *et al.*, 2010; Vander Werff and Burns, 2011; Marmel *et al.*, 2013) and with hearing loss (Plyler and Ananthanarayan, 2001).

Clinard *et al*. (2010) found age-related declines in FFR amplitude and phase coherence to pure tones around 1 kHz, and in behavioural pure-tone frequency discrimination at 1 kHz. At and around 500 Hz however, only frequency discrimination deteriorated with age. Neither FFR amplitude, nor phase coherence, predicted the age-related changes in frequency discrimination. Marmel *et al*. (2013) found that age was related both to pure-tone frequency discrimination at 600 Hz and to FFR strength, and that FFR strength accounted for the age-related changes in frequency discrimination. They found that absolute threshold was also related to frequency discrimination, and that this relation was independent of FFR strength.

Purcell *et al*. (2004) tested envelope-FFR magnitude to white noise with sinusoidal 25% amplitude modulation across a range of modulation rates from 20 to 600 Hz in sleeping young and old listeners. Between 45 and 60 Hz, and above 100 Hz, Purcell *et al*. (2004) found that the younger group had higher envelope-FFR magnitude than the older group. They suggested that this may be due to reduced temporal acuity in the aging nervous system.

5.2.4 Phase locking and hearing loss

As well as age, sensorineural hearing loss is associated with poorer than normal processing of TFS, including the discrimination of IPDs (Lacher-Fougère and Demany, 2005; Strelcyk and Dau, 2009; Hopkins and Moore, 2011; King *et al.*, 2014). Whether this is due to poorer phase locking is not clear yet. Marmel *et al*. (2013) did not find a relation between FFR to pure tones and absolute threshold, but Plyler and Krishnan (2001) found that listeners with elevated thresholds had poorer FFR phase locking to the second formant of a speech sound. Miller *et al*. (1997) showed diminished phase locking to the second and third formants of the /ϵ/ vowel in cats' auditory-nerve fibres after acoustic trauma, but normal *or better* phase locking at the fibre's most sensitive frequency. There is evidence for (Woolf *et al.*, 1981) and against (Harrison and Evans, 1979) impaired phase locking in the auditory nerve of animals after drug-induced damage to outer hair cells. Kale and Heinz (2010) found noise-induced hearing loss in chinchillas did not reduce phase locking to the TFS, but rather increased the envelope phase locking in the cells recorded.

The present study aimed to determine whether changes in TFS- and envelope-IPD discrimination thresholds seen with increasing age and absolute threshold could be explained by changes in FFR that reflect neural coding of TFS and envelope respectively. Furthermore, this study also aimed to determine whether age-related and hearing-loss-related changes in FFR occur in FFR generated earlier or later in the auditory pathway. This was done by measuring the FFR simultaneously with two montages, one oriented vertically and one oriented horizontally as in Chapter 4. The results of Chapter 4 suggest that the horizontal montage records earlier-generated TFS FFR than the vertical montage. If age-related or hearing-loss-related deficits in FFR are only seen in the FFR recorded with the vertical montage, this may indicate that deficits in phase locking occur in central processing. On the other hand, if age or hearing loss is only associated with deficits in the FFR recorded with the horizontal montage, this may because age or hearing loss affects phase locking on the auditory nerve.

5.3 Methods

5.3.1 Listeners

Thirty-seven listeners completed both the behavioural session (IPD discrimination) and the electrophysiological session (FFR recording). They ranged from 18 to 83 years old. The mean age was 50.3 years, but the sample distribution was bimodal, with one mode around 20 to 30 years and another around 60 to 70 years. Air- and bone-conduction pure-tone audiometry was performed on all listeners in accordance with the British Society of Audiology (2011) recommended procedure. Mean audiometric thresholds across 0.25, 0.5, 0.75, 1 and 1.5 kHz were 7 dB HL for the listeners below 40 years old and 16 dB HL for the listeners 40 years old and above. Mean audiometric thresholds across 2, 3, 4, 6 and 8 kHz were 13 dB HL for the listeners below 40 years old and 39 dB HL for the listeners 40 years old and above. Listeners with suspected conductive hearing loss, or asymmetry between ears greater than 15 dB below 1 kHz, were excluded. Listeners with suspected mild or moderate sensorineural hearing loss were included. However, only 10 listeners had low-frequency (250 and 500 Hz average) absolute thresholds equal to, or greater than, 20 dB SPL. Pure-tone absolute thresholds were determined monaurally by a three-interval, three-alternative forced-choice task with a two-down, one-up adaptive tracking method (71% correct; Levitt, 1971). See Chapter 3 or King *et al*. (2014) for more details. Listeners' absolute thresholds (averaged across 250 and 500 Hz and across left and right ears) and audiometric thresholds at high frequencies are plotted as a function of listener age in Figure 5.1 with open symbols indicating the listeners who also volunteered for King *et al*. (2014).

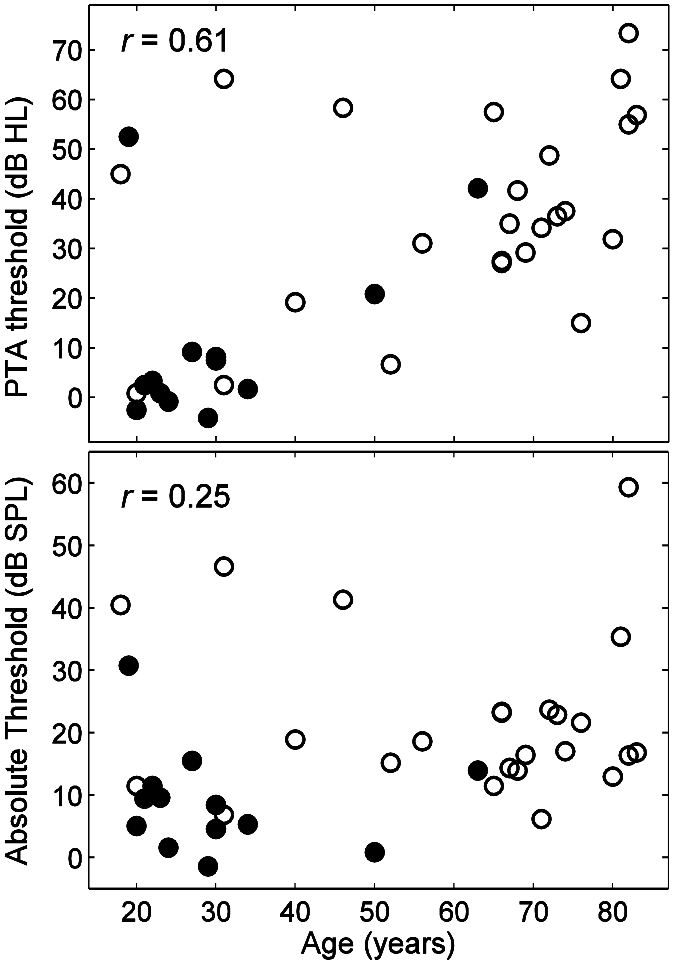

Figure 5.1: Listeners' low-frequency absolute thresholds (250 and 500 Hz average; bottom panel) and high-frequency audiometric thresholds (2 kHz to 8 kHz average; top panel) are plotted as a function of listener age. In both panels, thresholds are an average of left and right ears for each listener. Pearson's r is inset for each panel. Open symbols indicate listeners whose data were also presented in King *et al.* [2014; JASA 135(1), 342–351]. Filled symbols indicate additional listeners who were recruited for the current study specifically.

5.3.2 **IPD discrimination**

Thresholds for discriminating an IPD in an amplitude-modulated (AM) tone were measured using the stimuli and procedure described in King *et al.* (2014).

5.3.2.1 Stimuli

Pure tones of 250 or 500 Hz were sinusoidally amplitude modulated at 20 Hz. The tones were 500 ms long with 50 ms onset and offset ramps. The ramps were synchronous across ears. For the target stimuli, an IPD was generated in either the

TFS or the envelope by adding a positive starting phase to the carrier or modulation, respectively, in the signal played to the ear with the lower absolute threshold at the carrier frequency. In the reference stimuli, both TFS and envelope started in sine (0°) phase in both ears. The binaural, AM tones used for IPD discrimination were played at either 30 dB above the absolute threshold in each ear, or 30 dB above the mean absolute threshold across ears, whichever was reported by the listener as sounding the least lateralized when presented without an IPD.

5.3.2.2 Procedure

IPD discrimination thresholds were determined using a two-interval, two-alternative forced-choice task with a two-down, one-up adaptive tracking method. The reference interval consisted of four bursts of the reference stimulus (RRRR), whilst the target interval had an IPD in the second and fourth bursts, creating an alternating pattern (RTRT). Each burst lasted 500 ms and bursts were separated by 20 ms; the two intervals were separated by 500 ms. The adaptive tracks started with an IPD of 180° and stepped by a factor of 1.25^2 until the third reversal, then by a factor of 1.25 for a further eight reversals. The adaptive track threshold was calculated as the geometric mean of the IPDs at the last eight reversals. The maximum IPD was limited to 180°, above which an IPD does not become easier to detect. If a listener failed to detect an IPD of 180° in any trial after the initial three reversals, the adaptive track stopped and 40 trials with an IPD of 180° were presented. From the percent correct in these 40 trials, an estimate of IPD threshold at 71% correct was extrapolated (see the procedure in King *et al*., 2014).

Four adaptive tracks of IPD discrimination were conducted for each of the four conditions: a carrier frequency (f_c) of 250 Hz and a modulation frequency (f_m) of 20 Hz, with IPDs in the carrier (TFS250) or in the modulation (ENV250), and an f_c of 500 Hz and a f_m of 20 Hz, with IPDs in the carrier (TFS500) or the modulation (ENV500). For each condition, the listener’s threshold was taken as the geometric mean of the four adaptively tracked, or extrapolated, threshold estimates.

5.3.2.3 Apparatus

All audio stimuli for absolute and IPD threshold measurement were created in MATLAB 7.6 (The MathWorks, 2008). Sounds were converted from digital to analog at a sample rate of 48 kHz and a 24-bit depth and amplified using a Creative E-MU

0202 USB soundcard. Sounds were played over Sennheiser HD 650 circumaural headphones. Listeners sat in a double-walled listening booth and made responses via a computer keyboard. Audiometric thresholds were measured using VIASYS GSI-Arrow and Kamplex AC30 audiometers coupled to TDH39P supra-aural headphones and Radioear B-71 bone vibrators.

5.3.3 **FFR recording**

FFRs to AM tones were used to examine the representation of both envelope and TFS. AM tones without binaural information were used to test the extent to which neural synchrony to signals at each ear might determine IPD sensitivity.

5.3.3.1 Stimuli

Four AM tones were played to the listeners simultaneously. Each AM tone was created in MATLAB 7.12 (The MathWorks, 2011) by summing three equal-amplitude pure tones, one at f_c, one at $f_c – f_m$, and one at $f_c + f_m$ (see Table 5.1). Using equal-amplitude components for each AM tone resulted in non-sinusoidal modulation with smaller peaks half-way between the main peaks at the f_m[1]. The tones were presented dichotically, with two tones to each ear.

The left ear received the following:

- A tone with an f_c of 307 Hz and an f_m of 16 Hz (Figure 5.2, top left),
- A tone with an f_c of 578 Hz and an f_m of 145 Hz (Figure 5.2, middle left).

The right ear received the following:

- A tone with an f_c of 357 Hz and an f_m of 115 Hz (Figure 5.2, top right),
- A tone with an f_c of 537 Hz and an f_m of 27 Hz (Figure 5.2, middle right).

---

[1] The use of equal-amplitude components in the AM tones used for measuring FFR resulted in over-modulation, that is, modulation over 100% sinusoidal modulation. This means that instead of one peak and trough every cycle, there was a main peak, then a second peak ‘filling in’ the trough. This may have resulted in a peak in the FFR at the second harmonic ($2f_m$). However, as the peripheral auditory filters are not truly rectangular in the frequency domain, the filter centred on f_c will attenuate the sidebands, potentially reducing the "over-modulation". Auditory filters centred near the f_c may only pass two of the three components. More basally centred filters may pick up the over-modulation, but it is likely that the larger temporal envelope peaks at intervals corresponding to f_m will dominate the FFR response over the smaller peaks in-between.

This way, each ear received a tone with a ‘high’ f_c roughly corresponding to TFS500 and a tone with a ‘low’ f_c roughly corresponding to TFS250. Each ear received a ‘low’ f_m roughly corresponding to the 20 Hz f_m used for the ENV250 and ENV500 conditions. However, envelope FFR at such low f_m frequencies is likely to contain strong contributions from cortical activity (Herdman *et al.*, 2002; Kuwada *et al.*, 2002). The ‘high’ f_m rates of 115 and 145 Hz were included in order to measure envelope FFR that would likely be generated in the subcortical auditory pathway. Frequencies were selected to limit the possibility of shared components or harmonics.

Presenting multiple complex tones simultaneously and measuring the envelope FFR to various modulation rates simultaneously has precedent (Dolphin, 1996; Dolphin, 1997), and is used in the measurement of brainstem auditory steady-state responses as an objective measure of audiometric threshold (e.g., Lins *et al.*, 1996). Lins *et al.* (1994; 1996) found no significant changes in envelope FFR at any of their combinations of AM tones. Dolphin (1997) found envelope FFR to a probe tone pair could be slightly reduced (by up to 3 dB) by adding higher frequency tone pairs (with higher envelope rates), but could be slightly enhanced when the added tone pair (and envelope rate) were lower in frequency than the probe. Both enhancement and reduction were less marked at higher stimulus levels (Dolphin, 1997).

Table 5.1: The modulation and spectral components of the AM tones presented to the listener during the FFR recording.

f_m region	Low		High	
f_c region	Low	High	Low	High
f_m	16	27	115	145
$f_c - f_m$	291	510	242	433
f_c	307	537	357	578
$f_c + f_m$	323	564	472	723
Ear	Left	Right	Right	Left

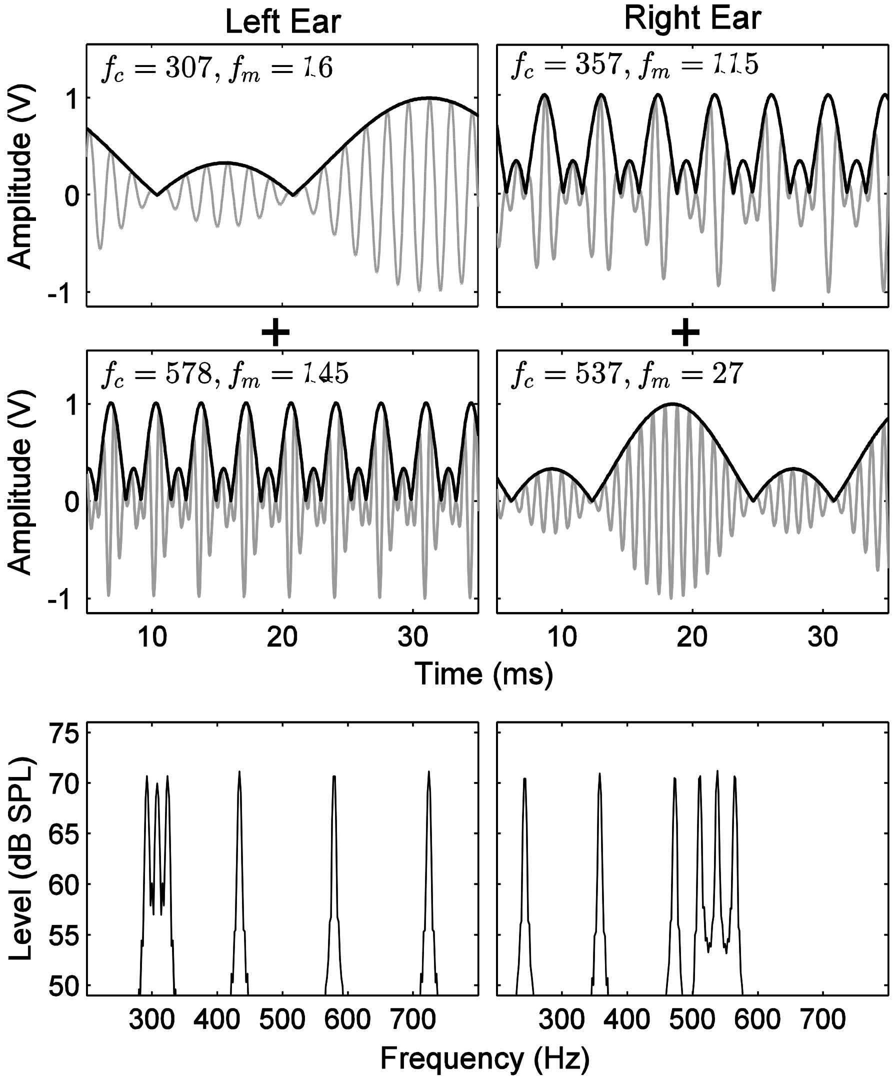

Figure 5.2: The four separate AM tones shown in the time domain with f_c and f_m inset (top and middle panels) and the combined pairs used for each channel (for each ear, given at the top of the figure) shown in the frequency domain (bottom two panels).

Simultaneous presentation of stimuli allows the number of responses collected to be increased by a factor equal to the number of stimuli. Another benefit to simultaneous presentation is that all stimulus conditions share the same noise in the response. As wakefulness and myogenic noise can vary over the course of a recording session, this is an important advantage over consecutive presentation blocks for each stimulus condition. Figure 5.3 shows a comparison of the response to a single AM tone (f_m=16 Hz; f_c=307 Hz) played to the left ear, the response to both this tone and the tone with an f_m of 145 Hz (f_c=578 Hz) played to the left ear simultaneously, and the response to all four tones in the experimental setup (both ears). The envelope-FFR (addition of FFR_0 and FFR_{180}; see Procedure section 5.3.3.3) appears enhanced at the

16 Hz when multiple AM tones are played simultaneously (Figure 5.3, panels A and C). Panels B and D in Figure 5.3 show TFS FFR (subtraction of FFR_0 and FFR_{180}) at the shared components is similar when either one or two AM tones are played monaurally. When all four tones are played, TFS FFR at these components is slightly reduced, but peaks in the spectrum are still evident. These preliminary recordings suggest that presenting four tones simultaneously allowed good phase-locked responses to most of the tone components and modulation rates, without much detriment compared to separate presentation.

AM tones were created at a sample rate of 24414 Hz and a bit depth of 2^{16}. The tones were summed into the left and right ear channels (see Figure 5.2, bottom left and right panels) and the stimuli in each channel were presented at an overall level of 80 dB SPL. This reasonably high level was used to ensure that all participants could hear the stimuli, including those with low-frequency hearing loss. Furthermore, as FFR amplitude increases with stimulus intensity (Marsh *et al.*, 1975; Krishnan, 2002), a high sound level should improve the signal-to-noise ratio of the recorded responses.

5.3.3.2 Apparatus

Stimuli were presented using Tucker Davis Technologies (TDT) SigGen and BioSigRP software. Stimuli were converted to analogue signals by a TDT RP2.1 processor and transduced to acoustic waves outside the listening booth by ER30 earphones (Etymotic Research, Illinois) to minimize stimulus artefact. The transducers were connected to the listener inside the listening booth by 6 m of tubing terminating in the listener’s external auditory canals through foam earplugs. Pilot tests showed that no stimulus artefact was recorded when the stimulus was presented with the tubing detached from the earplugs and sealed with tape.

The electrical field potential was recorded by two montages of gold-plated passive electrodes. The horizontal montage recorded at the ipsilateral mastoid (referenced to the contralateral mastoid) and the vertical montage recorded at the seventh cervical vertebra (referenced to the hairline the forehead on the sagittal line). Both montages shared a common ground electrode on the listener’s brow. Based on the results of Chapter 4, the horizontal montage was expected to mostly record TFS FFR from an earlier source in the auditory pathway than the vertical montage. Therefore, comparing FFR measured by the horizontal and vertical montages may be able to distinguish whether age or hearing loss affected phase locking to the TFS of stimuli differently in the earlier and later stages of the subcortical auditory pathway. This was done by testing the interactions between the effects of age and montage and between the effects of hearing loss and montage on the phase coherence of the TFS FFR.

The electrodes were wired into a TDT RA16LI-D head-stage pre-amplifier linked to a TDT RA4PA preamplifier which digitized the analogue signal at 24414 Hz and sent this to a TDT RA16 Medusa Base Station for processing. The base station and RP2.1 processor were synchronized and both communicated with the BioSigRP software. The BioSigRP software filtered the recordings online with a 10 Hz high-pass filter, a 3 kHz low-pass filter and a 50 Hz notch filter.

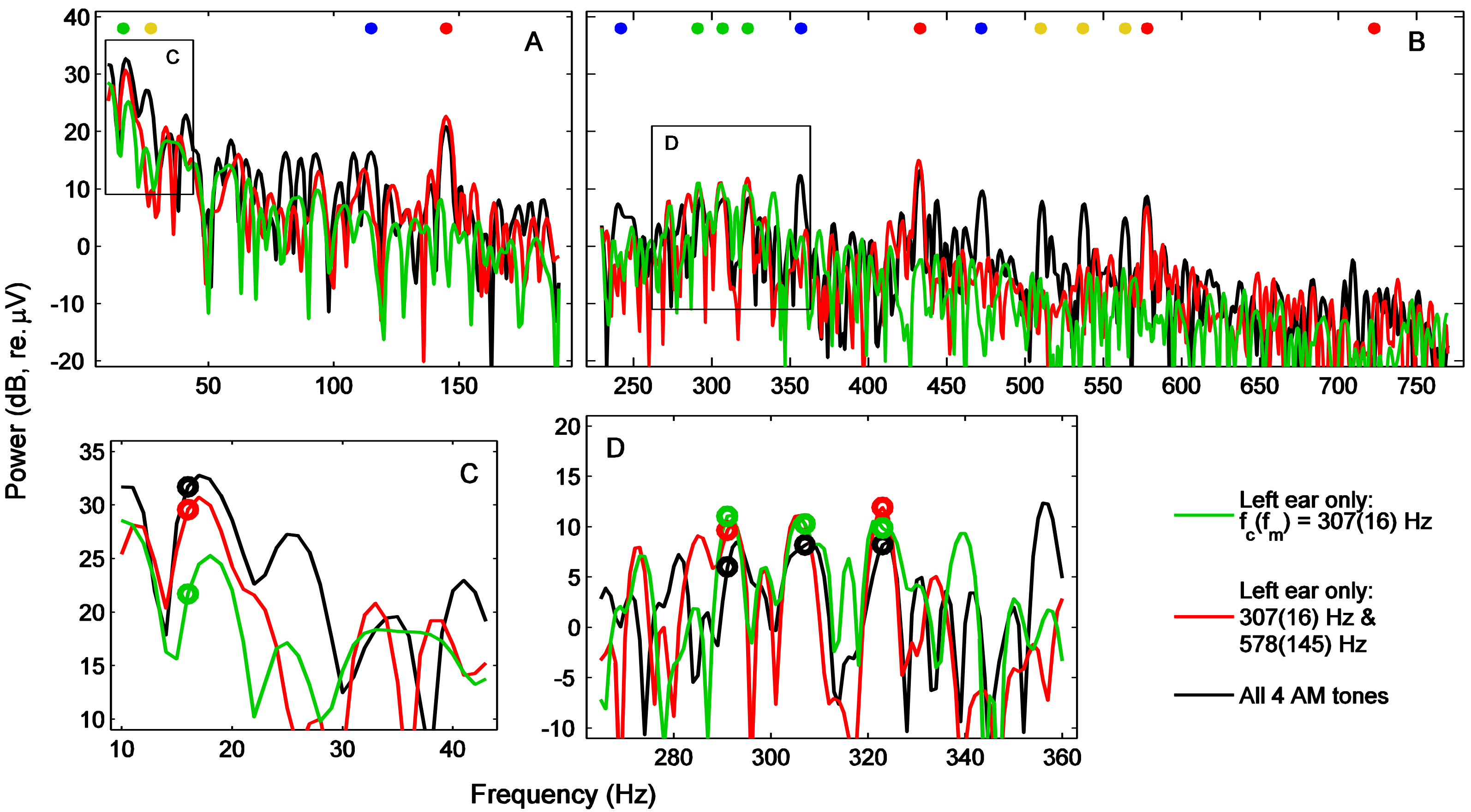

Figure 5.3: Comparison of FFR to a single, monaural AM tone (green trace), FFR to the two AM tones played to the left ear (red trace), and FFR to all four tones, two played to each ear (black trace), all from an individual listener (author AK). Each trace is an average of 80 sub-averages. Filled circles in panels A and B denote where on the frequency axis FFR should be seen (with each AM tone represented by a color). Open circles highlight the FFR power in each trace at 16 Hz f_m (panel C) and that tone's component frequencies (panel D).

5.3.3.3 Procedure

AM tones were presented in alternating starting polarity, so each recording had a response to the stimulus with all components starting in 0° phase (FFR_0) followed by a response to the stimulus with all components starting in 180° phase (FFR_{180}). By adding FFR_0 and FFR_{180}, the phase-locked responses reflecting the envelope interfere constructively, whilst the responses reflecting the TFS interfere destructively. On the other hand, subtracting FFR_{180} from FFR_0 results in constructive interference of the responses reflecting the TFS and destructive interference of the responses reflecting the envelope (Aiken and Picton, 2008). These methods of addition and subtraction were used to quantify the envelope-FFR and TFS-FFR respectively.

Individual trials with peak amplitudes exceeding ±100 μV in the vertical montage, and ±80 μV in the horizontal montage, were rejected from the sub-averages. Sub-averages consisted of 50 accepted trials. One hundred sub-averages in total were collected from each listener, taking approximately 80 minutes. Listeners sat in an arm-chair and were asked to relax as much as possible and try to let their head be supported by the chair headrest. This was done to reduce the myogenic noise in the recordings. The listeners watched a film of their choice, with the audio disabled and sub-titles enabled, on a portable digital-versatile-disc (DVD) player sat upon a desk in front of the listener. This was done to try to keep the listener awake, as at low modulation rates, wakefulness affects envelope FFR strength (e.g., Purcell *et al.*, 2004). However, wakefulness was not recorded. To test for DVD-player artefacts in either amplitude or phase coherence spectrum, two sessions of 100 sub-averages, without stimuli, were run in a pilot test. One session included the DVD player in the booth playing a film; the other session did not. Comparison of spectra with and without the DVD player did not reveal any artefact.

Phase coherence was calculated to quantify the quality of phase locking in the FFR. Phase coherence is the vector average (see equation 5.1) of the phase angle (θ) at a given frequency of the Fourier transform across *n* (sub-averages in the current study). Equation 5.1 was taken from Dobie and Wilson (1989):

$$PC = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos\theta_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin\theta_i\right)^2} \qquad (5.1)$$

where *i* is the index of sub-averages.

5.4 Results

5.4.1 **IPD discrimination**

Figure 5.4 shows the envelope- and TFS-IPD thresholds as a function of age (top-left and top-right panels, respectively) and as a function of absolute threshold (bottom-left and bottom-right panels, respectively). The absolute thresholds used in all analyses were the average across 250 and 500 Hz and across ears for each listener. IPD thresholds were log-transformed for correlation (and partial correlation) with age and absolute threshold, because this transform provided normally distributed IPD thresholds. Correlation coefficients are given in Table 5.2. There were moderate, positive correlations between age and ENV250 and between age and ENV500. These correlations remained significant after absolute threshold was partialled out, suggesting that old age is related to poorer envelope-IPD sensitivity, regardless of absolute threshold. There was a moderate, positive correlation between TFS500 and age that remained significant after absolute threshold was partialled out. On the other hand, the weak positive correlation between TFS250 and age did not survive after absolute threshold was partialled out. TFS-IPD thresholds also became more varied with age. The difference in threshold variation in the listener group under 60 years old was found to be significantly smaller ($X^2=4.39$, $p<0.05$) than the variation in the over 60 years old listener group by Bartlett's test of population variances (Bartlett, 1937).

Table 5.2: The Pearson's *r* correlation and partial correlation coefficients between thresholds in each IPD condition (columns) and age and absolute threshold. One, two and three asterisks denote significant correlations for alphas of 0.05, 0.01 and 0.001, respectively.

	IPD condition			
	ENV250	ENV500	TFS250	TFS500
Age	0.55***	0.53***	0.33*	0.46**
(absolute threshold partialled out)	0.56***	0.59***	0.30	0.43**
Absolute threshold	0.01	−0.16	0.20	0.20
(age partialled out)	−0.17	−0.34*	0.11	0.11

There were no significant correlations between absolute threshold and IPD threshold. ENV500 was weakly correlated with absolute threshold when age was partialled out (*r*=−0.34, *p*=0.04), but this was not significant after correction for multiple comparisons. Observed power for the correlations was 0.75, based on the mean coefficient of determination (r^2) of the partial correlations as the effect size, 37 listeners and an α of 0.05. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

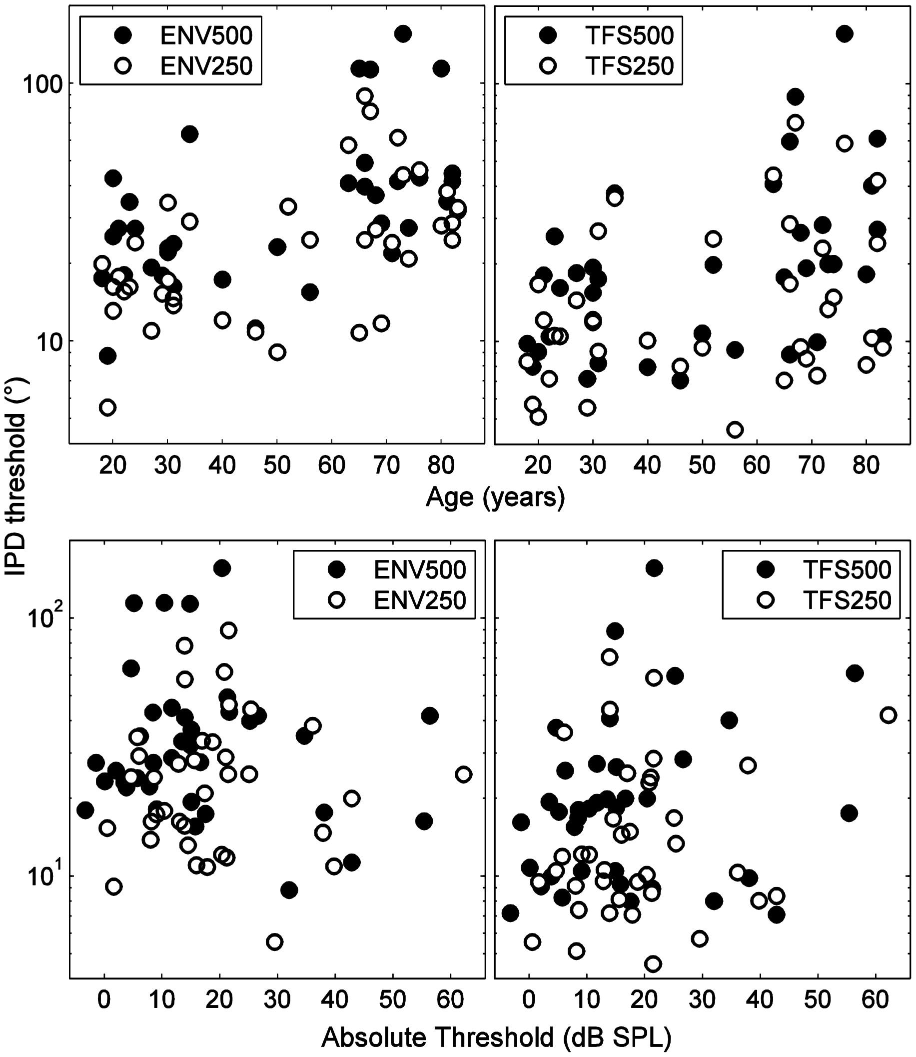

Figure 5.4: Envelope-IPD and TFS-IPD thresholds (left and right panels, respectively) are plotted as a function of listener age in the top two panels and as a function of absolute threshold (250 and 500 Hz average) in the bottom two panels. Open symbols show the IPD thresholds with the 250 Hz carrier (ENV250 and TFS250) and filled symbols show the IPD thresholds with the 500 Hz carrier (ENV500 and TFS500).

5.4.2 **FFR**

5.4.2.1 Envelope FFR

Mean phase-coherence spectra are plotted for listeners grouped by age in Figure 5.5 (top four panels). Herein, the labels EFFR16, EFFR27, EFFR115 and EFFR145 are used to refer to the phase coherence in the envelope FFR (addition waveforms) at an f_m of 16, 27, 115 and 145 Hz respectively. The vertical montage spectrum from the addition waveforms (enhancing the FFR that reflects phase locking to the envelope) indicated slightly stronger EFFR16 for the older listeners than younger listeners, but slightly weaker EFFR145 for older listeners than younger listeners. Weaker EFFR145 was also seen for older listeners in the horizontal montage, but no other differences between age groups are clear in this montage. When listeners were grouped by absolute threshold (averaged across 250 and 500 Hz and across ears) into those above the median (15 dB) and those below the median (bottom four panels of Figure 5.5), mean EFFR145 in both montages was only slightly stronger for those with absolute thresholds below the median compared to those with thresholds above it. Otherwise, mean envelope-FFR phase coherence spectra did not notably differ between absolute threshold groups in either vertical or horizontal montage.

Effects of age and absolute threshold (averaged across 250 and 500 Hz and across ears for each listener, as in IPD discrimination analysis) were tested as mean-centred covariates in a repeated-measures analysis-of-covariance (ANCOVA) with montage and f_m as within-subjects factors. There was no significant overall effect of age on envelope-FFR phase coherence [F(1,34)=1.9, $p>0.05$], and no significant overall effect of absolute threshold on envelope-FFR phase coherence [F(1,34)=1.1, $p>0.05$]. The ANCOVA showed a significant effect of montage, the vertical montage recorded stronger phase coherence than the horizontal montage [F(1,34)=23.4, $p<0.001$]. There was also an effect of f_m [F(2.7,90.2)=21.8, $p<0.001$]; a degrees of freedom correction was applied due to unequal variances in the data across conditions (Huynh and Feldt, 1976); pairwise comparisons showed that EFFR115 was significantly lower than EFFR16, EFFR27 and EFFR145, and that EFFR145 was significantly higher than EFFR16.

The interaction between the effects of montage and age was not significant [F(1,34)=0.4, $p>0.05$]. The effect of age interacted with the effect of f_m [F(2.7,90.2)=7.0, $p<0.001$]. As suggested above, EFFR16 increased with age and

EFFR145 decreased with age; correlation analysis supports this: phase coherence was averaged across the two montages for EFFR16, EFFR27, EFFR115 and EFFR145 and each of these variables was correlated against age. Table 5.3 gives the correlation coefficients and significance *p* values. The correlations between EFFR16 and age and between EFFR145 and age were still significant when absolute threshold was accounted for.

Table 5.3 Pearson's *r* correlation coefficients and significance values for mean FFR across both montages at the four envelope rates (f_m) correlated against age and against age with absolute threshold partialled out.

	EFFR16	EFFR27	EFFR115	EFFR145
Age	0.38	−0.07	−0.28	−0.49
Significance value (*p*)	0.021	0.691	0.090	0.002
Age (absolute threshold partialled out)	0.34	−0.05	−0.22	−0.46
Significance value (*p*)	0.041	0.781	0.198	0.005

Figure 5.6 shows the envelope FFR at the four f_m's as a function of listener age in the vertical montage (top row of panels) and horizontal montage (bottom row) separately. Pearson's *r* between the envelope-FFR phase coherence at each f_m and age is inset for each panel. Because the phase coherence at other frequencies than those stimulated at or modulated at was not uniform across listeners or frequency, the phase coherence noise floor in the frequency regions around the FFR of interest is also plotted for each listener in Figure 5.6 as black dots. This was the average phase coherence, excluding FFR at and around every f_m (*i.e.* f_m−4 to f_m+4). Specifically, this was between 10 and 49 Hz for EFFR16 and EFFR27, and between 90 and 170 Hz for EFFR115 and EFFR145.

The phase coherence noise floor was designed to show, for each listener, the extent to which phase coherence in response to the stimulus exceeded the ambient phase coherence unrelated to the stimulus. It should be noted that harmonics and distortion products in the FFR may have contributed to the phase coherence noise floor, despite being related to the stimulus.

There was no significant interaction between the effects of absolute threshold and montage [F(1,34)=0.04, $p>0.05$], between the effects of absolute threshold and f_m [F(2.7,90.2)=1.4, $p>0.05$], or between the effects of absolute threshold, montage and f_m [F(2.9,99.5)=0.2, $p>0.05$]. Because of the lack of a significant main effect of absolute threshold on envelope FFR, or any significant interaction, no correlational analyses or plots of envelope FFR as a function of absolute threshold are given.

There was a significant interaction between montage and f_m [F(2.9,99.5)=4.1,$p<0.05$](degrees of freedom adjusted with the Huynh–Feldt correction, Huynh and Feldt, 1976); the difference in phase coherence between montages was greater for EFFR16 and EFFR145 (0.14 for both) than EFFR27 and EFFR115 (0.07 and 0.02, respectively).

Lastly, there was a significant interaction between montage, f_m, and age [F(2.9,99.5)=5.4, $p<0.01$]. We interpret this as the interaction between age and f_m being greater for the vertical montage than for the horizontal montage.

Observed power for the correlations between envelope-FFR phase coherence and age was 0.68, based on the mean coefficient of determination (r^2) of the correlations as the effect size, 37 listeners and an α of 0.05. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

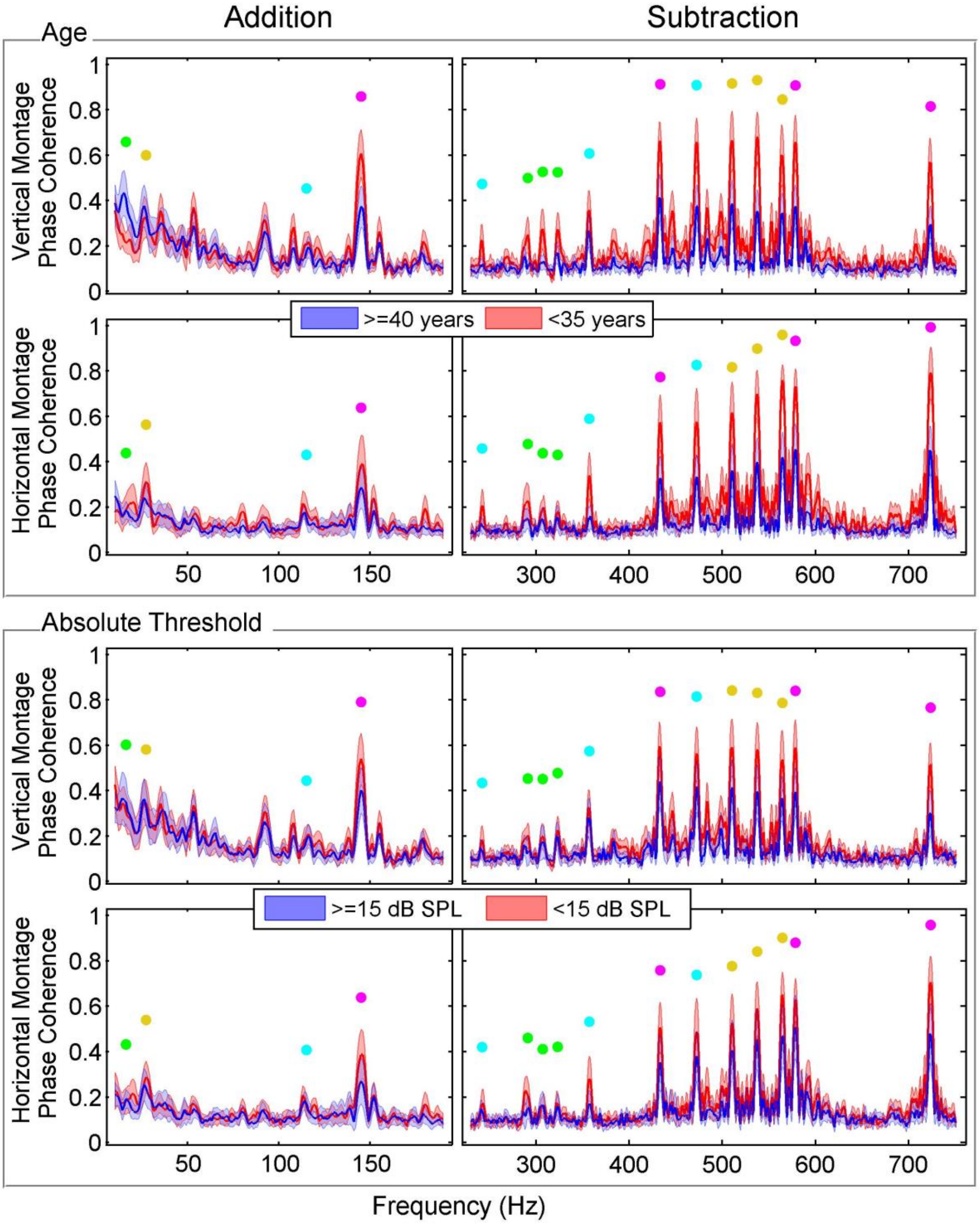

Figure 5.5: Mean phase-coherence spectra in the envelope (left panels) and TFS (right panels) FFR recorded with the vertical and horizontal montages. The top four panels show the mean spectra for listeners grouped by age: under 35 years old (red; N=15) and over 40 years old (blue; N=22). The bottom four panels show the mean spectra for the same listeners grouped by absolute threshold (250 and 500 Hz average): below the median absolute threshold (15 dB SPL) in red (N=18), and above the median in blue (N=19). Pale red and blue boundaries around the mean lines denote the 95% confidence intervals. Filled circles in each panel denote where on the frequency axis FFR should be seen (with each AM tone represented by a colour).

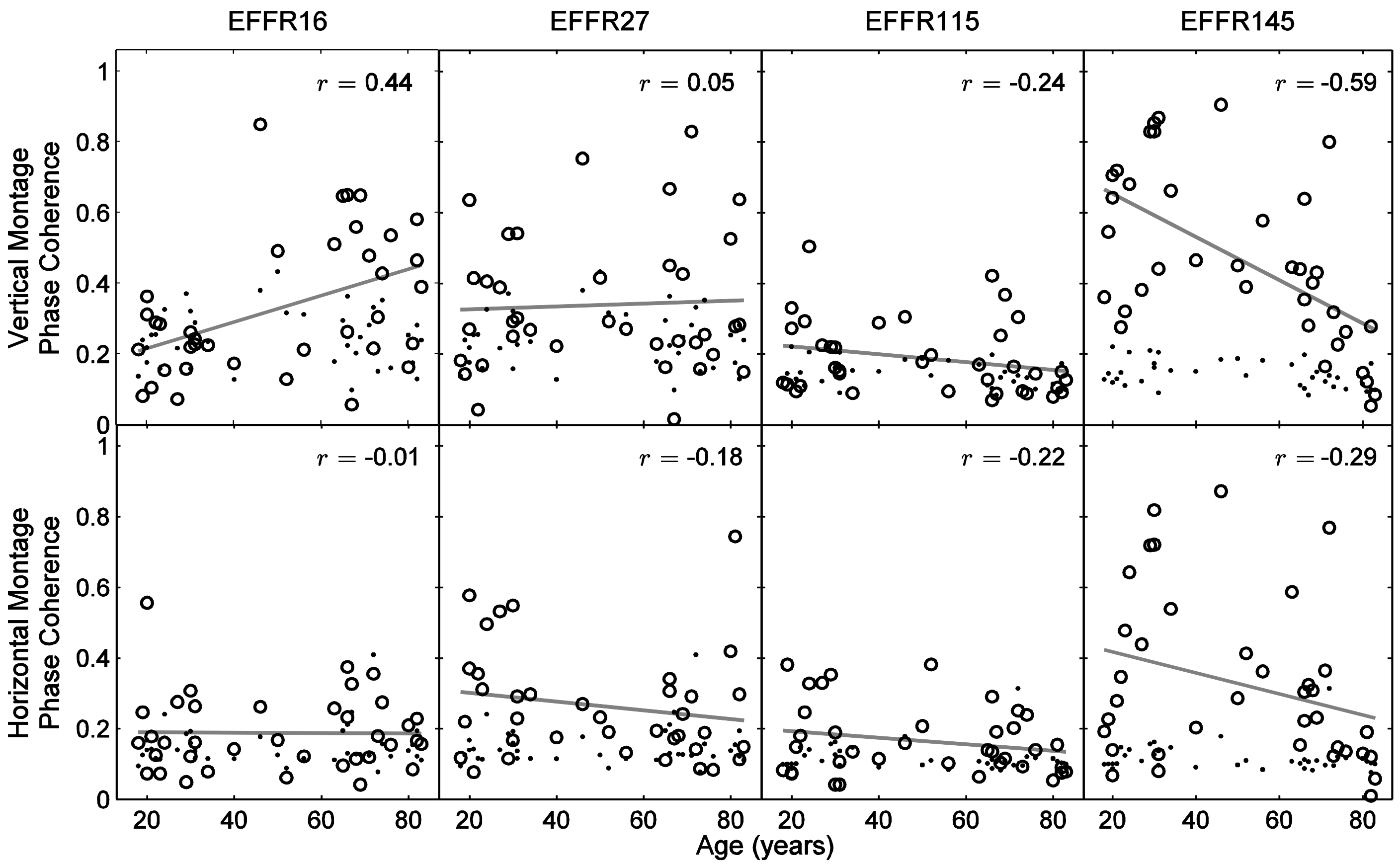

Figure 5.6: The phase coherence (unfilled circles) as a function of listener age, in the envelope FFR from the vertical montage (top row of panels) and from the horizontal montage (bottom row of panels). From left to right, the panels in both rows show the envelope FFR at 16 Hz (EFFR16), 27 Hz (EFFR27), 115 Hz (EFFR115) and 145 Hz (EFFR145). The black dots show the phase coherence noise floor in the addition waveform for each listener (see main text for details of noise floor calculation). The grey lines show the linear fits between FFR and age in a least squares sense with correlations (Pearson’s r) inset.

5.4.2.2 TFS FFR

In the phase-coherence spectra from the subtraction waveforms (right panels of Figure 5.5) differences between the two age groups and between the two absolute-threshold groups can be seen in the FFR reflecting phase locking to the TFS. Younger listeners had stronger phase coherence than older listeners at most AM-tone component frequencies. Listeners with lower absolute thresholds had slightly stronger phase coherence than listeners with higher absolute thresholds. Again, effects of age and absolute threshold (averaged across ears and 250 and 500 Hz for each listener) were tested as covariates in a repeated-measures ANCOVA with montage and component frequency region as within-subjects factors. Rather than test each component frequency as an individual factor level, the phase coherences were grouped into two groups of components, a lower region (242, 291, 307, 323 and 357 Hz; $\mathrm{TFSFFR}_{\mathrm{LOW}}$) and a higher region (433, 472, 510, 537, 564, 578 and 723 Hz; $\mathrm{TFSFFR}_{\mathrm{HIGH}}$). This was based on the observation that the phase coherences at the AM-tone component frequencies in the lower region were consistently weaker than at the AM-tone component frequencies in the higher region.

There was a significant main effect of age [F(1,34)=34.6, p<0.001]; older listeners had weaker phase coherence than younger listeners. There was a significant main effect of frequency region [F(1,34)= 238.8,p<0.001]; phase coherence in $\mathrm{TFSFFR}_{\mathrm{LOW}}$ was weaker (mean=0.19) than in $\mathrm{TFSFFR}_{\mathrm{HIGH}}$ (mean=0.49). There was no significant main effect of montage [F(1,34)=0.03, p>0.05]. However, the interaction between montage and frequency region was significant [F(1,34)=6.9, p<0.05]; the difference in phase coherence between $\mathrm{TFSFFR}_{\mathrm{LOW}}$ and $\mathrm{TFSFFR}_{\mathrm{HIGH}}$ was greater in the horizontal montage (0.33) than in the vertical montage (0.27).

There was a significant interaction between frequency region and age [F(1,34)=26.9,p<0.001]. Figure 5.5 shows that there was a greater difference between young and older listeners' phase coherence in $\mathrm{TFSFFR}_{\mathrm{HIGH}}$ than in $\mathrm{TFSFFR}_{\mathrm{LOW}}$ in both the vertical and horizontal montages.

This interaction is also shown in the top and bottom panels of Figure 5.7 (for vertical and horizontal montages respectively) as a function of listeners' age with Pearson's *r* inset. The phase coherence noise floor for the TFS FFR is plotted also. Similar to the noise floor for the envelope FFR in Figure 5.6, it was calculated as the average phase coherence, excluding FFR at and around the component frequencies (*i.e.* *f*−4 to *f*+4 for every component) between 200 and 400 Hz for $\mathrm{TFSFFR}_{\mathrm{LOW}}$ and between 410 and 740 Hz for $\mathrm{TFSFFR}_{\mathrm{HIGH}}$.

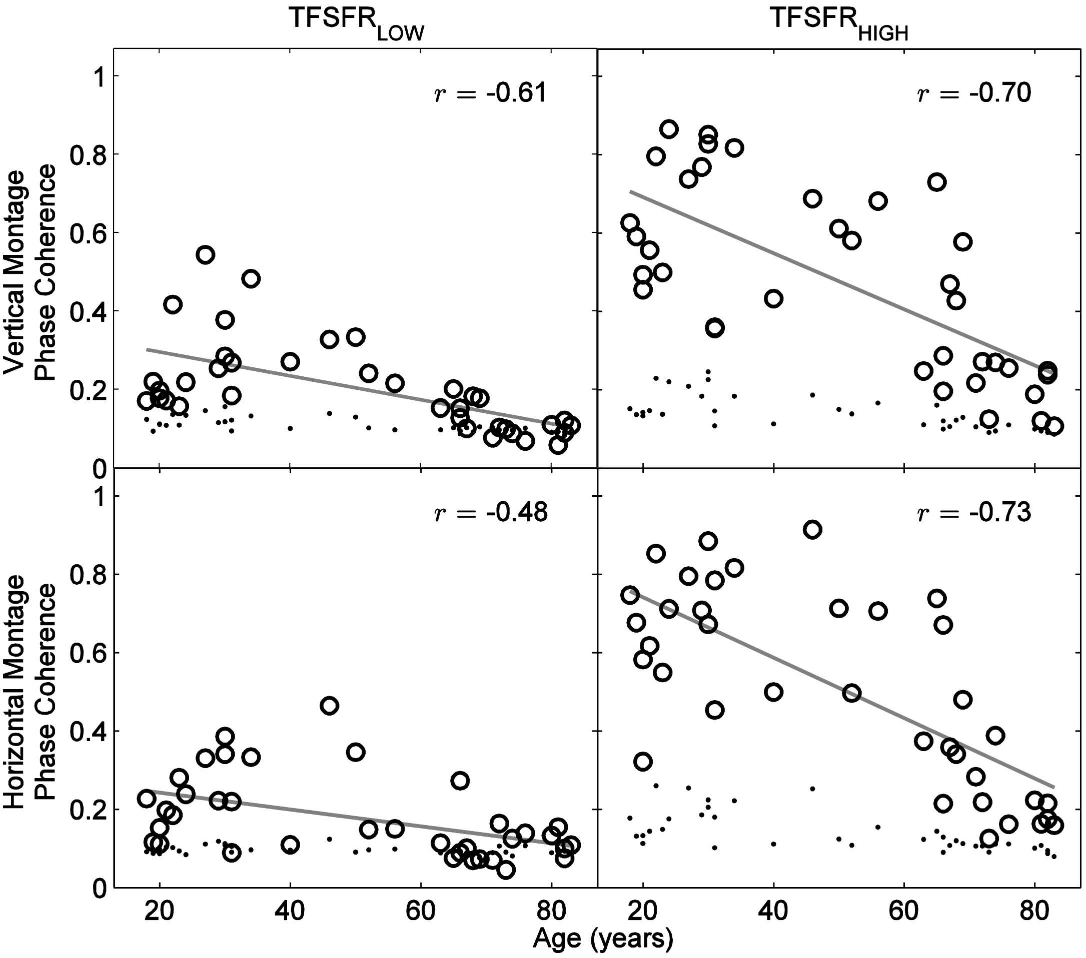

Figure 5.7 The TFS-FFR phase coherence from the subtraction waveforms recorded with the vertical (top panels) and horizontal (bottom panels) is plotted as a function of listener age (unfilled circles). The left panels show TFS-FFR phase coherence averaged across the stimulus components in the lower frequency region (242 to 357 Hz; $\mathrm{TFSFFR}_{\mathrm{LOW}}$). The right panels show TFS-FFR phase coherence averaged across the stimulus components in the higher frequency region (433 to 723 Hz; $\mathrm{TFSFFR}_{\mathrm{HIGH}}$). The black dots show the phase coherence noise floor in the subtraction waveform for each listener (see main text for details of noise floor calculation). The grey lines show the linear fits between FFR and age in a least squares sense with correlations (Pearson's *r*) inset.

Age was still negatively correlated with phase coherence in TFS FFR from the vertical montage when absolute threshold was partialled out ($\mathrm{TFSFFR}_{\mathrm{HIGH}}$: r=−0.69,

$p<0.001$; $\mathrm{TFSFFR}_{\mathrm{LOW}}$: $r=-0.58$, $p<0.001$). This was also true for TFS FFR from the horizontal montage ($\mathrm{TFSFFR}_{\mathrm{HIGH}}$: $r=-0.71$, $p<0.001$; $\mathrm{TFSFFR}_{\mathrm{LOW}}$: $r=-0.45$, $p<0.01$). There was no significant interaction between the effects of age and montage [$F(1,34)=0.001$, $p>0.05$], or between the effects of age, montage and frequency region [$F(1,34)=2.8$, $p>0.05$].

There was no significant main effect of absolute threshold [$F(1,34)=1.1$, $p>0.05$], interaction between the effects of absolute threshold and frequency region [$F(1,34)=1.4$, $p>0.05$], interaction between the effects of absolute threshold and montage [$F(1,34)=0.7$, $p>0.05$], or between the effects of absolute threshold, frequency region and montage [$F(1,34)=1.2$, $p>0.05$]. Therefore, no correlational analyses or plots of TFS FFR as a function of absolute threshold are given.

Observed power for the correlations between TFS-FFR phase coherence and age was 0.997, based on the mean coefficient of determination (r^2) of the correlations as the effect size, 37 listeners and an α of 0.05. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

5.4.3 **Relations between FFR strength and IPD discrimination**

For comparisons between envelope-IPD thresholds and envelope FFR the vertical montage was used as phase coherence was stronger than in the horizontal montage, and thus a better representation of the FFR above the noise floor.

EFFR16 and EFFR115 had lower f_c's (307 and 357 Hz respectively) than the f_c's for EFFR27 and EFFR145 (537 and 578 Hz respectively). Therefore, EFFR16 and EFFR115 were compared to ENV250 thresholds and EFFR27 and EFFR145 were compared to ENV500 thresholds. Correlations between envelope-FFR phase coherence and envelope-IPD discrimination thresholds are given in Table 5.4. EFFR16 and EFFR115 did not correlate significantly with ENV250. EFFR27 did not correlate significantly with ENV500, but there was a weak, negative correlation between EFFR145 and ENV500. This correlation is shown in the left panel of Figure 5.8. However, this correlation was not significant after correcting α for multiple comparisons or after partialing out age ($r=-0.05$, $p>0.05$). The effect of age can be seen in the left panel of Figure 5.8 as the progression from older listeners (pink) mostly at the top of the graph (poor IPD discrimination) and younger listeners (blue) mostly at the bottom (good IPD discrimination).

Table 5.4 Pearson’s *r* correlation coefficients for correlations between envelope-FFR phase coherence (columns) and envelope-IPD discrimination thresholds (rows). The *em* dashes indicate variable pairs not tested for correlation. Asterisk denotes a significant correlation for an α of 0.05.

	EFFR16	EFFR27	EFFR115	EFFR145
ENV250	0.02	—	−0.03	—
ENV500	—	0.15	—	−0.34*

TFS500 thresholds were compared to $\text{TFSFFR}_{\text{HIGH}}$; in the vertical montage there was a trend for a weak negative correlation, but this was not significant (r=−0.30, p>0.05). In the horizontal montage there was a moderate negative correlation between TFS500 and $\text{TFSFFR}_{\text{HIGH}}$ (r=−.40, p<0.05; see right panel of Figure 5.8), but this was abolished when age was partialled out (r=−0.11, p>0.05). TFS250 thresholds were not correlated with $\text{TFSFFR}_{\text{LOW}}$ (vertical montage: r=−0.15, p>0.05; horizontal montage: r=−0.11, p>0.05).

To quantify the extent to which phase locking to the envelope and TFS of AM tones accounted for the age-related deficits in envelope- and TFS-IPD discrimination, partial correlations between age and IPD threshold were calculated with phase coherence partialled out. In Figure 5.9 the ENV500 and TFS500 thresholds are again plotted as a function of age (as in Figure 5.4), but this time with each listener’s phase coherence (EFFR145 from the vertical montage in the left panel and $\text{TFSFFR}_{\text{HIGH}}$ from the horizontal montage in the right panel) indicated by the colour of the datum marker. The correlation reported earlier between age and ENV500 (r=0.53), was still significant after vertical-montage EFFR145 was partialled out (r=0.44, p<0.01). The correlation between age and TFS500 (r=0.46) was weak, but still significant, after vertical-montage $\text{TFSFFR}_{\text{HIGH}}$ was partialled out (r=0.36, p<0.05). However, when horizontal-montage TFSFFRHIGH was partialled out, the correlation between age and TFS500 was not significant (r=0.26, p>0.05).

Observed power for the correlations between FFR phase coherence and IPD discrimination thresholds was 0.38, based on the mean coefficient of determination (r^2) of the correlations as the effect size, 37 listeners and an α of 0.05. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

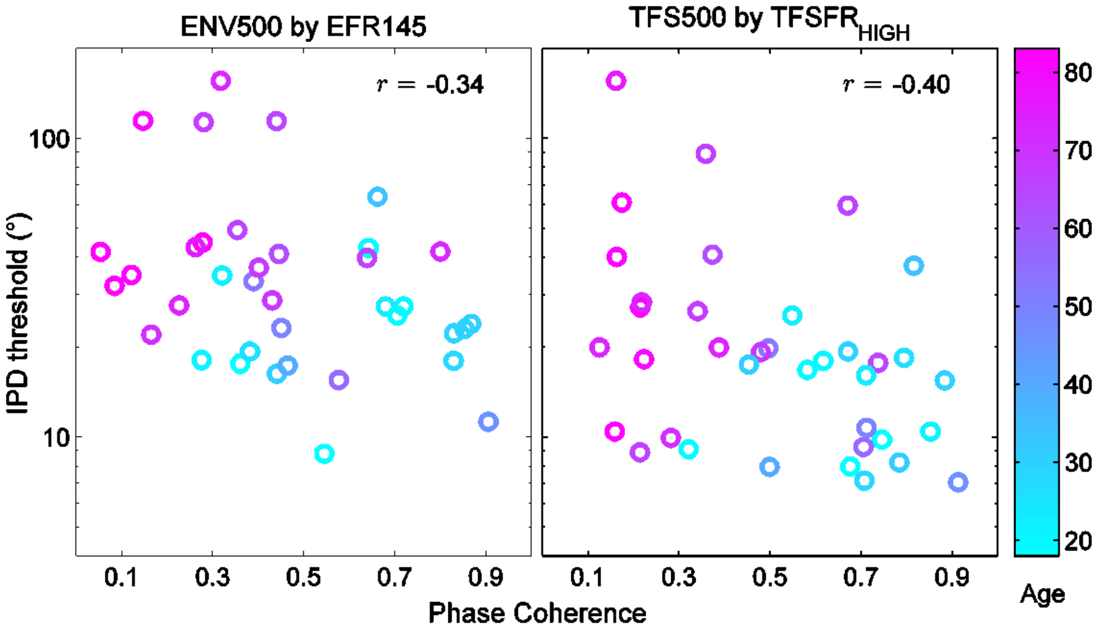

Figure 5.8: ENV500 (right panel) and TFS500 (left panel) thresholds as a function of EFFR145 from the vertical montage and TFSFFR$_{HIGH}$ from the horizontal montage respectively. Pearson's r is inset in each panel. Listener age is depicted by the colour of the symbol, from young (blue) to old (pink).

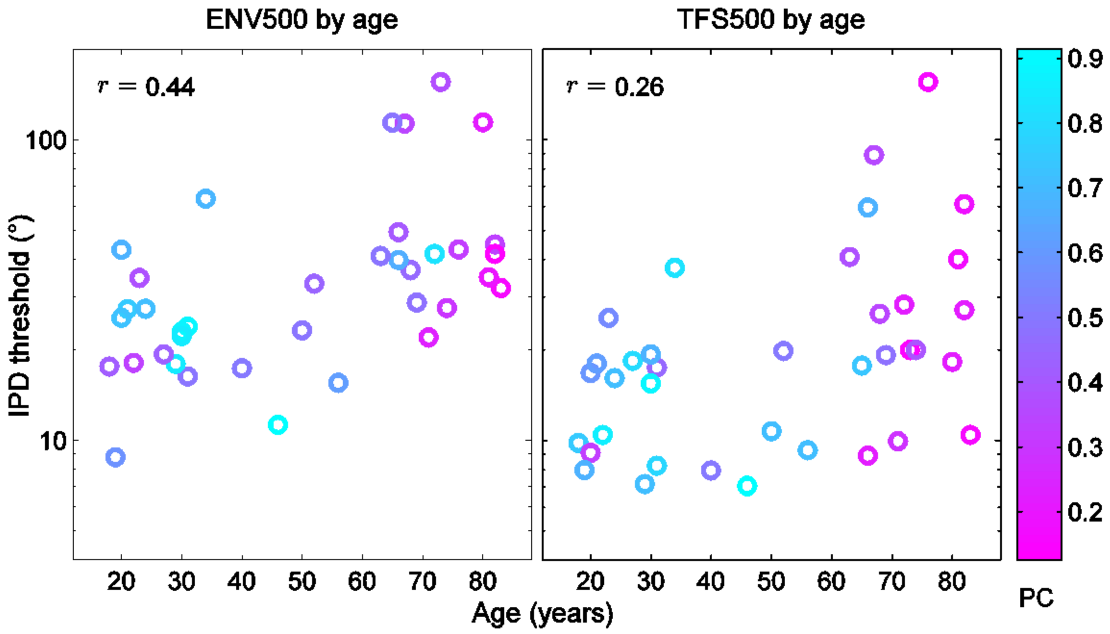

Figure 5.9: ENV500 (right panel) and TFS500 (left panel) thresholds as a function of age (as in the top two panels of Figure 5.4), but with EFFR145 from the vertical montage (right panel) and TFSFFR$_{HIGH}$ from the horizontal montage (left panel) depicted by the colour scale, from poor (pink) to good (blue). Pearson's r between age and IPD threshold, with FFR phase coherence (PC) partialled out, is inset in each panel.

5.5 Discussion

5.5.1 Hearing-loss-related effects

Neither envelope-IPD, nor TFS-IPD, thresholds correlated with absolute threshold. Previous reports of absolute threshold and TFS-IPD thresholds are consistent with this finding (Strelcyk and Dau, 2009; Hopkins and Moore, 2011; Moore *et al.*, 2012a). However, King *et al*. (2014) found a significant correlation between TFS-IPD thresholds and absolute thresholds, with mostly overlapping data to those reported here. Despite a large proportion of the listeners reported by King *et al* (2014) also taking part in the current study, the sample in the current study had fewer listeners with moderate-to-severe hearing loss. Strelcyk and Dau (2009) and Moore *et al*. (2012a) also had a majority of listeners with normal absolute or audiometric thresholds below 1 kHz. Strelcyk and Dau (2009) noted that, with a limited number of hearing-impaired listeners with homogeneous audiograms, absence of a significant correlation does not necessarily imply an absence of a relation. It is possible that there was not enough variation in hearing loss in the current sample to reveal an effect of absolute threshold on IPD thresholds or FFR phase coherence. However, Hopkins and Moore (2011) included a large number of listeners with elevated audiometric thresholds even at low frequencies, but only found a weak correlation between TFS-IPD threshold and absolute threshold at 750 Hz, and no correlation at 250 or 500 Hz.

Although TFS FFR was slightly weaker with increased absolute threshold at high component frequencies, there was no significant effect of absolute threshold on either TFS FFR or envelope FFR. Marmel *et al*. (2013) also did not find a significant relation between absolute threshold and FFR strength. Instead, they found that FFR and absolute threshold contributed independently to a linear model of frequency discrimination performance, with FFR strength accounting for most of the effect of age.

5.5.2 Age-related effects

5.5.2.1 IPD results

IPD discrimination thresholds increased (worsened) with increasing age. This may underlie a decrease in ability to use IPDs to localize and laterally separate sounds in old age. There was wider variation in TFS-IPD thresholds for listeners above 60 years than those below. Identifying causes of the spread in TFS IPD sensitivity in old age may be useful for understanding why some old people struggle to localize sounds and comprehend speech in noisy environments.

Previous studies have shown age-related increases in TFS-IPD thresholds (Grose and Mamo, 2010; Hopkins and Moore, 2011; Moore *et al.*, 2012a; Moore *et al.*, 2012b; Füllgrabe, 2013), but the current authors are unaware of any previous studies of age-related changes in envelope-IPD thresholds. Changes in envelope processing with age have been studied monaurally with temporal modulation transfer functions (TMTFs), which describe changes in modulation detection thresholds across modulation rates (Purcell *et al.*, 2004; He *et al.*, 2008). With broadband Gaussian noise carriers, old listeners have worse (elevated) modulation detection thresholds than younger listeners at modulation rates above 90 Hz, but not at lower rates (Purcell *et al.*, 2004). He *et al*. (2008) found age-related modulation detection deficits with modulation rates from 20 to 200 Hz and a 4 kHz pure-tone carrier (implicating envelope processing deficits). With a 500 Hz pure-tone carrier, deficits occurred at high modulation rates (i.e., when the partials were resolved, implicating TFS processing deficits; He *et al.*, 2008). The current results suggest that age-related deficits in *binaural* processing of modulation are not limited to high modulation rates, even with low-frequency carriers.

5.5.2.2 FFR results

FFRs from the subtraction waveforms, which the authors assume reflect phase locking to the TFS of the AM tones, were generally weaker in $\text{TFSFFR}_{\text{LOW}}$ than in $\text{TFSFFR}_{\text{HIGH}}$. In both regions the average phase coherence at the tone frequencies declined with increasing listener age; the correlations between phase coherence and age were similar. This age-related decline is similar to that found by Marmel *et al.* (2013), but also suggests that age-related declines in FFR phase coherence can occur at lower frequencies than previously reported by Clinard *et al*. (2010). This could be due to calculating FFR phase coherence over 10000 responses in the current study, whereas Clinard *et al.* (2010) only used 1000 responses. This larger recording number may have revealed age effects that were obscured in the data of Clinard *et al*. (2010).

As no exclusion criteria was used to limit absolute thresholds to normal-hearing thresholds, it is possible that slight differences in absolute threshold between young and old listeners confounded the results. However, low-frequency absolute thresholds

(250 and 500 Hz average) and age were not significantly correlated and the negative correlations between age and FFR phase coherence remained when absolute thresholds were partialled out.

Poorer envelope FFR with increasing listener age was seen in EFFR145, but not in EFFR16—an f_m similar to the 20 Hz rate used to test envelope-IPD discrimination. This is consistent with previous envelope-FFR results suggesting that aging affects phase locking predominantly at higher modulation frequencies (Purcell *et al.*, 2004; Leigh-Paffenroth and Fowler, 2006; Grose *et al.*, 2009a; Parthasarathy and Bartlett, 2012; Parthasarathy *et al.*, 2014). This common finding matches the changes in TMTFs with age discussed above (Purcell *et al.*, 2004; He *et al.*, 2008). Purcell *et al.* (2004) found a strong correlation between the maximum frequency at which envelope-FFR was measurable and at which 25% modulation could be detected behaviourally. This suggests that perception of modulation may be limited by a maximum frequency limit of envelope phase locking that reduces with age. Whether or not IPD discrimination is also limited by envelope phase locking is discussed below.

5.5.2.3 Peripheral vs central age-related changes

The current results did not find an interaction between age and montage for the TFS FFR. This suggests that either:

a) Vertical montages and horizontal montages did not selectively measure age-related changes to central (possibly brainstem) and auditory nerve phase locking respectively, or

b) Ageing does not differently affect phase locking in the brainstem and auditory nerve.

Envelope FFR was generally better represented in the vertical montage than in the horizontal montage. Possibly as a result of this, age-related changes in envelope-FFR phase coherence showed up better in the vertical montage also. Previous research has suggested that horizontal montages do not pick up envelope-FFRs well (Galbraith, 1994).

5.5.3 Relation between aging, the FFR, and IPD discrimination

Envelope-IPD discrimination was not well related to envelope-FFRs to the lower modulation rates. Age-related changes in envelope-IPD thresholds and EFFR16 were in opposing directions. Thresholds elevated (worsened) with age, but phase coherence strengthened (improved) with age. EFFR145 was related to envelope-IPD thresholds, but this weak relation could be accounted for by the deleterious effect of age on both EFFR145 and envelope-IPD thresholds. However, when EFFR145 was partialled out, increasing age was still related to worse envelope-IPD thresholds. A similar pattern was seen between TFS-IPD thresholds and TFS-FFR in the horizontal montage. Only in $\text{TFSFFR}_{\text{HIGH}}$ did FFR relate to TFS-IPD thresholds. Again, although age accounted for this relation, TFS-FFR could not account for the effect of age on TFS-IPD thresholds. This suggests that either phase locking to the envelope or the TFS of AM tones at each ear may not be a major determinant of IPD discrimination thresholds for binaural AM tones, or that the measurement of phase locking via the FFR is influenced by other variables, such as variation in FFR source or number of sources. The bottom-up input from the cochlea may be necessary, but other factors, such as binaural integration of the signals from each ear, may have more influence on IPD discrimination thresholds, as suggested by Moore *et al*. (2012b). Strelcyk and Dau (2009) took the strong correlation between frequency modulation detection and IPD-based lateralisation as evidence that lateralisation is also limited by peripheral TFS sensitivity. However, the current results do not provide support for this argument.

IPD discrimination may not be strongly related to the phase coherence of FFR in the current study because the AM-tone FFR stimuli contained no interaural information at any given f_c or f_m. Other studies have considered changes in FFR when interaural time or level differences are included in the stimuli (Clark *et al.*, 1997; Krishnan and McDaniel, 1998; Ballachanda and Moushegian, 2000), but these studies did not compare FFR to behavioural measures such as IPD discrimination. FFR amplitude to tone bursts is greatest with diotic stimuli, and decreases with increasing IPD (Clark *et al.*, 1997; Ballachanda and Moushegian, 2000). The interaction between neural outputs from both ears converging at the brainstem level may be observed in the binaural interaction component (BIC; Krishnan, 2007). The BIC is the residual potential after subtracting the sum of left- and right-ear monaural auditory brainstem responses (ABR) from the binaural ABR. With a click stimulus, the BIC systematically changes with IPDs and ILDs, but requires a fused audio percept (Furst *et al.*, 1985). Krishnan and McDaniel (1998) showed that 500 Hz FFR BIC decreased with increasing ILD, disappearing at 20 dB ILD, but Krishnan and McDaniel did not test the effects of IPDs in the BIC. Ballachanda and Moushegian (2000) found that the FFR BIC was not affected by IPD. Whilst the BIC is evidence that binaural processing is more than simply a linear sum of monaural inputs, it is not clear what binaural processing it actually involves. Further study may look at the relation between IPD sensitivity and FFRs to AM tones with binaural information that reveals the binaural integration of phase-locked responses.

5.6 Conclusions

Both FFR phase coherence and IPD discrimination deteriorate with age, but the age-related deterioration in IPD sensitivity is not strongly related to neural synchrony as measured by the FFR. It is possible that the FFR recorded in the current study reflects phase-locking in the auditory system that is independent of the binaural processing of phase-locked signals that encodes IPDs. Age-related deterioration in IPD sensitivity may be due to other processing changes, such as binaural integration.

Chapter 6. Old hearing-impaired listeners use temporal fine structure in spatial release from masking

Andrew King, Kathryn Hopkins and Christopher J. Plack

School of Psychological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom

Niels Henrik Pontoppidan, Lars Bramsløw, Renskje K. Hietkamp, Marianna Vatti, Atefeh Hafez,

Eriksholm Research Centre, Oticon A/S, Rørtangvej 20, Snekkersten, Denmark

Portions of this chapter were presented at the 4th International Symposium on Auditory and Audiological Research, Nyborg, Denmark, August 2013. This chapter was prepared for publication as an original research article in the Journal of the Acoustical Society of America.

6.1 Abstract

Old, hearing-impaired listeners benefit little from lateral separation of multiple talkers when listening to one of them. Sensitivity to temporal fine structure (TFS) may mediate spatial hearing. This study aimed to determine how spatial release from masking (SRM) in such listeners is affected when the interaural time differences (ITDs) in the TFS are removed by tone-vocoding (TVC) at the ears by a master hearing aid system. Word recall was compared, with and without TVC, when target and maskers sentences were played simultaneously from the front loudspeaker (co-located) and when the maskers were played 45° to the left and right of the listener (separated). SRM was significantly smaller with TVC than without TVC. Old, hearing-impaired listeners, on average, gained a 3.7 dB benefit from TFS-ITDs to their SRM. SRM correlated with monaural TFS sensitivity (discrimination of frequency-shifts in identically filtered complexes), but not when mean audiometric threshold below 1.5 kHz was partialled out. SRM correlated with audiometric thresholds below 1.5 kHz when TFS sensitivity was partialled out, suggesting that low-frequency audiometric thresholds may be a good indicator of candidacy for hearing aids that preserve ITDs. SRM was not correlated with age, pure-tone ITD thresholds, or fundamental frequency difference limens.

6.2 Introduction

Old, hearing-impaired (HI) people struggle to understand speech when they are in a noisy environment (Duquesnoy, 1983; Gatehouse and Noble, 2004; Helfer and Freyman, 2008). This may be due, in part, to a degraded ability to separate sounds that come from different directions. Interaural time differences (ITDs) and interaural level differences (ILDs) occur when a sound source is located outside the median plane (Rayleigh, 1907; Kuhn, 1977). The wave front reaches one ear before the other ear (causing an ITD), and is also acoustically shadowed by the head for wavelengths below the head size (causing an ILD). It is possible to use these differences to separate sounds that arrive from different directions (Middlebrooks and Green, 1991; Best *et al.*, 2004; Bremen and Middlebrooks, 2013). In noisy environments, speech intelligibility can be improved dramatically by presenting the masker sounds to the left and right of the target speech, rather than all from the same location (Bronkhorst, 2000). When the maskers are speech signals, the spatial release from masking (SRM) is typically 10 to 15 dB for normal-hearing (NH) listeners (Behrens *et al.*, 2008; Marrone *et al.*, 2008b). HI people (particularly if they are elderly) vary widely in the SRM they gain, from close-to-normal to none at all (Marrone *et al.*, 2008a), even with compensation for reduced audibility (Marrone *et al.*, 2008c; Neher *et al.*, 2009). This suggests that supra-threshold deficits in spatial separation of sounds may occur with some forms of hearing impairment. A better understanding of why some people benefit from spatial cues, and others do not, would help development and prescription of hearing aids that preserve or sacrifice spatial cues depending on a listener's ability to use these cues.

Precise temporal coding is needed to use ITDs to separate sounds arriving from different angles, as the maximum ITDs for a human head are less than 1 ms (Kuhn, 1977). Auditory nerve fibres synchronize their firing to the phase of basilar membrane displacement in each ear (Rose *et al.*, 1967), from which the brain can determine the interaural difference. By coding time intervals between corresponding peaks in the band-pass filtered output from the basilar membrane, phase locking can also represent the temporal fine structure (TFS) of sounds that would not be available from the place of excitation on the basilar membrane alone (Cariani and Delgutte, 1996). TFS may benefit pitch perception, lateralisation and speech perception in noise (Moore, 2008).

The benefit of TFS to speech perception can be demonstrated by comparing speech intelligibility with the TFS intact and with the TFS altered by a process such as vocoding (Dudley, 1939). Vocoding involves filtering the original waveform into a series of band-limited frequency channels. The envelopes are extracted from each band and then multiplied by synthesized carrier signals. In the case of tone-vocoding (TVC), the carrier is a sinusoid at the centre frequency of the pass-band. Finally, the synthesized channels are summed. In quiet, vocoded speech can be understood even with the use of a few frequency channels (Shannon *et al.*, 1995; Loizou *et al.*, 1999; Baskent, 2006; Lorenzi *et al.*, 2006). However, in the presence of noise, vocoded speech requires higher signal-to-noise ratios than non-vocoded speech to be intelligible (Qin and Oxenham, 2003; Stone and Moore, 2003).

It is worth noting that, for band-limited signals, vocoding may not manipulate the TFS completely independently from the envelope. Modelled neural representations of the TFS and envelope are comparably degraded by vocoding with noise carriers (Shamma and Lorenzi, 2013). Therefore, it may be incorrect to assume that envelope information is faithfully preserved, after vocoding, at a neural level. Instead, it may be more appropriate to think of vocoding as disrupting the neural representation of temporal information generally. Shamma and Lorenzi (2013) demonstrate the need to use caution when inferring distinct contributions of TFS and envelope to hearing abilities from manipulations of stimuli via vocoding. However, a narrow-band noise has more inherent fluctuations than a tone, so noise carriers may affect the envelope more than tone carriers (Kates, 2011).

It is possible that ITDs carried in the TFS are used to improve separation and intelligibility of speech streams. Presuming that vocoding primarily affects TFS-ITDs, this can be tested by comparing the amount of SRM achieved with preserved TFS to the SRM achieved with the TFS altered by TVC. Andersen *et al.* (2010) did this over headphones using head-related impulse responses (HRIRs) from a head and torso simulator (Algazi *et al.*, 2001) to simulate lateral separation of speech signals, and TVC to alter the TFS. Andersen *et al.* (2010) showed that young, NH listeners’ speech reception thresholds (SRTs) were lowest (best) when no TVC was applied; SRTs were elevated by 5.9 dB (a deficit) when the TVC was applied to the signals before the HRIRs (preserving the ITDs but not the original TFS). This shows that vocoding decreases performance even if spatial cues are preserved. Importantly, however, SRTs elevated a further 2.4 dB when the TVC (which was in phase between the ears) was applied after the HRIRs, effectively removing the ITDs. This difference in performance between vocoding before and after the application of spatial cues suggests that temporal binaural cues, even carried by TVC signals, help speech intelligibility.

Using model HRIRs (i.e., from a head and torso simulator, rather than a listener's own HRIRs) provides a reasonable facsimile of auditory space for NH listeners, but this is not how hearing aid users perceive auditory space. Firstly, model HRIRs commonly include pinna cues, whilst behind-the-ear hearing aids remove pinna cues. Secondly, using headphones and HRIRs results in the auditory scene artificially moving with the listener's head, whereas using hearing aids and a loudspeaker setup would allow head movements without moving the auditory scene. This may lead to SRM that is more ecologically valid for hearing aid users, but it is unclear whether this will improve SRM or not.

Hearing loss appears to reduce the ability to use TFS, as demonstrated by poorer performance at low-rate frequency modulation (FM) detection and discrimination (Lacher-Fougère and Demany, 1998; Moore and Skrodzka, 2002; Buss *et al.*, 2004), discrimination of frequency shifts in harmonic complex tones (Moore *et al.*, 2006b; Hopkins and Moore, 2007), pure-tone ITD discrimination and ITD-based lateralisation (Hawkins and Wightman, 1980; Smoski and Trahiotis, 1986; Lacher-Fougère and Demany, 2005; Strelcyk and Dau, 2009; King *et al.*, 2014), and speech perception (Lorenzi *et al.*, 2006; Hopkins *et al.*, 2008; Lorenzi *et al.*, 2009; Hopkins and Moore, 2010c; 2011). For example, Hopkins *et al.* (2008) studied the effect of vocoding speech on intelligibility in the presence of a continuous competing talker. They compared the performance of NH listeners with HI listeners. NH listeners performed much worse with TVC speech than with the original speech. HI listeners however, regardless of the speech processing, performed about the same as the NH listeners did with the TVC speech. This suggested that HI listeners do not benefit from the TFS in the original speech as much as NH listeners.

Increasing age is also associated with poorer TFS sensitivity, even in the absence of substantial hearing loss below 2 kHz (Ross *et al.*, 2007; Grose and Mamo, 2010; Hopkins and Moore, 2011; Moore *et al.*, 2012b), or with hearing loss statistically controlled for (Hopkins and Moore, 2011; Marmel *et al.*, 2013; King *et al.*, 2014). Age also reduces the advantage of binaural hearing (Warren *et al.*, 1978; Pichora-

Fuller and Schneider, 1992). This may underlie poor speech perception in noisy environments (Working Group on Speech Understanding and Aging, 1988; Divenyi *et al.*, 2005; Helfer and Freyman, 2008) and less SRM (Gelfand *et al.*, 1988; Marrone *et al.*, 2008a; Marrone *et al.*, 2008c). It is possible that age exacerbates the effect of hearing loss on the ability to use TFS to benefit speech perception.

The first aim of the current study was to determine whether or not older, HI listeners are able to use TFS to achieve SRM. This may help determine whether or not older HI listeners benefit more from hearing devices that preserve TFS ITDs or from devices that sacrifice these cues in favour of increasing the target-to-masker ratio (TMR) through existing or future processing strategies (such as directional microphone sensitivity; Van de Bogaert *et al.*, 2005). If TFS is used to separate the speech stimuli, then speech understanding should be diminished by TVC if it disrupts the TFS (and possibly other temporal cues). An interaction between SRM and the effect of TVC would suggest the ITDs in the TFS were being used to separate the masker speech from the target speech. The second aim was to identify a good measure to predict who can benefit from TFS in separating speech (see section 6.4); a measure that might be applicable in clinics for hearing aid prescription and fitting.

Previous studies have compared individual differences in speech intelligibility to psychoacoustic discrimination or detection performance that reflects basic auditory acuity such as temporal resolution or frequency selectivity (e.g., Tyler *et al.*, 1982; Schneider *et al.*, 1994; Strouse *et al.*, 1998; Strelcyk and Dau, 2009; Hopkins and Moore, 2011). For example, Hopkins and Moore (2011) measured SRTs for short sentences in notionally steady and modulated speech-shaped noise and compared them with psychoacoustic measures of frequency selectivity (threshold tone detection in notched-noise) and TFS sensitivity.

Hopkins and Moore (2011) measured TFS sensitivity binaurally with pure-tone ITD discrimination thresholds. They also measured TFS monaurally via discrimination of harmonic from inharmonic complexes filtered with identical pass-bands (the TFS1 task). As the components in the inharmonic complex are shifted in frequency equally, the modulation rate (envelope) is the same for both harmonic and inharmonic complexes. This task is thought to rely exclusively on TFS sensitivity if the components of the complex within the pass-band are of sufficiently high harmonic number to remain unresolved by the cochlea. Component magnitudes in both harmonic and inharmonic complexes are shaped by the same band-pass filter (Moore *et al.*, 2006a). However, small excitation pattern differences might still allow discrimination (Micheyl *et al.*, 2010).

Pitch discrimination has been used previously as a predictor of temporal coding acuity even when excitation pattern differences exist; for example, frequency discrimination of pure tones (e.g., Turner *et al.*, 1983; Freyman and Nelson, 1991). Performance is thought to reflect the advantage of coding frequency via phase locking (Rose *et al.*, 1967) when it is better than that expected from the frequency tuning of the basilar membrane (Moore, 1973a; b; Srulovicz and Goldstein, 1983; Sek and Moore, 1995; Micheyl *et al.*, 1998).

Strouse *et al.* (1998), Strelcyk and Dau (2009), and Neher *et al.* (2011; 2012) considered the relation between binaural advantages to SRTs and ITD sensitivity. Strouse *et al.* (1998) found no correlation between ITD sensitivity and the binaural masking level difference (BMLD) in speech for young and old NH listeners. The BMLD is the difference between the SRT with a diotic target and the SRT with a target in anti-phase across ears, when the masker is diotic in both cases. Strouse *et al.* (1998) used click trains for ITD discrimination, which do not test TFS-ITD sensitivity exclusively. Detection of ITDs in pure tones (with synchronous onsets and offsets) does require TFS sensitivity however. Strelcyk and Dau (2009) measured ITD-based lateralisation of tones in noise, and BMLDs for tones in noise, and found both to be correlated with SRTs in laterally separated, speech-shaped noise.

Lower SRTs in laterally separated speech maskers (similar to the Separated condition without TVC in the current study) have been correlated with a higher maximum frequency of a tone in which an ITD could be discriminated (Neher *et al.*, 2011) and a lower ITD discrimination threshold at a fixed frequency (Neher *et al.*, 2012). These results suggest that sensitivity to TFS ITDs might contribute to an individual’s ability to isolate and understand one talker in a background of speech from other azimuths. The current study aimed to determine if pure-tone ITD discrimination predicted not just SRM, but specifically the effect of TVC on SRM.

6.3 Speech test

6.3.1 **Listeners**

Twenty listeners ranged from 64 to 86 years old (mean=72 years old) and all had bilateral, gently-sloping sensorineural hearing loss (see Figure 6.1). Mean audiometric pure-tone thresholds across listeners and frequencies (0.125, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, and 8 kHz) were 47 dB HL for the right ears and 46 dB HL for the left ears. Listeners mean audiometric thresholds from 0.125 to 1.5 kHz (PTA_{LF}) were not correlated with their age (r=0.01). Listeners mean audiometric thresholds from 2 to 8 kHz were also not correlated with their age (r=0.19). Listeners were screened for conductive or mixed hearing losses. Asymmetry across ears was < 10 dB for the frequency-average (125 to 8000 Hz) audiogram. Three listeners had a 20 dB asymmetry at a single frequency. All listeners spoke Danish as their first language.

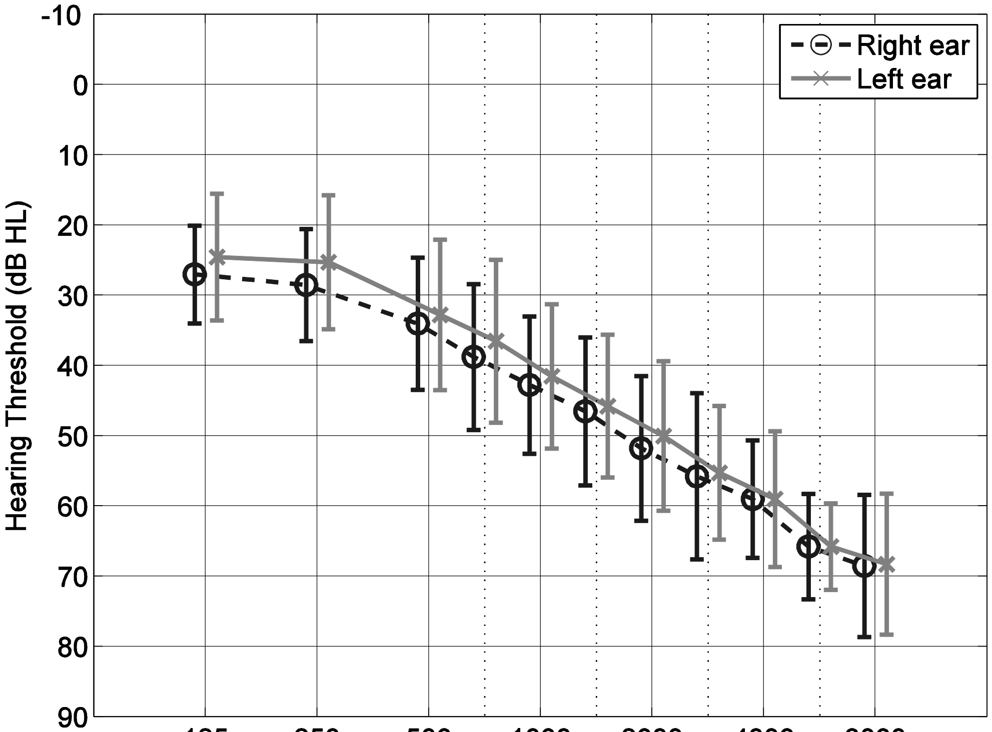

Figure 6.1: The mean audiograms of the 20 listeners (± 1 Standard Deviation) for left (grey crosses) and right (black circles) ears separately.

Based on the effect size of vocoding on speech-in-noise perception for HI listeners from Hopkins *et al.* (2008) and the effect size of SRM for HI listeners from Neher *et* *al.* (2009), 16 listeners were needed to achieve statistical power of 0.8 and 20 listeners provided a statistical power of 0.9.

6.3.2 **Speech stimuli**

A special version of the Danish Dantale II corpus (Wagener *et al.*, 2003) designed for spatial speech-on-speech testing was used (Behrens *et al.*, 2008). Recorded words were spoken by three Danish females. Words were selected from a closed set. Sentences were five words long and always followed the same structure and order: a person's name, verb, number, adjective, and object. For example, "*Henning købte tre smukke ringe"* ("Henning bought three beautiful rings"). The listener was scored on how many words in the target sentence they correctly recalled, in each trial. The target sentence was always played from the front loudspeaker (0° relative to the listener) in an anechoic room. Two masker sentences (M_1 and M_2) were played at the same time as the target, at various TMRs. The three sentences—each spoken by a different female talker—began with a "Ready" prompt. The first word (the person's name) in the target sentence was displayed to the listener via a computer screen hanging above the front loudspeaker to cue which sentence to listen to and recall. Two spatial configurations were used (see Figure 6.2). For the Co-located configuration, M_1 and M_2 were played from the same front loudspeaker as the target sentence. For the Separated configuration, M_1 and M_2 were played from loudspeakers −45° and +45° azimuth relative to the listener, respectively. Symmetrically separated maskers were used to minimize the benefits of increased TMR at one ear that occur with asymmetrical maskers (Marrone *et al.*, 2008a). The maximum root mean square (RMS) sound pressure level at the centre of the listener's head position was 70 dB sound pressure level (SPL). For positive TMRs, M_1 and M_2 were attenuated whilst the target level remained fixed. For negative TMRs, the target was attenuated whilst M_1 and M_2 remained fixed in level.

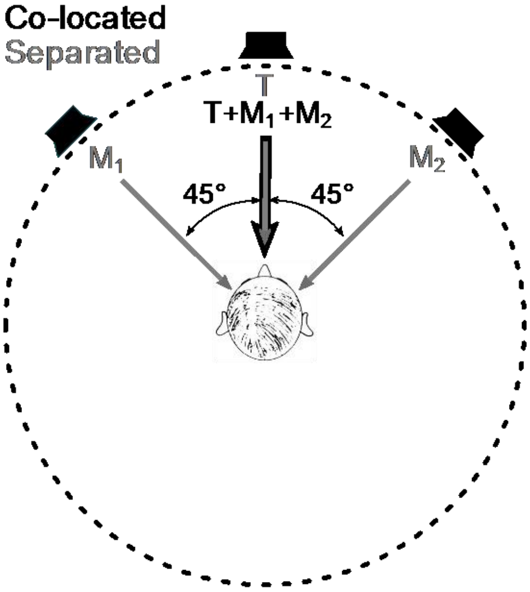

Figure 6.2: A plan of the loudspeaker positions around the listener and which loudspeakers played the target speech-signals (T) and the two masker speech-signals (M1 and M2) in the Co-located (black) and Separated (grey) configurations.

6.3.3 Master Hearing Aid signal processing

Listeners wore behind-the-ear hearing aid microphones which recorded the sounds at the listener's ears. The signals were sent to a control computer running the Master Hearing Aid system (MHA; HörTech, 2008; described in Grimm *et al.*, 2006) which split the signals into 512 linear sub-bands using a fast Fourier transform. These sub-bands were summed into 32 logarithmically spaced sub-bands the equivalent rectangular bandwidths of NH listeners’ auditory bandwidths at moderate sound levels (ERBN, Glasberg and Moore, 1990). The 32 non-linear sub-bands had a combined pass-band from 100 to 10,000 Hz. Since the complex value in each sub-band had an imaginary part that was a 90° phase-shifted version of the real part of any given time sample, the envelopes were extracted as the absolute value of each sub-band complex value (equivalent to the Hilbert envelope). For each sub-band, the cut-off frequency of the extracted envelope was equal to half the sub-band bandwidth. Two MHA conditions were used, one with TVC and one without TVC. In the TVC condition the envelopes were multiplied by a pure tone, in phase across ears, at the band centre frequency, thus removing the ITD in the TFS. The modulated tones were combined and amplified with linear gain following CAMEQ specifications (Moore and Glasberg, 1998) to correct for hearing loss and also for outer-ear gain and the hearing aid receiver (output) frequency response. In the condition without TVC, the processing followed the same procedure, except that each extracted envelope was multiplied by the phase angle of the complex signal of the corresponding sub-band, rather than pure tones, to restore the original TFS. The MHA signal processing (either with, or without, TVC) was performed in almost real time, producing a delay of approximately 40 ms. Figure 6.3 describes the stages of processing performed by the MHA platform. The output was presented to the listener from the hearing aid receiver via highly-damping Etymotic Research foam plugs for each listener.

Figure 6.4 shows the effects of TVC on the ITD of a speech signal from 45° left of a head and torso simulator wearing the MHA devices. The ITD is clear in both the broadband signal and envelope without TVC (top row), but with TVC (bottom row) the ITD (in the interaural cross-correlation) is removed from the broad-band signal whilst the ITD is partially preserved in the envelope. The envelope ITD and the ILD are reduced, but remain to some extent, whereas the TFS-ITD is entirely removed. The very broad peak in the envelope cross-correlation suggests that the envelope ITD may not be a precise cue for lateralising or localising sounds, if the auditory system extracts comparable information to this analysis.

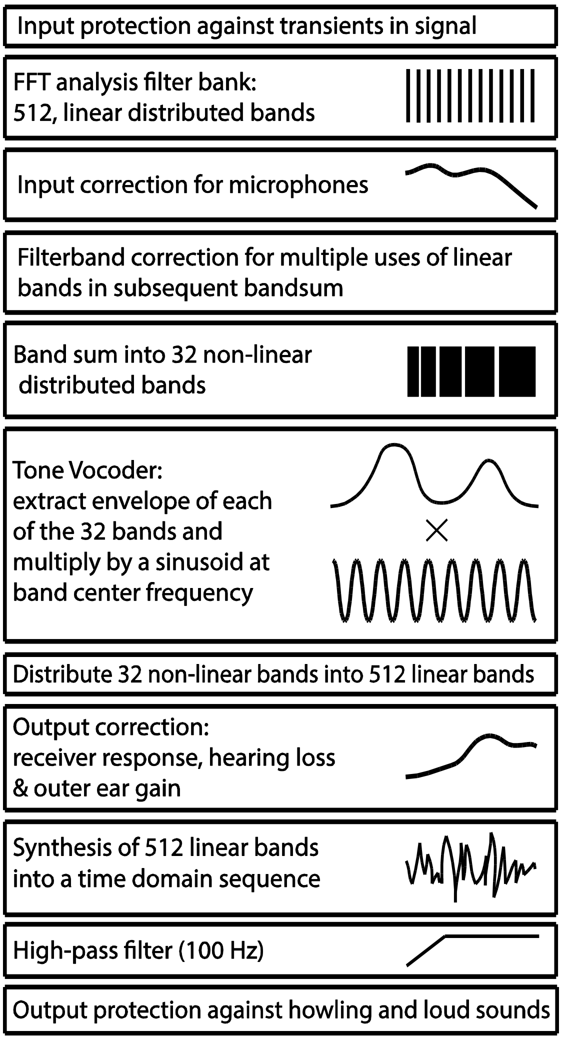

Figure 6.3: A flowchart of the processing stages involved in the Master Hearing Aid system.

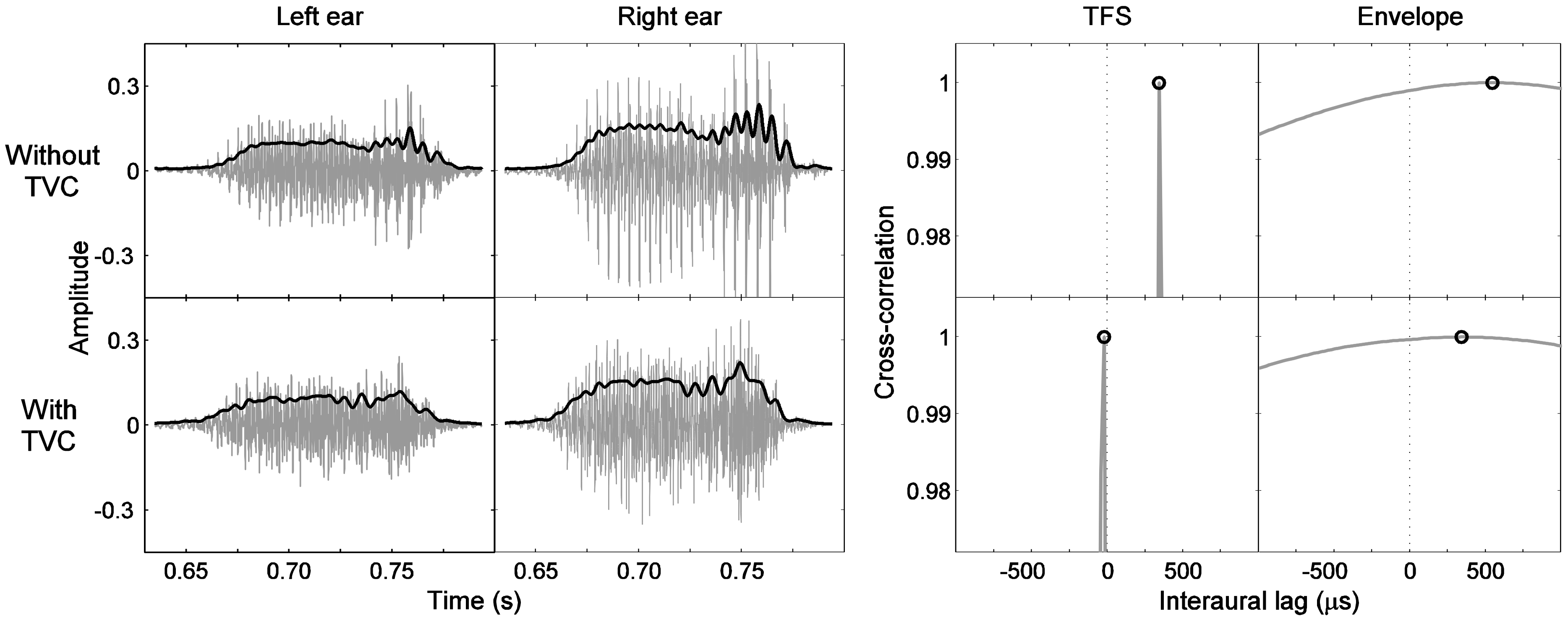

Figure 6.4: Example excerpts of sentences played from 45° to the right of the listener’s position, recorded on a head and torso simulator (B&K). The top row of panels shows the sentence recorded without tone-vocoding and the bottom row of panels show the same sentence recorded with tone-vocoding. The time domain waveform (in grey) of a portion of speech (the word “*valgte*”) and its envelope (half-wave rectified and low pass filtered at 160 Hz; overlaid in black) at the left ear (far-left panels) and the right ear (middle-left panels), the interaural cross-correlation of these two signals (middle-right panels) and of the two signals’ envelopes (far-right panels). Maximum cross-correlations (ITDs) indicated by black circles.

6.3.4 **Training**

To minimize the confounding effect of learning the task or the nature of the stimuli, listeners were familiarized with the stimuli in a training session at least one week prior to the test session. The training session consisted of four pairs of blocks of trials with the first block of each pair (odd-numbered steps) without TVC, and the second (even-numbered steps) with TVC. Blocks one and two each consisted of 12 trials of a target sentence from 0° in quiet. Blocks three and four each consisted of 12 trials of a target sentence from 0° and one masker sentence from either +45° or −45° (TMR=5 dB). The last four blocks each consisted of 12 trials at each of two TMRs (see Table 6.1) and had two maskers. The maskers were in the Separated (±45°) configuration for blocks five and six, and in the Co-located configuration for block seven and eight. Listeners had to recall as many words from the target sentence as possible. Sentences were not repeated.

Table 6.1: The TMRs, the percent correct of word recall the TMR was expected to produce, and the mean (across listeners) percent correct that was observed for the final four training blocks.

	TMR (dB)	Estimated percent correct	Observed percent correct
Separated without TVC	8	70%	74%
	2	40%	41%
Separated with TVC	9	70%	76%
	3	40%	45%
Co-located without TVC	10	80%	87%
	4	40%	49%
Co-located with TVC	11.5	80%	90%
	6	40%	56%

6.3.5 **Procedure**

The test session began with a ‘warm-up’ block of 12 trials in the Separated condition without TVC, then 12 with TVC (TMR=8.7 dB for both). After this, four test blocks of 50 trials were performed with short breaks in between. Test-condition order was pseudo-randomized with spatial configuration nested inside MHA condition, so both spatial configurations were completed for one MHA condition before beginning the second MHA condition. Randomisation resulted in eight participants completing the conditions with TVC before the conditions without TVC, and 12 completing the conditions without TVC before the conditions with TVC. Although a counter-balanced ordering would have been more appropriate, Student's *t*-tests between those who completed conditions with TVC first and those who completed conditions without TVC first found no significant differences in performance in the four test conditions.

For each test condition, a listener's psychometric function was estimated from the proportion of correctly recalled words in the 50 trials using a logistic function of TMR (Green and Swets, 1974):

$$\Psi = \left\{1 + e^{4s_{50}(L_{50} - TMR)}\right\}^{-1} \tag{6.1}$$

where s_{50} and L_{50} denote the slope and TMR (respectively) at 50% correct word recall fitted using a negated logarithmic maximum-likelihood estimation procedure. Using the mean group performance in the training session as L_{50} and a shallow function slope to minimize floor and ceiling effects (s_{50}=0.02), a wide range of TMRs estimated to produce 30, 40, 50, 60, 70, 80, and 90% correct word recall were calculated for the first 24 trials (pre-defined TMRs). The pre-defined TMRs are given in Table 6.2. The discrepancy between the TMRs in the last four blocks of training (Table 6.1) and the pre-defined TMRs in the test sessions (Table 6.2) occurred because of the shallow function slope to calculate the pre-defined TMRs. Three trials were presented at each TMR after an initial three at the TMR expected to produce 80% correct.

Table 6.2: A table of the pre-defined TMRs (dB) used for the first 24 trials of each block by spatial configuration (rows) and by the percent correct word recall the TMR was estimated to produce (columns).

	Estimated percent correct word recall						
Configuration	30%	40%	50%	60%	70%	80%	90%
Co-located, test	−1.3	1.9	5.1	8.3	11.5	14.7	17.9
Separated, test	−7.3	−4.1	−0.9	2.3	5.5	8.7	11.9

Further trials were presented at TMRs that were estimated to be focused around each listener’s psychometric function. To do this, an interim psychometric function was estimated from the listener’s performance on the pre-defined trials, and a further 26 trials were presented at TMRs from the interim function estimating 40, 60, 70, and 90% correct. Six trials were presented at each personalised TMR after two trials at the TMR expected to produce 70% correct. A final psychometric function was calculated from the results of both the pre-defined and personalised TMR trials (excluding the initial three pre-defined and two personalised). Inclusion of responses to both pre-defined and personalised TMR trials allowed as many data points as possible to be used in calculating the final psychometric function. The TMR that would give 50% correct was taken as threshold ($TMR_{50\%}$) for analysis.

SRM was defined as the $TMR_{50\%}$ in the Separated condition subtracted from the $TMR_{50\%}$ in the Co-located condition.

6.3.6 Results and discussion

The $TMR_{50\%}$ mean and individual values are plotted in Figure 6.5. With TVC, $TMR_{50\%}$ was similar in the Co-located (mean=3.7 dB) and Separated configurations (mean=3.4 dB); individual SRM ranged between −1.8 and 2.7 dB. In the Co-located condition, $TMR_{50\%}$ without TVC (mean=2.8 dB) was similar to $TMR_{50\%}$ with TVC. However, $TMR_{50\%}$ in the Separated condition without TVC was lower (mean=−1.2 dB), indicating better performance, but also more varied across listeners (−6.5 to 5.5 dB), leading to SRM ranging from −0.9 to 8.4 dB.

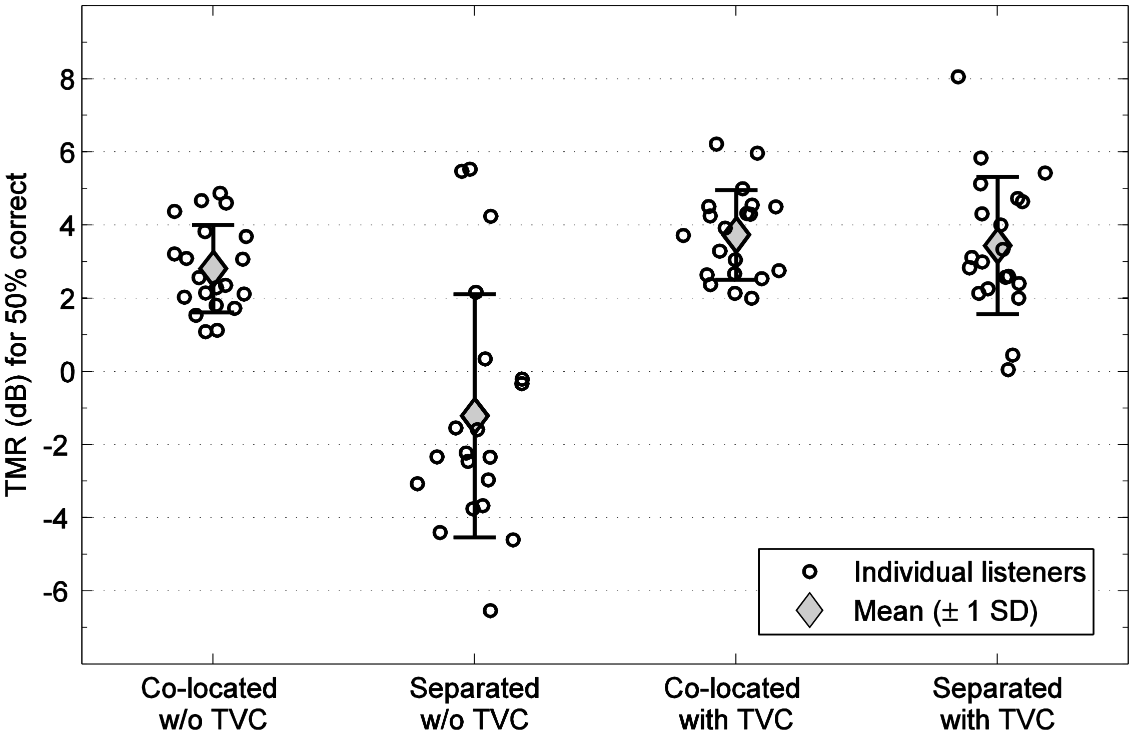

Figure 6.5: Estimated TMRs for 50% correct word recall ($\text{TMR}_{50\%}$), plotted by speech test condition. Black circles indicate $\text{TMR}_{50\%}$ for individuals, grey diamonds indicate mean ± 1 SD.

Performance was analysed with a repeated-measures ANOVA with two within-subjects factors: spatial configuration (Co-located *vs*. Separated) and processing (without TVC *vs*. with TVC). Both factors produced significant main effects. $\text{TMR}_{50\%}$ was lower in the Separated conditions than the Co-located conditions [$F(1,19)=37.8$, $p<0.001$] and $\text{TMR}_{50\%}$ was lower without TVC than with TVC [$F(1,19)=87.0$, $p<0.001$]. The interaction was also significant [$F(1,19)=38.6$, $p<0.001$], confirming that the SRM without TVC was greater than SRM with TVC. This suggests that the limited SRM available to old HI listeners can be disrupted by TVC. Because the TVC removed the ITDs in the TFS in each sub-band by generating the sine carriers in phase across the ears, it appears that old HI listeners gained significantly by using ITDs in the TFS of sounds from different azimuths. However, as TVC may also affect the spectro-temporal envelope (Kates, 2011; Shamma and Lorenzi, 2013), the interaural envelope cues may not have been entirely preserved in the TVC conditions. Therefore, the effect of changes to the envelope cues on SRM cannot be discounted, although contribution of ITDs cues in the envelope to SRM is unknown.

Without TVC, the mean SRM and the variation in SRM was very similar to that found for HI listeners by Marrone *et al.* (2008a) and Neher *et al.* (2009; 2011). Behrens *et al*. (2008), Marrone *et al.* (2008a), and Neher *et al.* (2009) found HI listeners' mean SRM was approximately 8-14 dB less than NH listeners' SRM and HI listeners had larger individual differences in SRM than NH listeners. In the present study, the difference between $TMR_{50\%}$ in the Separated configuration with TVC and without TVC was 4.6 dB, which is roughly half that found for young NH listeners in a simulated spatial setup using headphones and HRIRs (8.3 dB; Andersen *et al.*, 2010). Whilst old HI listeners are less sensitive to TFS-ITDs than young NH listeners (Hawkins and Wightman, 1980; Smoski and Trahiotis, 1986; Lacher-Fougère and Demany, 2005; Ross *et al.*, 2007; Grose and Mamo, 2010; Hopkins and Moore, 2011; King *et al.*, 2014), it should not be assumed that they are unable to use ITDs in the TFS of speech. Differences in the nature and severity of hearing loss may explain the variation in old HI listeners' ability to use TFS-ITDs in SRM.

To determine whether individual differences in the effects and interaction of TVC and spatial configuration were driven by age or hearing loss, each listener's age and PTA_{LF} were used as covariates in an extension of the ANOVA model described above. Higher (worse) PTA_{LF} was related to a higher (worse) $TMR_{50\%}$, across all conditions [F(1,17)=8.8, $p<0.01$], but age was not related to a change in $TMR_{50\%}$ over all conditions [F(1,17)=0.7, $p>0.05$]. Age did not interact with the effects of MHA processing [F(1,17)=0.1, $p>0.05$] or spatial configuration [F(1,17)=1.2, $p>0.05$]. Higher PTA_{LF} interacted was associated with a smaller difference in performance between the vocoded and non-vocoded conditions [F(1,17)=17.3, $p<0.01$]. PTA_{LF} did not make a significant difference to the effect of spatial configuration on $TMR_{50\%}$ [F(1,17)=3.6, $p>0.05$]. PTALF also had an influence on the interaction between MHA processing and spatial configuration [F(1,17)=23.2, $p<0.001$]; the increase in SRM with non-vocoded speech, compared to vocoded speech, was smaller with increasing PTA_{LF}. This is shown in Figure 6.6, where $TMR_{50\%}$ is plotted as a function of PTA_{LF} with least-squares linear fits for each test condition separately. The larger spread of performance variance in the Separated condition without TVC can be partially explained by an increase in $TMR_{50\%}$ with increasing PTA_{LF}. This suggests that low-frequency audiometric hearing loss is related to how well a listener can understand speech from a conversational partner when others are talking around them.

If performance in the Separated condition without TVC is dictated by sensitivity to TFS IPDs in the maskers, then increasing $TMR_{50\%}$ with increasing PTA_{LF} is consistent with evidence of poorer TFS-IPD discrimination with increasing absolute threshold at the carrier frequency (e.g., King *et al.*, 2014). Low-frequency audiometric thresholds may be a convenient way to determine which type of hearing aid processing may be best for an individual with hearing loss. If an individual's PTA_{LF} is low, they may benefit from hearing aids that preserve the binaural cues in the TFS. If their PTA_{LF} is high, they may benefit from other processing strategies that focus the microphones in a specific direction even at the expense of binaural TFS cues.

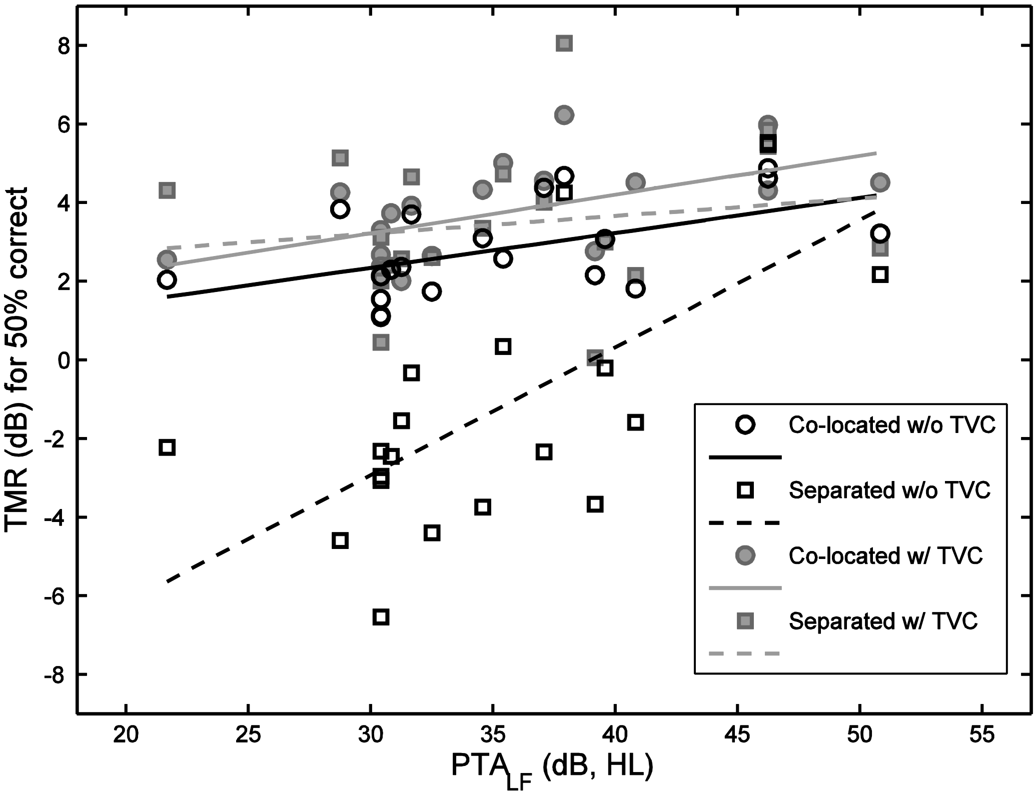

Figure 6.6: $TMR_{50\%}$ plotted for each test condition as a function of the listeners' low-frequency-average audiometric threshold (PTA_{LF}). Lines indicate least-squares best linear fit for each condition.

The similarity between $TMR_{50\%}$ in both spatial configurations with TVC suggests the envelope ITDs and ILDs are not sufficient for speech unmasking, assuming that TVC does not disrupt these cues substantially. This may be because old listeners have poorer sensitivity than younger listeners to ITDs in the envelope of amplitude-modulated tones (King *et al.*, 2014). Comparing performance in the Separated condition with TVC in the current study and in Andersen *et al.* (2010) suggests that

$\mathrm{TMR}_{50\%}$ is about 12 dB lower for young NH listeners than for old HI listeners when only envelope ITDs and ILDs are available for SRM. Envelope ITDs may be a poor cue for separating speech streams that overlap temporally. Monaurally, speech signals are not well unmasked by modulated noise or speech maskers compared to steady noise maskers when only envelope information is available (Lorenzi *et al.*, 2006; Hopkins and Moore, 2009). Old listeners may have to rely almost exclusively on whatever TFS-ITD cues they retain, but when these are disrupted the listeners receive very little benefit from the remaining spatial cues.

There was a small, but significant, difference in $\mathrm{TMR}_{50\%}$ between the Co-located conditions with TVC and without TVC [mean=0.9 dB, $t(19)=4.8$, $p<0.001$]. Better performance without TVC suggests that the listeners were also able to use information that was disrupted in the conditions with TVC to help separate the speech signals, even when no ITD information existed. The effect of TVC in the Co-located conditions was smaller than the monaural effect of TVC on speech intelligibility found by Hopkins *et al.* (2008); Hopkins *et al.* (2008) found that old, HI listeners had SRTs of 7 dB with TVC and 2 dB without TVC. This may possibly be due to differences in the speech corpora. In a similar study to Hopkins *et al.* (2008), Lunner *et al.* (2012) compared speech intelligibility in competing speech with three different speech corpora:

1) The Dantale II closed-set corpus (Wagener *et al.*, 2003),
2) The Danish version of the open-set Hearing In Noise Test (Nielsen and Dau, 2009),
3) The open-set adaptive sentence list (MacLeod and Summerfield, 1990) used by Hopkins *et al.* (2008).

Lunner *et al.* (2012) found that there was a larger effect of TVC in the open-set corpora than the closed-set corpus. However, this was only evident with young NH listeners, and not old HI listeners.

6.4 Discrimination tests

The individual differences in speech recall performance were compared to performance on various psychoacoustic discrimination tasks. It was hypothesized that performance on these tasks would correlate with individual differences in SRM, and the effect of TVC on SRM, for older HI listeners.

Hopkins and Moore (2011) found that, after controlling for audiometric threshold, frequency selectivity was still correlated with SRTs in steady noise and TFS sensitivity was still correlated with SRTs in modulated noise. As the maskers in the current study were not steady noise, and the speech test was designed to reveal the benefits of TFS to SRM, frequency selectivity was not measured and instead the experiment focused on measures of TFS sensitivity.

In addition to the TFS1task and pure-tone ITD discrimination, fundamental frequency difference limens (F0DLs) were measured in the current study. An F0DL is a listener's threshold for discriminating two harmonic complexes with different modulation rates (which corresponds to the frequency spacing of the components). Therefore F0DLs may reflect a listener's ability to use both the envelope and TFS cues to discriminate complexes if the harmonics are unresolved, although the use of TFS cues may be limited (Oxenham *et al.*, 2009). If the harmonics are resolved, envelope cues may be weak. Comparisons of TFS1 scores and F0DLs may reveal the extent to which older HI listeners use envelope cues to discriminate sounds.

Hopkins and Moore (2011) measured SRTs in modulated maskers monaurally. Therefore, SRTs may have been more related to monaural, than to binaural, measures of TFS. SRTs were correlated with both TFS1 and ITD discrimination, but only with TFS1 if audiometric threshold was controlled for. In the current study however, the speech test was assumed to test binaural processing abilities that use interaural cues to improve speech separation. Binaural TFS sensitivity may play a greater role in SRM than monaural TFS sensitivity, and ITD discrimination may reflect this better than TFS1 score. Comparison of TFS1, ITD discrimination, and the effect of TVC on SRM may indicate whether SRM is degraded by an overall lack of sensitivity to TFS or by binaural processing specifically.

Neher *et al.* (2012) suggested that all the measures they tested may be affected by a common, age-related mechanism. Therefore, age was included as a regressor in the current study.

6.4.1 **Stimuli**

6.4.1.1 Fine structure based pitch discrimination

Reference stimuli were harmonic complexes with components spaced by a modulation rate (f_m). Each component began in a random phase. They were band-pass filtered around a centre frequency (f_c) of 1.2 kHz. The bandwidth was dictated by f_m, passing five components with a 30 dB/octave roll off. Two different f_m's were tested: 100 and 200 Hz, giving two different ranges of harmonics. When f_m=100 Hz, the 10th to 14th harmonics were passed by the filter, with the twelfth harmonic at f_c (TFS1$_{H12}$, middle panel of Figure 6.7). When f_m=200 Hz, the fourth to eighth harmonics were passed by the filter, with the sixth harmonic at f_c (TFS1$_{H6}$, top panel of Figure 6.7). Listeners discriminated between these harmonic complexes and inharmonic versions where all frequency components were shifted by a δ Hz. Both shifted and reference stimuli were identically band-pass filtered based on the reference f_c and f_m. δ started at 50 Hz for both TFS1$_{H12}$ and TFS1$_{H6}$ and was limited to a maximum of half f_m. Above $f_m/2$ the shifted stimulus becomes increasingly similar to the reference stimulus as δ approaches f_m. If a listener cannot discriminate an $f_m/2$ shift, it is impossible to measure a threshold TFS1 score. Thresholds can be obtained consistently from NH listeners up to harmonic number 14 (Moore and Sek, 2009b), but HI listeners appear to have a much lower harmonic number above which thresholds cannot be measured (Hopkins and Moore, 2007). TFS1$_{H6}$ was included in case TFS1$_{H12}$ thresholds could not be obtained. However, TFS1$_{H6}$ thresholds may be partly based on discrimination of resolved harmonics (reflecting tonotopic, as well as temporal, encoding), rather than purely TFS from unresolved harmonics.

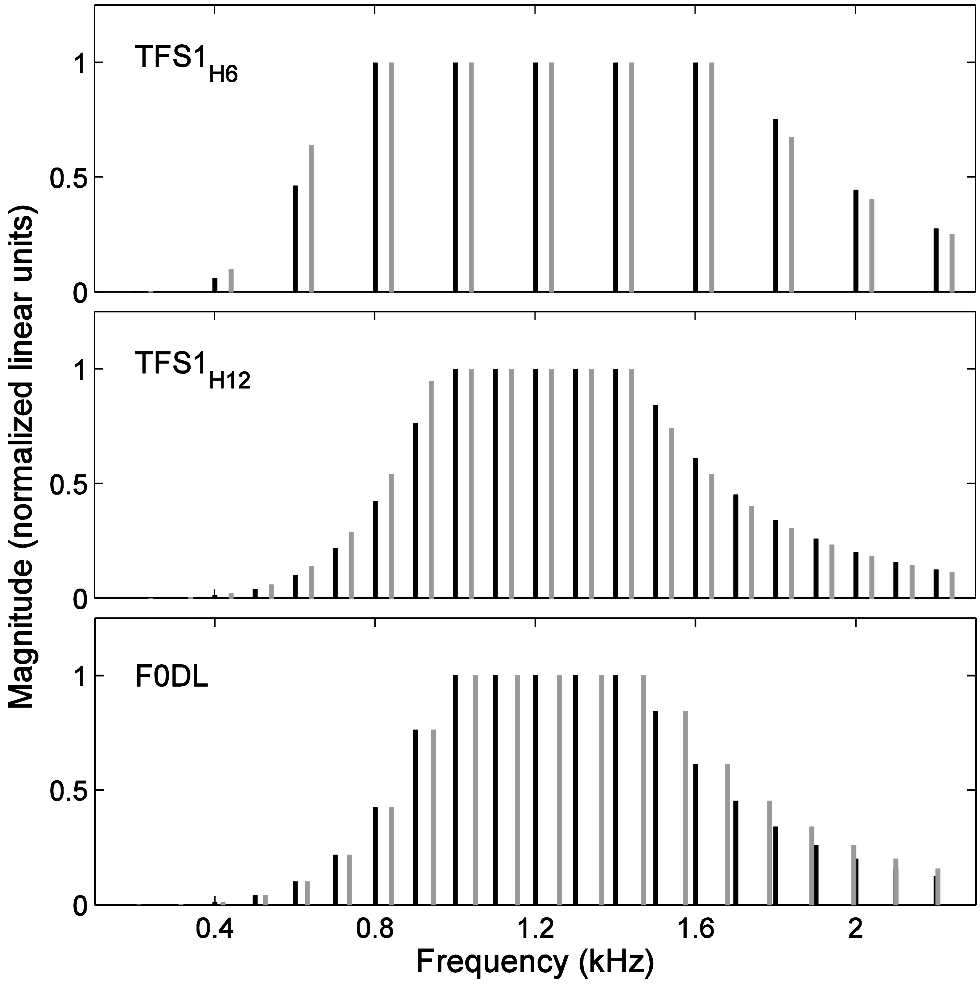

Figure 6.7: Schematic diagrams of the frequency spectra of the pitch discrimination stimuli on a linear frequency scale. The panels show, from top to bottom, $TFS1_{H6}$, $TFS1_{H12}$ and F0DL. Black lines show the reference stimuli and the grey lines show the shifted stimuli. The top two panels show a shift in all components of +40 Hz, and the bottom panel shows the shift in components due to a modulation rate shift of +5 Hz.

6.4.1.2 Envelope and fine structure based pitch discrimination

For the measurement of F0DLs, both reference and shifted stimuli were harmonic; δ was multiplied by the harmonic number of each component (f / f_m) to produce the shifted stimulus. Like $TFS1_{H12}$, the F0DL reference f_c was 1.2 kHz, and f_m was 100 Hz. Again, both reference and shifted stimuli were band-pass filtered with a 30 dB/octave roll off, passing five components. However, for F0DLs the filter shifted with the stimulus; so whilst the reference filter pass-band f_c was 1.2 kHz, the shifted filter pass-band moved with δ. The shifted filter pass-band was centred at $f_c + \delta(f_c / f_m)$, or 1.2+ δ (1.2/0.1) kHz. The F0DL stimuli are given in the bottom panel of Figure 6.7). The starting δ was 5 Hz and no maximum limit was imposed.

6.4.1.3 ITD discrimination

ITDs were presented using 500 Hz pure tones (ITD500). Onset and offset ramps were synchronous across ears so there was no ITD in the envelope. The left ear stimulus was $\sin(2\pi\cdot 500t)$ where t is the sampling time-vector, the right ear stimulus was $\sin(2\pi\cdot 500t + \Delta)$. The starting δ was π radians (1 ms) and this also the maximum δ limit.

The stimuli in the monaural tests (TFS1 and F0DL) were presented to the left ear. All stimuli were presented in the presence of a threshold-equalizing noise (TEN; Moore *et al.*, 2000) to mask combination tones and components of the complex tones falling outside of the pass-band of the filter. The TEN level at 1 kHz was 15 dB/$\mathrm{ERB_N}$ RMS below the overall RMS level of the test stimulus. This corresponded to an effective signal-to-noise ratio of 25 dB. In the monaural conditions, the TEN was played in the same ear as the test stimulus. For ITD500, uncorrelated samples of TEN were played to each ear. Uncorrelated noise was chosen to avoid any competing spatial cues in the noise, as an interaural correlation could produce an ITD coherent across frequency.

Listeners were tested in a sound insulating listening booth. All stimuli were created via MATLAB (MathWorks, 2012), an RME Hammerfall II digital-to-analog converter, and a custom-made amplifier. Stimuli were presented over a pair of Sennheiser HD200 circumaural headphones at 30 dB sensation level (SL) based on the listener's pure-tone audiogram.

6.4.2 **Procedure**

A two-interval, two-alternative, forced-choice task was used where one interval contained four 200 ms bursts of the reference stimuli (RRRR) and the other interval contained four 200 ms bursts, alternating between reference and shifted stimuli (RSRS). Each burst was separated by 100 ms, and ramped on and off by the rising and falling halves (respectively) of a 20 ms Hanning window. The two intervals were separated by 400 ms of silence. This paradigm is described further elsewhere (Moore and Sek, 2009a; Hopkins and Moore, 2010a). For all four discrimination tasks, a geometric, two-down, one-up adaptive procedure tracked 71% correct (Levitt, 1971) over eight reversals. Step sizes were a factor of 1.5^3 until the first reversal, 1.5^2 until the second reversal and 1.5 thereafter. The geometric mean of the last six reversals was used as the threshold estimate from a given track. For each test, up to three tracks were performed by each listener and their threshold was taken as the geometric mean of the thresholds from the completed adaptive tracks. If the listener failed to discriminate the maximally shifted stimulus from the reference stimulus three times within a track, 40 extra trials were presented with the maximum shift. In these cases, percent correct was calculated for all the trials in which the maximum shift was presented. If the listener scored better than 63% correct (above which performance can be assumed not to be due to chance with greater than 95% confidence) the adaptive track was run again. Otherwise, another 40 trials at maximum shift were presented and percent correct was recalculated from these trials. If the listener still scored worse than 63% correct, testing for that condition stopped and no threshold was obtained; otherwise, a final adaptive track was run.

Before beginning the experiment, the listeners were given a brief training period. In this period they heard example trials with the maximum shift in the shifted stimuli for TFS1_{H12}, TFS1_{H6} and ITD500, and a 10 Hz shift for F0DL. They heard eight trials without TEN and eight with TEN per condition.

6.4.3 **Results and discussion**

For each listener, the effect of TVC on SRM was calculated as the SRM without TVC minus TVC minus the SRM with TVC. These scores were then correlated against various variables variables including the discrimination task scores (see

Table 6.3). Before Pearson's product–moment correlations (*r*) were carried out, TFS1_{H6} and ITD500 thresholds were logarithmically transformed to normally distribute these data. As the F0DLs had a bimodal distribution (see right panel of Figure 6.8), they were correlated against other variables using Spearman's ranked correlation (ρ). A sequentially-rejective Bonferroni correction (Holm, 1979) was applied to compensate for multiple comparisons. TFS1_{H12} thresholds could only be measured for two listeners, so this condition was discarded from analysis. Observed power for the correlations, based on the mean coefficient of determination (r^2) of the correlations as the effect size and an α of 0.05, was 0.55 for 18 listeners and 0.59 for 20 listeners. Calculations were performed with GPower 3.1.3 (Kiel, Germany).

Table 6.3: A correlation matrix between listeners' age, PTA$_{LF}$, ITD500, TFS1$_{H6}$ score, F0DL and TVC effect on SRM. Pearson's product–moment and Spearman's rank correlations are denoted by r and ρ respectively; N and p denote listener number and significance value, respectively. Asterisks denote correlations significant ($\alpha < 0.05$) after Holm–Bonferroni correction. Daggers denote correlations not included in the Holm-Bonferroni correction as their significance was not tested.

		TVC effect on SRM	PTA$_{LF}$	Age	TFS1$_{H6}$	F0DL	ITD500
TVC effect on SRM	r						
	N	1					
	p						
PTA$_{LF}$	r	−0.757					
	N	20	1				
	p	< 0.001*					
Age	r	−0.098	0.011†				
	N	20	20	1			
	p	0.680					
TFS1$_{H6}$	r	−0.636	0.642	0.677			
	N	18	18	18	1		
	p	0.005*	0.004*	0.002*			
F0DL	ρ	−0.402	0.419†	0.179†	0.352		
	N	20	20	20	18	1	
	p	0.079			0.152		
ITD500	r	−0.166	0.020†	0.176	0.138		
	N	18	18	18	16	—	1
	p	0.511		0.485	0.609		

The effect of TVC on SRM was strongly correlated with PTA$_{LF}$. This can be seen in the left panel of Figure 6.9, where the effect of TVC on SRM decreases as PTA$_{LF}$ increases. PTA$_{LF}$ also correlated with SRM without TVC (r=−0.63, p=0.003; still significant after Bonferroni correction of α to 0.006). These results are consistent with the ANOVA with PTA$_{LF}$ included as a covariate in section 6.3.6. Peissig and Kollmeier (1997) found that audiometric threshold was a poor predictor of SRM, whereas Neher *et al.* (2011) found that low frequency audiometric threshold was moderately correlated with SRM. The conditions for SRM without TVC in the current study were similar to the conditions tested by Neher *et al.* (2011).

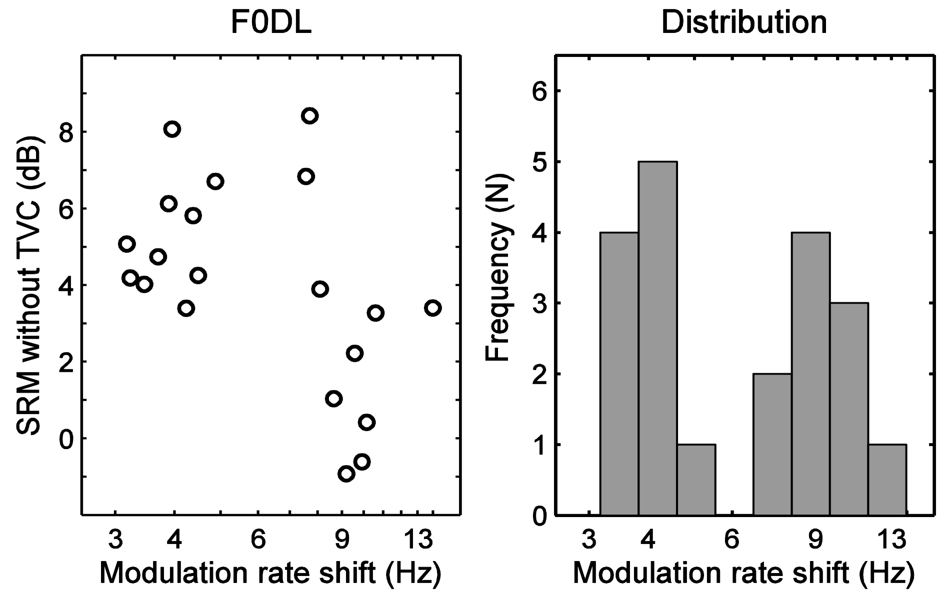

Figure 6.8: Individual listeners' SRM without TVC plotted as a function of F0DL (left) and the distribution of F0DLs (right).

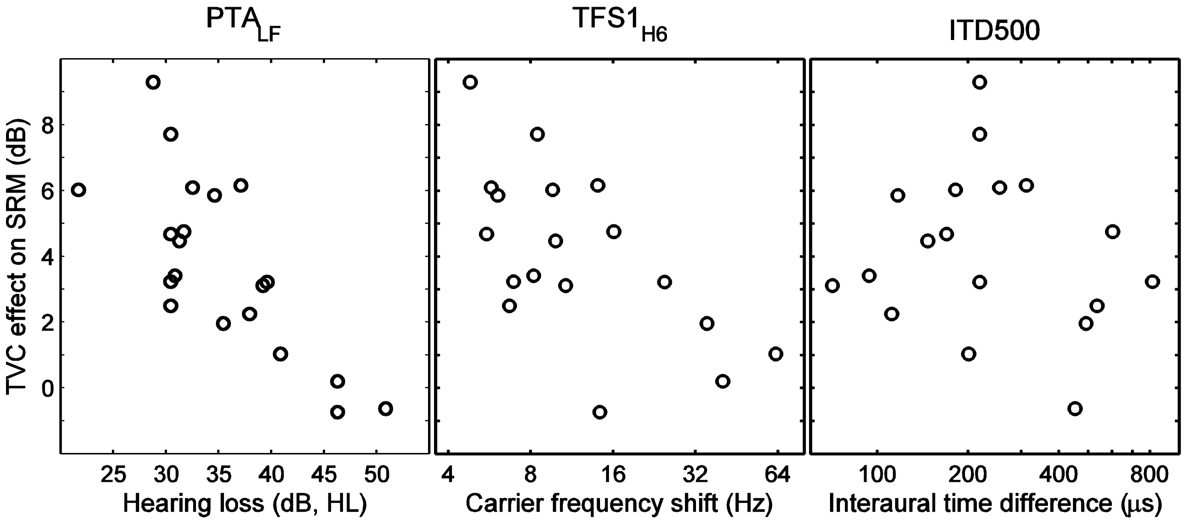

Figure 6.9: Individual listeners' effect of TVC on SRM plotted as a function of PTA_{LF} (left), $TFS1_{H6}$ score (middle), and ITD500 (right).

Higher PTA_{LF} thresholds may have been associated with less SRM due to insufficient audibility. However, foam plugs provide a closed seal which minimizes leakage, and the CAMEQ prescription to compensate for hearing loss has been verified to restore audibility of the whole dynamic range of speech stimuli (between 0.5 and 5 kHz) for listeners with audiometric thresholds greater than those in the current study (Hopkins *et al.*, 2008; Hopkins and Moore, 2011). Furthermore, in the current study SRM and the effect of TVC on SRM were analysed as difference measures; poor audibility of the speech would be expected to affect the four speech test conditions similarly.

It is possible that PTA_{LF} was related to the effect of TVC on SRM because sensorineural hearing loss affected SRM (in part) through TFS-ITD processing. Previous research shows relations between sensorineural hearing loss and poorer TFS-ITD processing (Hawkins and Wightman, 1980; Smoski and Trahiotis, 1986; Lacher-Fougère and Demany, 2005; Strelcyk and Dau, 2009; King *et al.*, 2014), and between TFS-ITD processing and SRM (Hopkins and Moore, 2011; Neher *et al.*, 2011; Neher *et al.*, 2012). However, in the current study ITD500 did not correlate with the effect of TVC on SRM (Figure 6.9, right panel), despite the ITD500 task relying on TFS-ITD processing. Additionally, ITD500 did not correlate with PTA_{LF} or with audiometric threshold at 500 Hz ($r=-0.12$, $p=0.628$).

Whereas Neher *et al.* (2011; 2012) and Hopkins and Moore (2011) tested ITD discrimination in quiet, the current study tested ITD discrimination in noise. It is possible that the noise confounded the mechanism that links ITD discrimination to SRM. The current study included binaurally uncorrelated TEN at 15 dB below target stimuli for two reasons. First, ITD500 was compared to speech reception in a competing-talker background rather than speech reception in quiet. Second, Henry and Heinz (2012) showed that the temporal representation of pure tones was more degraded as noise-masker level increased in auditory nerve fibres of chinchillas with noise-induced hearing loss. The current authors reasoned that presenting ITD500 with TEN may help reveal individual differences in TFS processing related to the listeners' hearing loss. However, this was not the case.

Strelcyk and Dau (2009) found a correlation between SRT in lateralized noise and ITD discrimination in noise 10 dB below masked threshold. However, Strelcyk and Dau (2009) only found this for diotic noise and at a high level above absolute threshold, not for uncorrelated noise and a lower SL similar to the ITD500 stimuli in the current study. Whilst dichotic maskers aided masking release in speech reception, they presented an additional challenge in lateralisation (footnote 7 in Strelcyk and Dau, 2009). Also, Strelcyk and Dau (2009) found that the differences in ITD discrimination between NH and HI listeners were most pronounced with a high level tone in quiet. Therefore, ITD discrimination of low-SL pure tones in the presence of uncorrelated TEN might use different mechanisms than using ITDs in SRM, which may be more closely related to ITD discrimination at high SLs.

It is possible that PTA_{LF} was correlated with the effect of TVC on SRM because both variables were dependent on a third, common variable, rather than threshold elevation directly affecting the effect of TVC on SRM. Age has been associated with poor TFS sensitivity (e.g., Ross *et al.*, 2007; Grose and Mamo, 2010; Hopkins and Moore, 2011; Moore *et al.*, 2012b) and reduced SRM (Gelfand *et al.*, 1988; Marrone *et al.*, 2008a; Marrone *et al.*, 2008c; Neher *et al.*, 2009; Neher *et al.*, 2012), so it may be a good candidate for the common cause of increased PTA_{LF} and decreased effect of TVC on SRM. However, in the current study age was not correlated with the effect of TVC on SRM or ITD500, and the correlation between the effect of TVC on SRM and PTA_{LF} remained after age was partialled out (see Table 6.4). Neher *et al.* (2011) argued that a narrow age range (60–78 years) limited the effect size of cognitive measures previously found to relate to age and SRTs with spatially separated speech (Neher *et al.*, 2009). The narrow age range (22 years) in the current study also may have limited the effects of age. However, $TFS1_{H6}$ score increased (poorer performance) as listener age increased (Figure 6.10, left panel), leading to a moderate-to-strong positive correlation. This indicates some age-related deficit in monaural TFS processing.

$TFS1_{H6}$ scores also correlated with the effect of TVC on SRM (r=−0.64, p<0.006; Figure 6.9, middle panel) and with PTA_{LF} (r=0.64, p<0.005; Figure 6.10, right panel). These correlations remained significant once age was partialed out (see Table 6.4). A partial correlation was tested to see if $TFS1_{H6}$ score correlated with the TVC effect on SRM after the variance due to PTA_{LF} was accounted for. This was not the case (r=−0.34, p=0.18). Conversely, the partial correlation between PTA_{LF} and the TVC effect on SRM, with the variance due to $TFS1_{H6}$ score accounted for, was significant (r=−0.50, p=0.04) before correction for multiple comparisons. This suggests that PTA_{LF} may have measured some individual differences in the TVC effect on SRM that was independent from the individual differences explained by $TFS1_{H6}$ score.

The difference in speech understanding (specifically, $TMR_{50\%}$) in the Co-located spatial configuration with TVC and without TVC was expected to be due to the ability to use TFS cues when the speech was not vocoded, but not when vocoding disrupted the TFS. However, $TFS1_{H6}$ score was not significantly correlated with the difference in TMR$_{50\%}$ between the Co-located conditions with TVC and without TVC (r=0.187, p=0.458). This may be because either TFS1$_{H6}$ did not measure TFS sensitivity exclusively and accurately, or TVC did not exclusively disrupt TFS cues (as asserted by others; e.g., Shamma and Lorenzi), or for both reasons.

Table 6.4: A matrix of partial-correlation coefficients (r) between PTA$_{LF}$, and TFS1$_{H6}$ score with listener age partialled out. N denotes listener number in correlation; p denotes significance value.

Controlled for age		TVC effect on SRM	PTA$_{LF}$	TFS1$_{H6}$
TVC effect on SRM	r	1		
	N			
	p			
PTA$_{LF}$	r	−0.760	1	
	N	20		
	p	< 0.001		
TFS1$_{H6}$	r	−0.696	0.739	1
	N	18	18	
	p	0.002	0.001	

There was a trend for F0DLs to decrease with increasing SRM without TVC ($\rho=-0.57$, p=0.009; Figure 6.8, left panel), but this was not significant after Bonferroni correction of α to 0.006. This moderate correlation suggests that the ability to discriminate pitch shifts may help using spatial cues to listen to speech in noisy backgrounds. However, F0DLs were not related to the effect of TVC on SRM or to TFS1$_{H6}$ scores, suggesting that F0DLs do not relate well to TFS processing that contributes to spatial unmasking. A lack of a relationship to TFS processing is perhaps not surprising if F0DLs are driven by spectral analysis, or envelope or distortion product analysis of unresolved components (Oxenham *et al.*, 2009).

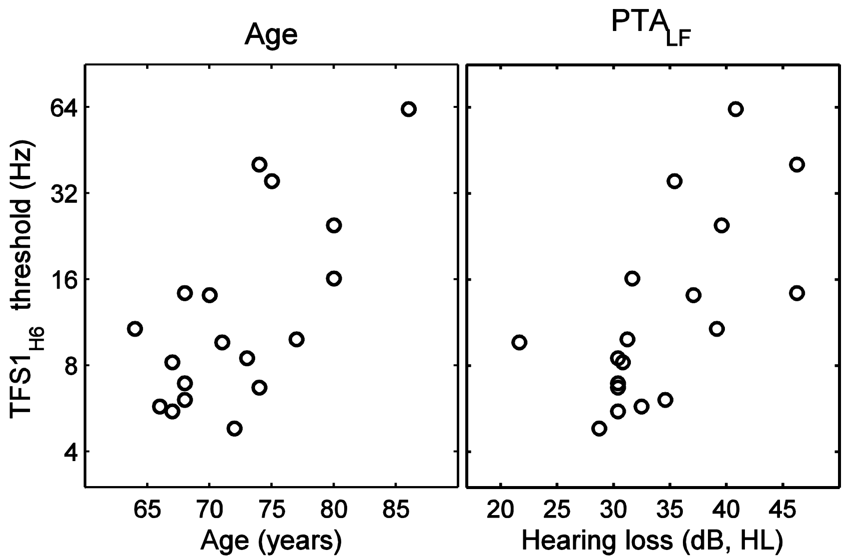

Figure 6.10: Individual listeners' TFS1$_{H6}$ thresholds plotted as a function of age (left), and PTA$_{LF}$ (right).

6.5 Conclusions

The current study aimed to determine the contribution of TFS ITDs to SRM by manipulation of TFS IPDs via vocoding, and to find a good predictor of individual differences in benefit from TFS ITDs to SRM. Older, HI listeners benefited from spatial separation, although more variably and on average not as much as younger, NH listeners in previous studies (Marrone *et al.*, 2008a; Marrone *et al.*, 2008c; Neher *et al.*, 2011). Removing the ITDs in the TFS of the signals as recorded at the ears, by applying TVC in phase across ears, significantly reduced the benefit of spatial separation. This suggests that some older HI listeners are capable of using fine-tuned, binaural temporal cues to aid speech perception in multi-talker environments. This has implications for the potential trade-offs, or sacrifices, of disrupting binaural temporal acoustic information with signal processing strategies employed in modern, digital hearing aids.

The current study assessed the predictive power of several measures that are thought to rely on auditory temporal processing in a manner similar to previous studies (e.g., Neher *et al.*, 2009; Strelcyk and Dau, 2009; Hopkins and Moore, 2011; Neher *et al.*, 2011; Neher *et al.*, 2012). The audiogram below 1.5 kHz and monaural TFS sensitivity were both well correlated with the effect of TVC on SRM, whilst pure-tone ITD discrimination was not. This is surprising, as both pure-tone ITD discrimination and SRM rely on binaural processing of TFS information. Low-frequency audiometric thresholds may be a convenient metric to determine who may benefit from hearing aids that preserve binaural cues in the TFS.

Chapter 7. General Discussion

7.1 Thesis objectives

This thesis aimed to determine how hearing loss and age affect the use of interaural time cues for binaural hearing. Specifically, this thesis tested how interaural phase difference (IPD) sensitivity, which requires precise temporal processing of binaural signals, changes with age and hearing loss. These relations were also tested with respect to understanding speech in a background of competing talkers from different directions.

The first phase of the thesis involved repeating a study of how cochlear hearing loss affects temporal fine structure (TFS) in IPD discrimination by Lacher-Fougère and Demany (2005), but without confounding the effects of old age and hearing loss. TFS- and envelope-IPD discrimination thresholds were separately tested in both young and old normal-hearing (NH) and hearing-impaired (HI) listeners. The second phase aimed to establish whether TFS- and envelope-IPD processing is affected by age and hearing loss due to poor phase locking in the auditory nerve or brainstem. To test this, the frequency following response (FFR) to sustained amplitude-modulated (AM) tones was measured and compared to age, hearing loss and IPD thresholds. The third phase determined how TFS IPDs in speech may help speech perception in competing-speech backgrounds, and whether the effects of age and hearing loss on IPD discrimination translate to the ability to take advantage of IPDs in day-to-day life.

7.2 Methodological issues

7.2.1 Minimising the need for training in IPD discrimination

IPD discrimination performance varies widely, even for young, NH listeners, with some listeners performing well without practice, some performing poorly even after lengthy practice, and others improving to varying degrees over time (Wright and Fitzgerald, 2001; Rowan and Lutman, 2006). Wright and Fitzgerald (2001) suggested that a substantial amount of training is required to obtain asymptotic discrimination thresholds. However, Hopkins and Moore (2010b) found that little training was needed for stable performance in their IPD discrimination protocol. In order to repeat the Lacher-Fougère and Demany (2005) study with both young and old, normal-hearing (NH) and hearing-impaired (HI) listeners, a large sample of listeners was required. With a large sample size, minimal training could dramatically expedite testing. In chapter 2, a systematic comparison of training with the presentation paradigms used by Lacher-Fougère and Demany (2005) and by Hopkins and Moore (2010b) showed that the latter paradigm (AAAA vs. ABAB) could produce thresholds as low as those produced after two hours of training within the first few estimates. This result was capitalised upon in chapter 3 by using the Hopkins and Moore (2010b) presentation paradigm to test TFS- and envelope-IPD discrimination, instead of the paradigm used by Lacher-Fougère and Demany (2005).

7.2.2 Continuous, rather than categorical, treatment of age and hearing-loss effects

Another modification to the protocol for repetition of the Lacher-Fougère and Demany (2005) study was to assess the effects of age and hearing loss as continuous variables, allowing them to vary across as broad a range as possible, rather than matching listener groups on age or hearing loss and comparing performance between discrete categories (young NH, young HI, old NH and old HI). In this way age was statistically controlled for when assessing the effect of absolute threshold, and *vice versa*, revealing the independent effects of hearing loss and age on TFS- and envelope-IPD discrimination.

The study of how age and hearing loss affect the use of TFS IPDs in speech perception, in competing speech backgrounds, did not include a comparison to young or NH listeners (chapter 6). Young and NH listeners were not tested due to constraints on the availability of participants and use of facilities. Only old HI listeners were tested, because the research was most relevant to this demographic. The effects of age and hearing loss within this single group were tested as covariates nonetheless. The use of TFS IPDs in speech perception in young NH listeners was tested previously at the Eriksholm Research Centre using headphones and a simulation of spatial cues using digital head-related transfer functions (Andersen *et al.*, 2010). To the author’s knowledge, the contribution of TFS IPDs to the speech perception of young HI and old NH listeners has not been tested by tone-vocoding previously.

7.2.3 Determining the sources of far-field recording of phase locking

In order to determine whether IPD sensitivity is affected by changes in phase locking with hearing loss or age, FFR was measured with electro-encephalography (EEG) in two different electrode montage orientations. The vertically-oriented montage (from seventh cervical vertebra to forehead) was expected to emphasise a dipole of neural activity in the rostral brainstem. The horizontally-oriented montage (from right mastoid to left) was expected to emphasise a dipole in the neural pathway closer to the cochlea (Stillman *et al.*, 1978; Scherg and Brinkmann, 1979; Scherg and Von Cramon, 1985; Galbraith, 1994; Galbraith *et al.*, 2000; Galbraith *et al.*, 2001). The assumption that horizontal and vertical electrode montage orientations would record, respectively, more peripheral and more central sources of FFR was tested in young, NH listeners via the latency of the FFR in both montages using group delay (chapter 4). It was assumed that more peripheral sources should have shorter latencies than more central sources. If the two montages could measure distinct sources of FFR, then they could be used to determine whether more peripherally-sourced FFR or more centrally-sourced FFR exhibited age-related declines or related to IPD sensitivity.

In Chapter 4, both montages appeared to measure sources of envelope-FFR in the rostral brainstem, around the level of the inferior colliculus, lateral lemniscus, or later (based on the group delay). However, for TFS-FFR, the group delay was shorter for the horizontal montage than for the vertical montage. This validated the use of horizontal and vertical montages to measure phase locking at more peripheral and more central sites in the sub-cortical auditory system. However, it is possible that the TFS-FFR (particularly in the horizontal montage) was contaminated by the cochlear microphonic, which is generated by the hair cells and is not neurogenic.

In chapter 5, age- and hearing-loss-related changes in FFR were measured with both vertical and horizontal montages. However, latency of the response in each montage was not measured by group delay in chapter 5, because in chapter 4 there were 25% missing data (17 out of 23 listeners had missing data), and such large amounts of missing data would be an obstacle to obtaining enough power for statistical analysis in chapter 5. If age- or hearing-loss-related deficits occurred peripherally, then they should be observable in the horizontal FFR; if such deficits occur more centrally, they should only appear in the vertical FFR.

The reason for the large amount of missing data in chapter 4 is that criteria for FFR strength and group delay limits were imposed. As old listeners have weaker FFR than young listeners (Clinard *et al.*, 2010; Marmel *et al.*, 2013), it is likely that the older listeners' FFRs in chapter 5 would be less likely to pass the FFR strength criterion, and therefore increase the amount of missing data. Whilst Marmel *et al*. (2013) measured group delay in 25 out of 27 listeners of varying age and hearing loss severity, they did not impose any criteria on FFR strength or group delay limits. For a vertical montage, Marmel *et al*. (2013) found group delay increased with age, but was highly variable also.

7.3 Hearing loss and temporal spatial hearing

7.3.1 **IPD discrimination**

In chapter 3, increasing absolute threshold (250 and 500 Hz average) was correlated with decreasing TFS-IPD sensitivity independent of the listener's age. However, no such decrease in envelope-IPD sensitivity was seen with hearing loss. Together these findings supported the assertions by Lacher-Fougère and Demany (2005) that cochlear damage leads to a specific deficit in TFS processing.

Monaural measures of TFS processing are associated with cochlear damage (e.g., Strelcyk and Dau, 2009; Hopkins and Moore, 2011) and correlate with absolute threshold (even in the NH threshold range, Moore *et al.*, 2012a). However, binaural TFS processing, such as pure-tone IPD discrimination, does not always correlate well with hearing loss (e.g., Hopkins and Moore, 2011; Moore *et al.*, 2012a; Moore *et al.*, 2012b). Differing stimulus levels may have been responsible for the different relations between hearing loss and IPD discrimination between Lacher-Fougère and Demany (2005) and Moore and colleagues (Hopkins and Moore, 2011; Moore *et al.*, 2012a; Moore *et al.*, 2012b). Whereas Lacher-Fougère and Demany (2005) used a presentation level of 75 dB SPL, Moore and colleagues used a sensation level of 30 dB relative to each listener's absolute threshold. However, in chapter 3 the sensation levels were similar to those used by Moore and colleagues, but the results were similar to those of Lacher-Fougère and Demany (2005). It seems unlikely, therefore, that level differences are responsible for the discrepancies in previous findings.

The use of AM tones by Lacher-Fougère and Demany (2005) and in the study in chapter 3, compared to the use of pure tones by Moore and colleagues, may underlie the discrepant results. In the AM tones the envelope provided conflicting IPD information to the TFS (IPD in one, but not the other). For pure tones there is no envelope information to carry an IPD. Envelope coding is enhanced in auditory nerve fibres from chinchilla ears with noise-induced hearing loss, possibly due to damaged inner hair cell stereocilia (Liberman and Kiang, 1984; Kale and Heinz, 2010), and modulation-depth perception is enhanced in the impaired ears of unilaterally-impaired listeners (presumably due to reduced cochlear compression, Moore *et al.*, 1996). If envelope encoding dominates TFS encoding in HI listeners, but the envelope contains no IPD (as in the TFS-IPD discrimination conditions with AM tones), then this may obscure perception of TFS IPDs more than when no envelope exists (*i.e.* for pure tones).

Envelope-IPD discrimination did not deteriorate with hearing loss, so HI individuals may be able to use envelope IPDs to compensate for poor TFS IPD sensitivity. However, there are reasons to believe that envelope IPDs do not provide as useful information as TFS IPDs. First, because distinguishing the envelope of the target sound from the envelope of any competing sounds may be difficult unless the envelopes are sufficiently different that they would be processed separately (Dau *et al.*, 1997). Second, the human head will only create small IPDs at low modulation rates; in chapter 3 envelope-IPD discrimination thresholds were rarely as low as the maximum IPD that would occur in the free-field (about 7° for 20 Hz modulation rate). Although, Henning (1974) found listeners could discriminate envelope IPDs below 7° for 50, 125 and 300 Hz modulation rates with a 3900 Hz carrier. Third, envelope-IPD sensitivity is more impaired by reverberation than TFS-IPD sensitivity (Devore and Delgutte, 2010), particularly with high-frequency carriers (Rakerd and Hartmann, 2010; Monaghan *et al.*, 2013). When this happens, it is likely interaural level differences (ILDs) are relied upon instead (Devore and Delgutte, 2010; Rakerd and Hartmann, 2010). To compound matters, more salient envelope information may obscure temporal gaps in masking noises (Glasberg and Moore, 1992) and impair speech intelligibility in fluctuating background noise (Moore and Glasberg, 1993).

In chapter 5, there was no evidence of phase locking to AM tones decreasing with increasing absolute threshold, and the correlations between absolute threshold and TFS-IPD discrimination found in chapter 3 were not subsequently found in chapter 5. This may be because fewer listeners with greater-than-mild hearing loss were tested in chapter 5 than chapter 3. This may have limited the statistical power to measure effects of elevated thresholds on binaural TFS processing. For the narrow range of absolute thresholds considered as normal, Moore *et al*. (2012b) found monaural TFS processing correlated with absolute threshold, but binaural TFS processing did not. It would be premature, however, to claim that cochlear hearing loss does not affect phase locking or the representation of TFS IPDs.

7.3.2 **The benefit of TFS to spatial release from masking**

In chapter 6, speech reception thresholds (SRTs) were measured in competing speech maskers that were located either directly in front of the listener (co-located with the target speech) or 45° to the left and right of the listener (separated from the target speech). Comparison of these SRTs revealed 4 dB of spatial release from masking (SRM), on average, for old HI listeners with symmetric sensorineural hearing loss. However, when the speech signals were tone vocoded synchronously at the ears (thus removing the IPDs from the TFS in each processed frequency-band) the same listeners achieved negligible SRM (0.3 dB). The listeners generally required target-to-masker ratios above zero in both spatial configurations with vocoding, and in co-located speech without vocoding. This implies that old, HI people will struggle to understand speech in competing speech backgrounds, but some may still be able to use TFS-IPD cues to gain a small, but potentially significant, advantage. However, the advantage is much less than that gained by young, NH listeners (Andersen *et al.*, 2010).

Low frequency (250 to 1500 Hz average) audiometric thresholds ($\mathrm{PTA}_{\mathrm{LF}}$) were strongly correlated with the difference between SRM with vocoding and without vocoding. The higher a listener's $\mathrm{PTA}_{\mathrm{LF},}$ the less benefit from TFS-IPDs to SRM they received. It was surprising that pure-tone IPD discrimination by these listeners did not relate well to either $\mathrm{PTA}_{\mathrm{LF}}$ or to the effect of TVC on SRM. Neher *et al*. (2012) found that pure-tone IPD discrimination in quiet related to SRTs in laterally-separated speech maskers. In chapter 6 it was SRTs in separated speech maskers that contributed the majority of the variation in individuals' difference in SRM with and without TVC. Neher *et al*. (2011; 2012) studied IPD discrimination in quiet, whereas in chapter 6 the tones were presented in threshold equalising noise at 15 dB signal-to-noise ratio (SNR). It is possible that this confounded the relationship between IPD discrimination and the effect of TVC on SRM. Strelcyk and Dau (2009) tested IPD-based lateralisation at different levels and in various noises; they came to the conclusion that background noise is a hindrance to IPD discrimination. Only lateralisation of a high-level pure-tone in a high-level diotic masker related to SRTs in laterally-separated maskers (Strelcyk and Dau, 2009). This may implicate cochlear neuropathy of high-threshold afferent nerve fibres (see section 7.4.4).

The conflicting reports of relations between IPD discrimination and hearing loss, and of relations between IPD discrimination and SRM, perhaps demonstrate that IPD discrimination can be affected by differences in the stimuli. This needs to be taken into account when considering how IPD discrimination ability may relate to an individual’s ability to use IPDs to separate sounds from different azimuths beneficially in day-to-day life.

7.4 Ageing and temporal spatial hearing

7.4.1 **IPDs**

In a sample of listeners ranging from 18 to 83 years old, TFS-IPD discrimination thresholds increased with increasing age (chapters 3 and 5). This suggests that older listeners are less sensitive than younger listeners to IPDs in the TFS of sounds. Many studies show poorer TFS-IPD discrimination by NH listeners in middle and old age (e.g., Grose and Mamo, 2010; Moore *et al.*, 2012b; Füllgrabe, 2013); this thesis shows that this remains the case even when variations in absolute thresholds (including NH and HI listeners) are accounted for. There also appears to be a larger variation in TFS-IPD discrimination performance in older listeners than in young listeners. Poorer TFS-IPD sensitivity is likely to impair older listeners’ perception of auditory space.

In a group of hearing aid users between 64 and 86 years old, all with gently sloping, bilateral hearing losses, pure-tone IPD discrimination ability did not vary systematically with age (chapter 6). Pure-tone IPD discrimination for this group of listeners was generally poorer than for the young NH listeners tested in chapters 2, 3 and 5. It is possible that pure-tone IPD discrimination did not correlate with age in chapter 6 because the age range was too narrow (Neher *et al.*, 2011). The benefit of TFS IPDs to SRM also did not correlate with age, despite the benefit being much smaller for this group of old HI listeners than for the younger NH listeners reported by Andersen *et al*. (2010). Comparing the benefit of TFS IPDs to SRM for young and old listeners, whilst controlling for the effects of their absolute thresholds, may be worthwhile.

7.4.2 FFR

In chapter 5, the phase coherence of FFR reflecting phase locking to the TFS of AM tones was weaker in older listeners than in young listeners. This is consistent with studies of the effect of age on the strength of FFR to pure tones (Clinard *et al.*, 2010; Marmel *et al.*, 2013) and to speech stimuli (Anderson *et al.*, 2012). The results in chapter 5 demonstrate an age-related phase locking deficit to TFS at lower frequencies than reported by Clinard *et al*. (2010). Clinard *et al*. (2010) found FFR magnitude decreased with age around 1 kHz but not around 500 Hz, whereas in chapter 5 FFR phase coherence decreased with age at frequencies from 242 to 723 Hz. This suggests that age-related declines in phase locking may not be limited to a decline at higher frequencies.

Chapter 4 suggested that a horizontal electrode montage may measure TFS-FFR from an earlier source than a vertical montage. However, chapter 5 did not show any significant differences between horizontally- and vertically-recorded TFS-FFR that interacted with listener age.

Age was also related to a decrease in the ability to discriminate envelope IPDs. The author is unaware of any other study of age-related declines in envelope-IPD discrimination. Age-related changes in temporal-envelope processing have been studied in other ways, such as with temporal modulation transfer functions (TMTFs; e.g., Purcell *et al.*, 2004; He *et al.*, 2008). TMTFs typically show a decrease in sensitivity to high modulation rates with age, but not to low rates, such as 20 Hz. Indeed, the age-related declines in FFR to the envelopes of the AM tones followed a similar pattern to the TMTFs: no age-related decline in phase coherence was seen at lower modulation rates (16 and 27 Hz), only at a relatively high rate (145 Hz). Similarly, others have found that the amplitude of envelope FFR decreases with age primarily at high modulation rates, rather than low rates (Purcell *et al.*, 2004; Parthasarathy and Bartlett, 2012; Parthasarathy *et al.*, 2014). Also, the upper frequency limit of envelope FFR correlates strongly with the highest modulation frequency at which modulation can be behaviourally detected (Purcell *et al.*, 2004).

Phase coherence to the TFS of AM tones, and to the 145 Hz modulation-rate envelope, declined with age, even when absolute threshold (250 and 500 Hz average) was partialled out. Clinard *et al*. (2010) restricted the selection of listeners to those that had normal hearing up to 4 kHz and only mild hearing loss between 4 and 8 kHz.

Chapter 5 showed that even when hearing loss severity (particularly at high frequencies) ranged more considerably, it did not contribute to the age-related deterioration in phase locking at low frequencies, despite high-frequency hearing loss correlating with age.

Performance deficits in old age are widely reported and some studies have reported a link between phase locking precision and understanding speech in noise (e.g., Anderson *et al.*, 2012). In chapter 5 we attempted to link phase locking precision to spatial hearing through correlating FFR phase coherence with IPD discrimination.

7.4.3 **Lack of relation between FFR and IPDs**

The ability to perform spatial listening tasks, such as localisation or detecting speech in noise, may be impaired in old age because of less consistent phase locking at each ear. In chapter 5, we found that, whilst both FFR strength and IPD discrimination thresholds worsened with age, FFR phase coherence and IPD discrimination only correlated weakly. When age was partialled out, the Pearson's *r* for correlations between FFR and IPD discrimination were reduced almost to zero. This suggests that any variability in phase locking that influences IPD discrimination are due to the effects of ageing. On the other hand, other processes involved in IPD discrimination, unrelated to FFR, also appear to deteriorate with age. The correlations between age and IPD thresholds changed little when FFR was partialled out, and remained significant. Further study would be useful to determine what these mechanisms are. A possible candidate is binaural integration occurring in the olivary complex of the brainstem, but higher-level processing such as auditory object classification or even cognitive capacity (such as attention or working memory) may also contribute.

A binaural listening task, which is thought to demonstrate a listener's ability to use interaural cues and selective attention (Bharadwaj *et al.*, 2014), has been found to relate to age-related differences in FFR to speech. Ruggles *et al.* (2012) compared the FFR to the /dah/ syllable with recall of digits spoken from in front of the listener whilst competing digits were spoken 15° to the left and right. The spatial characteristics of the speech, and three levels of reverberation, were simulated and presented over headphones. Reverberation distorts the TFS at each ear and reduces the interaural correlation (Ruggles and Shinn-Cunningham, 2011). Ruggles *et al.* (2012)

found that, for young listeners, FFR to the envelope, but not to the TFS, related to SRM in moderate reverberation. For older listeners, FFR to the TFS, but not the envelope, related to SRM. The results of Ruggles *et al*. (2012) are useful because they show a link between neural phase locking and a form of spatial hearing that is very similar to everyday listening situations (e.g. listening to one person talk whilst others talk in a reverberant environment). However, the study did not test whether the quality of phase locking affected SRM through the accuracy of IPD representation in the different speech streams.

Understanding how the auditory brain extracts and uses IPDs to facilitate spatial hearing in noisy and reverberant backgrounds may help explain why age affects IPD sensitivity and spatial hearing. Such conditions are very difficult to listen in, because the IPDs in target sounds may be easily confused with IPDs from reflections and competing sounds. Dietz *et al*. (2013) tested how IPDs might be used in such conditions using amplitude-modulated binaural beats (AMBBs). When two pure tones with a slight frequency mismatch are presented, one to each ear, a beating sound is produced at the difference in tone frequency. By amplitude-modulating these tones, synchronously across ears at the beat frequency, a full cycle of IPDs (e.g. −180° to +180°) is produced per modulation cycle. The phase relation between the AM and binaural beat dictates where in the modulation cycle a given IPD occurs. Dietz *et al*. (2013) argued that AMBBs contain similar spatial cues and acoustic properties to natural environments where reverberant energy reflects back to the listener with conflicting IPDs—environments to which the binaural system is likely well adapted.

Young listeners could match AMBBs to AM tones with constant TFS-IPDs (Dietz *et al.*, 2013), which typically produce a lateral percept. Dietz *et al*. (2013) found that young, NH listeners matched AMBBs to TFS-IPDs that occurred in the rising portion of the AM. Dietz *et al*. (2013) also reported magneto-encephalography (MEG) data that supported the behavioural data; MEG responses to AMBBs adapted more from (and were thus more related to) static IPDs below 180° than static IPDs above 180°. Dietz *et al*. (2013) reasoned that these rising portions would be less contaminated by conflicting IPDs in sound reflections. Consistent data were found from neurons in the medial superior olivary nuclei of gerbils and inferior colliculi of guinea pigs (Dietz *et al.*, 2014). Neurones typically responded most when the IPD dominated the total neural response in the rising portion of the modulation. It may be worth determining whether it is IPDs in the rising portion of modulations that older, and HI, listeners are less sensitive to than young NH listeners.

FFRs to monaural stimuli appear to relate well to monaural frequency discrimination thresholds (Marmel *et al.*, 2013). It is possible that the TFS-FFR to the AM tones that differed in each ear (chapter 5) would have related better to a behavioural test of monaural TFS sensitivity than to a binaural TFS processing task such as IPD discrimination. FFR to binaural stimuli may have related better to IPD discrimination. FFR can be measured to binaural stimuli and is greatest in magnitude when the signal is diotic. When IPDs are imposed on these binaural stimuli, systematic reductions in FFR amplitude are seen with increasing IPD (Clark *et al.*, 1997; Ballachanda and Moushegian, 2000). Binaural interaction is also seen when the sum of the left- and right-monaural auditory brainstem responses (ABRs) is subtracted from the binaural ABR, leaving a residual waveform. This binaural interaction component (BIC) changes systematically with IPDs and ILDs in click stimuli (Furst *et al.*, 1985). In FFR however, the BIC appears to be unaffected by IPDs (Ballachanda and Moushegian, 2000), but may be affected by ILDs (Krishnan and McDaniel, 1998).

Studies have shown how cortical responses to binaural stimuli with interaural information change with age, and have related these responses to behavioural measures of IPD sensitivity or localisation precision (Ross *et al.*, 2007; Briley and Summerfield, 2014). Using MEG, Ross *et al.* (2007; 2008) measured the cortical potential and the temporary lag in the steady-state response’s phase (relative to stimulus phase) after the carrier of a binaural AM tone is switched from in phase across ears to an IPD of 180° in a modulation trough. They found the cortical potential and steady-state phase deviation decreased with increasing carrier frequency. The highest carrier frequencies at which these physiological responses were observed (*i.e.* the threshold frequency) were similar to behavioural thresholds of detecting the IPD switch. Behavioural and physiological thresholds decreased with age; even by middle age (40 to 60 years) threshold frequencies were lower (poorer) than they were for young adults (Ross *et al.*, 2007; Ross, 2008). These measures show an early onset of binaural TFS processing deficits with age and may work well as early warning indicators of problems localising sounds in later life.

Briley and Summerfield (2014) found decreases localisation and associated physiological responses in later old age (after 73 years). These listeners were less able to behaviourally detect changes in sound azimuth at wide lateral angles than younger listeners. Briley and Summerfield (2014) also measured cortical EEG activity to changes in sound azimuth from these listeners and passed it through an opponent-channel model of azimuthal location coding (van Bergeijk, 1962; Magezi and Krumbholz, 2010). Poorer localisation in older-old age could be relatively well explained by shallower slopes for the two channels in the model. Briley and Summerfield (2014) note it may be useful to determine how temporal processing (particularly TFS) may contribute to differences between age groups in the cortical representation of a sound's angle.

Despite FFR and IPD discrimination both deteriorating with age, only a weak relation was found in chapter 5. This may be due to a lack of useful interaural information in the stimuli for FFR measurement. An inclusion of such information may reveal changes in temporal processing that influence localisation. However, the processing of IPDs may convert phase-locked information to rate-based information better measured through analysis of other electrophysiological responses than FFR.

7.4.4 **Cochlear neuropathy**

Cochlear neuropathy has been shown to occur with ageing and noise exposure even when absolute thresholds return to normal (Kujawa and Liberman, 2009; Lin *et al.*, 2011; Makary *et al.*, 2011; Sergeyenko *et al.*, 2013). This has been implicated in temporal processing of sounds at moderate-to-high levels, as there are fewer low-spontaneous-rate fibres to encode the sound (Furman *et al.*, 2013). However, this should not affect the phase locking of the remaining fibres. Instead, it should decrease the fidelity of the encoding relative to any noise within the auditory filter channel once the fibre firings are combined in the cochlear nucleus (Joris and Smith, 2008). IPD discrimination at moderate-to-high stimulus levels may be affected by cochlear neuropathy whilst absolute thresholds are not. It was expected that FFRs to high-level stimuli may reveal any effects of cochlear neuropathy (unrelated to absolute threshold) that may impair IPD discrimination. In chapter 5, the evidence for this is weak.

7.4.5 Combatting auditory ageing through training

Neural phase locking appears to be very plastic. Even after only one month of training in discriminating harmonic complexes, young adults show improved synchronisation in FFRs to these stimuli (Carcagno and Plack, 2011). Speech-based training also appears to show benefits for hearing in old age. Anderson *et al*. (2013) put 55 to 70 year-old participants through an eight-week training program that focused on formant transitions in speech, from syllables to whole stories (thus requiring attention and working memory too). Trainees showed improved formant transition representation in the FFR and better vowel representation in the FFR to syllables in noise than an age-matched control group who followed a programme of careful, attentive listening without a speech cue focus. Behaviourally, speech-in-noise perception, short-term memory and processing speed also improved through training (Anderson *et al.*, 2013).

Playing music requires players to parse melodies from each other and from background noise, and has been suggested to require similar auditory processes to speech perception, particularly in competing speech and noise (Parbery-Clark *et al.*, 2009). In old age, life-long musicians exhibit less delayed neural phase locking to the speech formant transitions than non-musicians (Parbery-Clark *et al.*, 2012). This suggests that regular music practice may be a useful tool to protect against age-related declines in speech encoding in later life (as well as a rewarding pastime). Encoding of speech formant transitions is poorer than normal in the hearing impaired (Plyler and Ananthanarayan, 2001), so musicianship may also help HI individuals as well as the elderly.

Auditory training regimes are yet to be tested for improving or preserving spatial hearing abilities. Do the training regimes described above have the capacity to generalise to spatial hearing benefits? Can spatial hearing be improved through other training regimes? One possible method could be to engage in a demanding spatially selective attention task (e.g., Shinn-Cunningham and Best, 2008; Ruggles *et al.*, 2012; Bharadwaj *et al.*, 2014), which requires trainees to follow competing speech in reverberant conditions, using cognitive functions as well as binaural processing. Musicianship appears to limit the distortion of neural representations of formants and fundamental frequencies by reverberation for young NH adults (Bidelman and Krishnan, 2010). It be may be useful to see if training in reverberant conditions with spatially separated competing sounds protects against the effects of ageing and hearing loss. This could improve the trainee's confidence in the difficult listening conditions they may usually avoid or feel isolated in (Strawbridge *et al.*, 2000).

7.5 Summary

The results reported in this thesis can be summarised as follows:

i. By testing IPD discrimination with a forced-choice task with an alternating pattern in the target interval (AAAA vs. ABAB), little training is needed for listeners to perform as well as they can after hours of practice.

ii. Absolute thresholds are related to a diminished sensitivity to TFS IPDs in AM tones, irrespective of the individual's age.

iii. Low-frequency absolute thresholds relate to how much old HI listeners benefit in speech perception when competing speech is laterally separated. This benefit seems to come mostly from TFS-IPD cues in the competing speech. However, pure-tone IPD discrimination in noise does not predict this benefit.

iv. With increasing age, people become less sensitive to both TFS and envelope IPDs, irrespective of absolute threshold.

v. Measuring FFR to the TFS of AM tones with a horizontal electrode montage appears to emphasise an earlier FFR source than a vertical electrode montage.

vi. With increasing age, the phase-locked neural representation of the TFS of AM tones deteriorates, as does that of the envelope of AM tones with high modulation rates (but not AM tones with low rates).

vii. Although both neural phase locking and IPD discrimination deteriorate with age, the latter is only partly explained by the former. This suggests that other factors, presumably higher-level processing, such as binaural integration, may play a part in age-related difficulties in IPD discrimination.

Both old and HI individuals have deficits in processing IPDs, struggling most with envelope and TFS IPDs respectively. Old individuals also exhibit concurrent declines in neural phase-locking to sounds, and hearing loss also relates to declines in the ability to use temporal spatial cues to improve speech understanding. However, the associations between these deficits and deficits in processing IPDs remain unclear.

References

Abel, S. M., Giguere, C., Consoli, A., and Papsin, B. C. (**2000**). "The effect of aging on horizontal plane sound localization," J. Acoust. Soc. Am. **108**, 743-752.

Aiken, S. J., and Picton, T. W. (**2008**). "Envelope and spectral frequency-following responses to vowel sounds," Hear. Res. **245**, 35-47.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001**). "The CIPIC HRTF database," in *2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 99-102.

Amitay, S., Irwin, A., and Moore, D. R. (**2006**). "Discrimination learning induced by training with identical stimuli," Nat. Neurosci. **9**, 1446-1448.

Andersen, M. R., Kristensen, M. S., Neher, T., and Lunner, T. (**2010**). "Effect of binaural tone vocoding on recognising target speech presented against spatially separated speech maskers," in *International Hearing Aid Research Conference 2010* (Lake Tahoe, California, USA).

Anderson, D. J., Rose, J. E., Hind, J. E., and Brugge, J. F. (**1971**). "Temporal position of discharges in single auditory nerve fibers within the cycle of a sine ‐ wave stimulus: Frequency and intensity effects," J. Acoust. Soc. Am. **49**, 1131-1139.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (**2012**). "Aging affects neural precision of speech encoding," J. Neurosci. **32**, 14156-14164.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (**2013**). "Reversal of age-related neural timing delays with training," Proc. Natl. Acad. Sci. **110**, 4357-4362.

Babkoff, H., Muchnik, C., Ben-David, N., Furst, M., Even-Zohar, S., and Hildesheimer, M. (**2002**). "Mapping lateralization of click trains in younger and older populations," Hear. Res. **165**, 117-127.

Baker, R. J., and Rosen, S. (**2006**). "Auditory filter nonlinearity across frequency using simultaneous notched-noise masking," J. Acoust. Soc. Am. **119**, 454-462.

Ballachanda, B. B., and Moushegian, G. (**2000**). "Frequency-following response: Effects of interaural time and intensity differences," J. Am. Acad. Audiol. **11**, 1-11.

Bartlett, M. S. (**1937**). "Properties of sufficiency and statistical tests," Proc. R. Soc. Math. Phys. Sci. **160**, 268-282.

Baskent, D. (**2006**). "Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels," J. Acoust. Soc. Am. **120**, 2908-2925.

Batra, R., Kuwada, S., and Maher, V. L. (**1986**). "The frequency-following response to continuous tones in humans," Hear. Res. **21**, 167-177.

Bear, M. F., Connors, B. W., and Paradiso, M. A. (**2007**). *Neuroscience, exploring the brain* (Lippincott Williams & Wilkins, Baltimore, MD).

Behrens, T., Neher, T., and Johannesson, R. B. (**2008**). "Evaluation of speech corpus for assessment of spatial release from masking.," in *Auditory signal processing in hearing-impaired listeners*, edited by T. Dau, J. M. Buchholz, J. Harte, and T. U. Christiansen (Centertryk A/S, Copenhagen, Denmark), pp. 449-457.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**, 1026-1036.

Best, V., van Schaik, A., and Carlile, S. (**2004**). "Separation of concurrent broadband sound sources by human listeners," J. Acoust. Soc. Am. **115**, 324-336.

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., and Shinn-Cunningham, B. G. (**2014**). "Cochlear neuropathy and the coding of supra-threshold sound," Front. Syst. Neurosci. **8**.

Bidelman, G. M., and Krishnan, A. (**2010**). "Effects of reverberation on brainstem representation of speech in musicians and non-musicians," Brain Res. **1355**, 112-125.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (**2002**). "Precise inhibition is essential for microsecond interaural time difference coding," Nature **417**, 543-547.

Bremen, P., and Middlebrooks, J. C. (**2013**). "Weighting of spatial and spectro-temporal cues for auditory scene analysis by human listeners," Plos One **8**, e59815.

Briley, P. M., and Summerfield, A. Q. (**2014**). "Age-related deterioration of the representation of space in human auditory cortex," Neurobiol. Aging **35**, 633-644.

British Society of Audiology (**2011**). "Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking: Recommended procedure," (British Society of Audiology, Reading), p. 32.

Bronkhorst, A. W. (**2000**). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," Acta Acust. Utd. Acust. **86**, 117-128.

Bronkhorst, A. W., and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. **83**, 1508-1516.

Bronkhorst, A. W., and Plomp, R. (**1989**). "Binaural speech intelligibility in noise for hearing-impaired listeners," J. Acoust. Soc. Am. **86**, 1374-1383.

Brugge, J. F., Anderson, D. J., Hind, J. E., and Rose, J. E. (**1969**). "Time structure of discharges in single auditory nerve fibres of the squirrel monkey in response to complex periodic sounds," J. Neurophysiol. **32**, 386-401.

Burkard, R. F., Eggermont, J. J., and Don, M. (**2007**). *Auditory evoked potentials: Basic principles and clinical application* (Lippincott Williams & Wilkins).

Buss, E., Hall, J. W., 3rd, and Grose, J. H. (**2004**). "Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss," Ear Hear. **25**, 242-250.

Buus, S., Scharf, B., and Florentine, M. (**1984**). "Lateralization and frequency selectivity in normal and impaired hearing," J. Acoust. Soc. Am. **76**, 77-86.

Byrne, D., and Noble, W. (**1998**). "Optimizing sound localization with hearing aids," Trends Amplif. **3**, 51-73.

Carcagno, S., and Plack, C. (**2011**). "Subcortical plasticity following perceptual learning in a pitch discrimination task," J. Assoc. Res. Otolaryngol. **12**, 89-100.

Cariani, P. A., and Delgutte, B. (**1996**). "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," J. Neurophysiol. **76**, 1698-1716.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., and Colburn, H. S. (**2002**). "Auditory phase opponency: A temporal model for masked detection at low frequencies," Acta Acust. Utd. Acust. **88**, 334-346.

Caspary, D. M., Milbrandt, J. C., and Helfert, R. H. (**1995**). "Central auditory aging: Gaba changes in the inferior colliculus," Exp. Gerontol. **30**, 349-360.

Clark, J. L., Moushegian, G., and Rupert, A. L. (**1997**). "Interaural time effects on the frequency-following response," J. Am. Acad. Audiol. **8**, 308-313.

Clinard, C. G., Tremblay, K. L., and Krishnan, A. R. (**2010**). "Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings," Hear. Res. **264**, 48-55.

Cohen, J. (**1992**). "A power primer," Psychological Bulletin **112**, 155-159.

Colburn, H. S., and Esquissaud, P. (**1976**). "An auditory ‐ nerve model for interaural time discrimination of high ‐ frequency complex stimuli," J. Acoust. Soc. Am. **59**, S23-S23.

Cranford, J. L., Andres, M. A., Piatz, K. K., and Reissig, K. L. (**1993**). "Influences of age and hearing-loss on the precedence effect in sound localization," J. Speech Hear. Res. **36**, 437-441.

Cranford, J. L., Boose, M., and Moore, C. A. (**1990**). "Tests of the precedence effect in sound localization reveal abnormalities in multiple-sclerosis," Ear Hear. **11**, 282-288.

Culling, J. F., Hawley, M. L., and Litovsky, R. Y. (**2004**). "The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources," J. Acoust. Soc. Am. **116**, 1057-1065.

Darwin, C. J. (**2006**). "Contributions of binaural information to the separation of different sound sources," Int. J. Audiol. **45**, 20-24.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers," J. Acoust. Soc. Am. **102**, 2892-2905.

Davis, A. C. (**1989**). "The prevalence of hearing impairment and reported hearing disability among adults in great britain," Int. J. Epidemiol. **18**, 911-917.

Davis, H., Fernández, C., and McAuliffe, D. R. (**1950**). "The excitatory process in the cochlea," Proc. Natl. Acad. Sci. U. S. A. **36**, 580-587.

Davis, R. L., and Britt, R. H. (**1984**). "Analysis of the frequency following response in the cat," Hear. Res. **15**, 29-37.

DESA, U. N. (**2007**). "World population prospects: The 2006 revision, highlights," edited by P. D. Department of Economic and Social Affairs (United Nations, New York).

Devore, S., and Delgutte, B. (**2010**). "Effects of reverberation on the directional sensitivity of auditory neurons across the tonotopic axis: Influences of interaural time and level differences," J. Neurosci. **30**, 7826-7837.

Dietz, M., Marquardt, T., Salminen, N. H., and McAlpine, D. (**2013**). "Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds," Proc. Natl. Acad. Sci. **110**, 15151-15156.

Dietz, M., Marquardt, T., Stange, A., Pecka, M., Grothe, B., and McAlpine, D. (**2014**). "Emphasis of spatial cues in the temporal fine-structure during the rising segments of amplitude modulated sounds ii: Single neuron recordings," J. Neurophysiol.

Divenyi, P. L., Stark, P. B., and Haupt, K. M. (**2005**). "Decline of speech understanding and auditory thresholds in the elderly," J. Acoust. Soc. Am. **118**, 1089-1100.

Dobie, R. A., and Wilson, M. J. (**1989**). "Analysis of auditory evoked potentials by magnitude-squared coherence," Ear Hear. **10**, 2-13.

Dobreva, M. S., O'Neill, W. E., and Paige, G. D. (**2011**). "Influence of aging on human sound localization," J. Neurophysiol. **105**, 2471-2486.

Dolphin, W. F. (**1996**). "Auditory evoked responses to amplitude modulated stimuli consisting of multiple envelope components," J. Comp. Physiol. A **179**, 113-121.

Dolphin, W. F. (**1997**). "The envelope following response to multiple tone pair stimuli," Hear. Res. **110**, 1-14.

Dreschler, W. A. (**1983**). "Impaired frequency/time resolution and its effect on speech intelligibility," in *Hearing - physiological bases and psychophysics,* edited by R. Klinke, and R. Hartmann (Springer-Verlag, Berlin).

Dreschler, W. A., and Plomp, R. (**1980**). "Relations between psychophysical data and speech perception for hearing-impaired subjects. I," J. Acoust. Soc. Am. **68**, 1608-1615.

Dreschler, W. A., and Plomp, R. (**1985**). "Relations between psychophysical data and speech perception for hearing-impaired subjects. Ii," J. Acoust. Soc. Am. **78**, 1261-1270.

Du, Y., Kong, L., Wang, Q., Wu, X., and Li, L. (**2011**). "Auditory frequency-following response: A neurophysiological measure for studying the "cocktail-party problem"," Neurosci. Biobehav. Rev. **35**, 2046-2057.

Dubno, J. R., Eckert, M. A., Lee, F.-S., Matthews, L. J., and Schmiedt, R. A. (**2013**). "Classifying human audiometric phenotypes of age-related hearing loss from animal models," J. Assoc. Res. Otolaryngol. **14**, 687-701.

Dudley, H. (**1939**). "Remaking speech," J. Acoust. Soc. Am. **11**, 169-177.

Duquesnoy, A. J. (**1983**). "Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. **74**, 739-743.

Eddins, D. A., and Hall, J. W. I. (**2010**). "Binaural processing and auditory asymmetries," in *The aging auditory system*, edited by S. Gordon-Salant, Frisina, R. D., Popper, A. N. and Fay, R. R. (Springer, New York).

Edmonds, B. A., and Culling, J. F. (**2005**). "The spatial unmasking of speech: Evidence for within-channel processing of interaural time delay," J. Acoust. Soc. Am. **117**, 3069-3078.

Elberling, C., and Parbo, J. (**1987**). "Reference data for abrs in retrocochlear diagnosis," Scand. Audiol. **16**, 49-55.

Felder, E., and Schrott-Fischer, A. (**1995**). "Quantitative evaluation of myelinated nerve fibres and hair cells in cochleae of humans with age-related high-tone hearing loss," Hear. Res. **91**, 19-32.

Freigang, C., Schmiedchen, K., Nitsche, I., and Rübsamen, R. (**2014**). "Free-field study on auditory localization and discrimination performance in older adults," Exp. Brain Res. **232**, 1157-1172.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2001**). "Spatial release from informational masking in speech recognition," J. Acoust. Soc. Am. **109**, 2112-2122.

Freyman, R. L., and Nelson, D. A. (**1991**). "Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **34**, 1371-1386.

Frisina, D. R., and Frisina, R. D. (**1997**). "Speech recognition in noise and presbycusis: Relations to possible neural mechanisms," Hear. Res. **106**, 95-104.

Frisina, R. D., and Walton, J. P. (**2006**). "Age-related structural and functional changes in the cochlear nucleus," Hear. Res. **216–217**, 216-223.

Füllgrabe, C. (**2013**). "Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss," Am. J. Audiol. **22**, 313-315.

Füllgrabe, C., Berthommier, F., and Lorenzi, C. (**2006**). "Masking release for consonant features in temporally fluctuating background noise," Hear. Res. **211**, 74-84.

Furman, A. C., Kujawa, S. G., and Liberman, M. C. (**2013**). "Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates," J. Neurophysiol. **110**, 577-586.

Furst, M., Levine, R. A., and McGaffigan, P. M. (**1985**). "Click lateralization is related to the β component of the dichotic brainstem auditory evoked potentials of human subjects," J. Acoust. Soc. Am. **78**, 1644-1651.

Galambos, R., and Davis, H. (**1943**). "The response of single auditory-nerve fibers to acoustic stimulation," J. Neurophysiol. **6**, 39-57.

Galbraith, G. C. (**1994**). "Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli," Electroen. Clin. Neuro. **92**, 321-330.

Galbraith, G. C., Bagasan, B., and Sulahian, J. (**2001**). "Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations," Percept. Mot. Skills **92**, 99-106.

Galbraith, G. C., Threadgill, M. R., Hemsley, J., Salour, K., Songdej, N., Ton, J., and Cheung, L. (**2000**). "Putative measure of peripheral and brainstem frequency-following in humans," Neurosci. Lett. **292**, 123-127.

Gardi, J., Merzenich, M., and McKean, C. (**1979**). "Origins of the scalp-recorded frequency-following response in the cat," Int. J. Audiol. **18**, 353-380.

Gatehouse, R. W., and Noble, W. (**2004**). "The speech, spatial and qualities of hearing scale (ssq)," Int. J. Audiol. **43**, 85-99.

Gelfand, S. A., Ross, L., and Miller, S. (**1988**). "Sentence reception in noise from one versus two sources: Effects of aging and hearing loss," J. Acoust. Soc. Am. **83**, 248-256.

Gerken, G. M., Moushegian, G., Stillman, R. D., and Rupert, A. L. (**1975**). "Human frequency-following responses to monaural and binaural stimuli," Electroen. Clin. Neuro. **38**, 379-386.

Gilchrist, A. L., Cowan, N., and Naveh-Benjamin, M. (**2008**). "Working memory capacity for spoken sentences decreases with adult ageing: Recall of fewer but not smaller chunks in older adults," Memory **16**, 773-787.

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

Glasberg, B. R., and Moore, B. C. J. (**1992**). "Effects of envelope fluctuations on gap detection," Hear. Res. **64**, 81-92.

Glaser, E. M., Suter, C. M., Dasheiff, R., and Goldberg, A. (**1976**). "The human frequency-following response: Its behavior during continuous tone and tone burst stimulation," Electroen. Clin. Neuro. **40**, 25-32.

Gleeson, M., and Felix, H. (**1987**). "A comparative study of the effect of age on the human cochlear and vestibular neuroepithelia," Acta Otolaryngol. **104**, 103-109.

Goblick, T. J., and Pfeiffer, R. R. (**1969**). "Time‐domain measurements of cochlear nonlinearities using combination click stimuli," J. Acoust. Soc. Am. **46**, 924-938.

Gockel, H. E., Carlyon, R. P., Mehta, A., and Plack, C. J. (**2011**). "The frequency following response (ffr) may reflect pitch-bearing information but is not a direct representation of pitch," J. Assoc. Res. Otolaryngol. **12**, 767-782.

Gordon-Salant, S., Frisina, R. D., Popper, A. N., and Fay, R. R. (eds). (**2010**). *The aging auditory system* (Springer, New York).

Goverts, S. T., and Houtgast, T. (**2010**). "The binaural intelligibility level difference in hearing-impaired listeners: The role of supra-threshold deficits," J. Acoust. Soc. Am. **127**, 3073-3084.

Green, D. M., and Swets, J. A. (**1974**). *Signal detection theory and psychophysics* (Krieger, New York), pp. 53-85, 154-161.

Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (**2006**). "The master hearing aid: A pc-based platform for algorithm development and evaluation," Acta Acust. Utd. Acust. **92**, 618-628.

Grose, J. H., Hall, J. W., 3rd, and Buss, E. (**2006**). "Temporal processing deficits in the pre-senescent auditory system," J. Acoust. Soc. Am. **119**, 2305-2315.

Grose, J. H., and Mamo, S. K. (**2010**). "Processing of temporal fine structure as a function of age," Ear Hear. **31**, 755-760.

Grose, J. H., Mamo, S. K., and Hall III, J. W. (**2009a**). "Age effects in temporal envelope processing: Speech unmasking and auditory steady state responses," Ear Hear. **30**, 568.

Grose, J. H., Mamo, S. K., and Hall, J. W., 3rd (**2009b**). "Age effects in temporal envelope processing: Speech unmasking and auditory steady state responses," Ear Hear. **30**, 568-575.

Grose, J. H., Poth, E. A., and Peters, R. W. (**1994**). "Masking level differences for tones and speech in elderly listeners with relatively normal audiograms," J. Speech Hear. Res. **37**, 422-428.

Guinan, J. J. J. (**2006**). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," Ear Hear. **27**, 589-607.

Hacker, M. J., and Ratcliff, R. (**1979**). "A revised table of d' for m-alternative forced choice," Percept. Psychophys. **26**, 168-170.

Hafter, E. R., and Carrier, S. C. (**1970**). "Masking-level differences obtained with a pulsed tonal masker," J. Acoust. Soc. Am. **47**, 1041-1047.

Hafter, E. R., and Carrier, S. C. (**1972**). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. **51**, 1852-1862.

Hall, J. W., Tyler, R. S., and Fernandes, M. A. (**1984**). "Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners," J. Speech Hear. Res. **27**, 145-154.

Harrison, R. V., and Evans, E. F. (**1979**). "Cochlear fibre responses in guinea pigs with well defined cochlear lesions," Scand. Audiol. Suppl. **9**, 83-92.

Hawkey, D. J., Amitay, S., and Moore, D. R. (**2004**). "Early and rapid perceptual learning," Nat. Neurosci. **7**, 1055-1056.

Hawkins, D. B., and Wightman, F. L. (**1980**). "Interaural time discrimination ability of listeners with sensori-neural hearing loss," Audiology **19**, 495-507.

Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (**1999**). "Speech intelligibility and localization in a multi-source environment," J. Acoust. Soc. Am. **105**, 3436-3448.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (**2004**). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**, 833-843.

He, N.-j., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (**2008**). "Age-related differences in the temporal modulation transfer function with pure-tone carriers," J. Acoust. Soc. Am. **124**, 3841-3849.

Heinz, M. G. (**2010**). "Across-fiber coding of temporal fine-structure: Effects of noise-induced hearing loss on auditory-nerve responses," in *Advances in Auditory Physiology, Psychophysics and Models,* edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 621-630.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neural Comput. **13**, 2273-2316.

Helfer, K. S., and Freyman, R. L. (**2008**). "Aging and speech-on-speech masking," Ear Hear. **29**, 87.

Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. **55**, 84-90.

Henry, K. S., and Heinz, M. G. (**2012**). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," Nat. Neurosci. **15**, 1362-1364.

Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., and Picton, T. W. (**2002**). "Intracerebral sources of human auditory steady-state responses," Brain Topogr. **15**, 69-86.

Herman, G. E., Warren, L. R., and Wagener, J. W. (**1977**). "Auditory lateralization: Age differences in sensitivity to dichotic time and amplitude cues," J. Gerontol. **32**, 187-191.

Hinchcliffe, R. (**1959**). "The threshold of hearing as a function of age," Acustica **9**, 303-308.

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

Hirsh, I. J. (**1950**). "The relation between localization and intelligibility," J. Acoust. Soc. Am. **22**, 196-200.

Holm, S. (**1979**). "A simple sequentially rejective multiple test procedure," Scand. J. Stat. **6**, 65-70.

Hopkins, K., and Moore, B. C. J. (**2007**). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," J. Acoust. Soc. Am. **122**, 1055-1068.

Hopkins, K., and Moore, B. C. J. (**2009**). "The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise," J. Acoust. Soc. Am. **125**, 442-446.

Hopkins, K., and Moore, B. C. J. (**2010a**). "Development of a fast method of measuring sensitivity to temporal fine structure information at low frequencies *in* abstracts of the british society of audiology short papers meeting on experimental studies of hearing and deafness," Int. J. Audiol. **49**, 141-183.

Hopkins, K., and Moore, B. C. J. (**2010b**). "Development of a fast method of measuring sensitivity to temporal fine structure information at low frequencies.," Int. J. Audiol. **49**, 940-946.

Hopkins, K., and Moore, B. C. J. (**2010c**). "The importance of temporal fine structure information in speech at different spectral regions," J. Acoust. Soc. Am. **127**, 1595-1608.

Hopkins, K., and Moore, B. C. J. (**2011**). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," J. Acoust. Soc. Am. **130**, 334-349.

Hopkins, K., Moore, B. C. J., and Stone, M. A. (**2008**). "The effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," J. Acoust. Soc. Am. **123**, 1140-1153.

Huis in't Veld, F., Osterhammel, P., and Terkildsen, K. (**1977**). "Frequency following auditory brain stem responses in man," Scand. Audiol. **6**, 27-34.

Huynh, H., and Feldt, L. S. (**1976**). "Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs," Journal of Educational and Behavioral Statistics **1**, 69-82.

Irwin, R. J., Hinchcliff, L. K., and Kemp, S. (**1981**). "Temporal acuity in normal and hearing-impaired listeners," Audiol. **20**, 234-243.

Jeffress, L. A. (**1948**). "A place theory of sound localization," J. Comp. Physiol. Psychol. **41**, 35-39.

Jerger, J., and Johnson, K. (**1988**). "Interactions of age, gender, and sensorineural hearing loss on abr latency," Ear Hear. **9**, 168-176.

Jewett, D. L., and Williston, J. S. (**1971**). "Auditory-evoked far fields averaged from the scalp of humans," Brain **94**, 681-696.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994a**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," J. Neurophysiol. **71**, 1022-1036.

Joris, P. X., Schreiner, C. E., and Rees, A. (**2004**). "Neural processing of amplitude-modulated sounds," Physiol. Rev. **84**, 541-577.

Joris, P. X., and Smith, P. H. (**2008**). "The volley theory and the spherical cell puzzle," Neuroscience **154**, 65-76.

Joris, P. X., Smith, P. H., and Yin, T. C. (**1994b**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. Ii. Responses in the tuning curve tail," J. Neurophysiol. **71**, 1037-1051.

Kale, S., and Heinz, M. G. (**2010**). "Envelope coding in auditory nerve fibers following noise-induced hearing loss," J. Assoc. Res. Otolaryngol. **11**, 657-673.

Kates, J. M. (**2011**). "Spectro-temporal envelope changes caused by temporal fine structure modification," J. Acoust. Soc. Am. **129**, 3981-3990.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). *Discharge patterns of single fibers in the cat's auditory nerve* (MIT Press, Cambridge, Mass.).

King, A., Hopkins, K., and Plack, C. J. (**2014**). "The effects of age and hearing loss on interaural phase difference discriminationa)," J. Acoust. Soc. Am. **135**, 342-351.

Klug, A., Bauer, E. E., Hanson, J. T., Hurley, L., Meitzen, J., and Pollak, G. D. (**2002**). "Response selectivity for species-specific calls in the inferior colliculus of mexican free-tailed bats is generated by inhibition," J. Neurophysiol. **88**, 1941-1954.

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

Koch, U., and Grothe, B. (**1998**). "Gabaergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat," J. Neurophysiol. **80**, 71-82.

Kramer, S. E., Kapteyn, T. S., and Festen, J. M. (**1998**). "The self-reported handicapping effect of hearing disabilities," Int. J. Audiol. **37**, 302-312.

Krishnan, A. (**2002**). "Human frequency-following responses: Representation of steady-state synthetic vowels," Hear. Res. **166**, 192-201.

Krishnan, A. (**2007**). "Frequency-following response," in *Auditory evoked potentials: Basic principles and clinical application*, edited by R. F. Burkard, M. Don, and J. J. Eggermont (Lippincott Williams and Wilkins, Baltimore, MD), pp. 313-333.

Krishnan, A., and McDaniel, S. S. (**1998**). "Binaural interaction in the human frequency-following response: Effects of interaural intensity difference," Audiol. Neurootol. **3**, 291-299.

Krishnan, A., and Plack, C. J. (**2011**). "Neural encoding in the human brainstem relevant to the pitch of complex tones," Hear. Res. **275**, 110-119.

Krishnan, A., Xu, Y., Gandour, J. T., and Cariani, P. A. (**2004**). "Human frequency-following response: Representation of pitch contours in chinese tones," Hear. Res. **189**, 1-12.

Krishnan, A., Xu, Y., Gandour, J. T., and Cariani, P. A. (**2005**). "Encoding of pitch in the human brainstem is sensitive to language experience," Cognitive Brain. Res. **25**, 161-168.

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

Kujawa, S. G., and Liberman, M. C. (**2006**). "Acceleration of age-related hearing loss by early noise exposure: Evidence of a misspent youth," J. Neurosci. **26**, 2115-2123.

Kujawa, S. G., and Liberman, M. C. (**2009**). "Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss," J. Neurosci. **29**, 14077-14085.

Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., and D'Angelo, W. R. (**2002**). "Sources of the scalp-recorded amplitude-modulation following response," J. Am. Acad. Audiol. **13**, 188-204.

Lacher-Fougère, S., and Demany, L. (**1998**). "Modulation detection by normal and hearing-impaired listeners," Audiology **37**, 109-121.

Lacher-Fougère, S., and Demany, L. (**2005**). "Consequences of cochlear damage for the detection of interaural phase differences," J. Acoust. Soc. Am. **118**, 2519-2526.

Lee, F. S., Matthews, L. J., Dubno, J. R., and Mills, J. H. (**2005**). "Longitudinal study of pure-tone thresholds in older persons," Ear Hear. **26**, 1-11.

Leigh-Paffenroth, E. D., and Fowler, C. G. (**2006**). "Amplitude-modulated auditory steady-state responses in younger and older listeners," J. Am. Acad. Audiol. **17**, 582-597.

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

Levitt, H., and Rabiner, L. R. (**1967**). "Binaural release from masking for speech and gain in intelligibility," J. Acoust. Soc. Am. **42**, 601-608.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442-455.

Liberman, M. C. (**1982**). "Single-neuron labeling in the cat auditory-nerve," Science **216**, 1239-1241.

Liberman, M. C., Dodds, L. W., and Learson, D. A. (**1986**). "Structure-function correlation in noise-damaged ears: A light and electron-microscopic study," in *Basic and applied aspects of noise-induced hearing loss*, edited by R. J. Salvi, D. Henderson, R. P. Hamernik, and V. Colletti (Plenum, New York), pp. 163-176.

Liberman, M. C., and Kiang, N. Y. S. (**1984**). "Single-neuron labeling and chronic cochlear pathology. Iv. Stereocilia damage and alterations in rate- and phase-level functions," Hear. Res. **16**, 75-90.

Licklider, J. C. R. (**1948**). "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. **20**, 150-159.

Lin, H. W., Furman, A. C., Kujawa, S. G., and Liberman, M. C. (**2011**). "Primary neural degeneration in the guinea pig cochlea after reversible noise-induced threshold shift," J. Assoc. Res. Otolaryngol. **12**, 605-616.

Lins, O. G., Picton, T. W., Boucher, B. L., Durieux-Smith, A., Champagne, S. C., Moran, L. M., Perez-Abalo, M. C., Martin, V., and Savio, G. (**1996**). "Frequency-specific audiometry using steady-state responses," Ear Hear. **17**, 81-96.

Lins, O. G., Picton, T. W., Picton, P. E., Champagne, S. C., and Durieux‐Smith, A. (**1994**). "Brainstem auditory steady‐state responses to multiple simultaneous stimuli," J. Acoust. Soc. Am. **95**, 2842-2843.

Liu, L.-F., Palmer, A. R., and Wallace, M. N. (**2006**). "Phase-locked responses to pure tones in the inferior colliculus," J. Neurophysiol. **95**, 1926-1935.

Loizou, P. C., Dorman, M., and Tu, Z. (**1999**). "On the number of channels needed to understand speech," J. Acoust. Soc. Am. **106**, 2097-2103.

Lopez-Poveda, E. A., and Barrios, P. (**2013**). "Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation," Front. Neurosci. **7**.

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., and Moore, B. C. (**2009**). "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal," J. Acoust. Soc. Am. **125**, 27-30.

Lorenzi, C., Gatehouse, S., and Lever, C. (**1999a**). "Sound localization in noise in hearing-impaired listeners," J. Acoust. Soc. Am. **105**, 3454-3463.

Lorenzi, C., Gatehouse, S., and Lever, C. (**1999b**). "Sound localization in noise in normal-hearing listeners," J. Acoust. Soc. Am. **105**, 1810-1820.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (**2006**). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Proc. Natl. Acad. Sci. U. S. A. **103**, 18866-18869.

Lunner, T., Hietkamp, R. K., Andersen, M. R., Hopkins, K., and Moore, B. C. J. (**2012**). "Effect of speech material on the benefit of temporal fine structure information in speech for young normal-hearing and older hearing-impaired participants," Ear Hear. **33**, 377-388.

MacLeod, A., and Summerfield, Q. (**1990**). "A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use," Br. J. Audiol. **24**, 29-43.

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," J. Acoust. Soc. Am. **111**, 2219-2236.

Magezi, D. A., and Krumbholz, K. (**2010**). "Evidence for opponent-channel coding of interaural time differences in human auditory cortex," J. Physiol. **104**, 1997-2007.

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., and Merchant, S. N. (**2011**). "Age-related primary cochlear neuronal degeneration in human temporal bones," J. Assoc. Res. Otolaryngol. **12**, 711-717.

Marmel, F., Linley, D., Carlyon, R. P., Gockel, H. E., Hopkins, K., and Plack, C. J. (**2013**). "Subcortical neural synchrony and absolute thresholds predict frequency discrimination independently," J. Assoc. Res. Otolaryngol. **14**, 757-766.

Marrone, N., Mason, C. R., and Gerald Kidd, J. (**2008a**). "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," J. Acoust. Soc. Am. **124**, 3064-3075.

Marrone, N., Mason, C. R., and Kidd, G. (**2008b**). "Tuning in the spatial dimension: Evidence from a masked speech identification task," J. Acoust. Soc. Am. **124**, 1146-1158.

Marrone, N., Mason, C. R., and Kidd, G., Jr. (**2008c**). "Evaluating the benefit of hearing aids in solving the cocktail party problem.," Trends Amplif. **12**, 300-315.

Marsh, J. T., Brown, W. S., and Smith, J. C. (**1975**). "Far-field recorded frequency-following responses: Correlates of low pitch auditory perception in humans responses d'entertainement enregistrees a distance: Correlates de la perception auditive des sons de faible hauteur chez l'homme," Electroen. Clin. Neuro. **38**, 113-119.

Micheyl, C., Dai, H., and Oxenham, A. J. (**2010**). "On the possible influence of spectral and temporal-envelope cues in tests of sensitivity to temporal fine structure," J. Acoust. Soc. Am. **127**, 1809-1809.

Micheyl, C., Moore, B. C. J., and Carlyon, R. P. (**1998**). "The role of excitation-pattern cues and temporal cues in the frequency and modulation-rate discrimination of amplitude-modulated tones," J. Acoust. Soc. Am. **104**, 1039-1050.

Middlebrooks, J. C., and Green, D. M. (**1990**). "Directional dependence of interaural envelope delays," J. Acoust. Soc. Am. **87**, 2149-2162.

Middlebrooks, J. C., and Green, D. M. (**1991**). "Sound localization by human listeners," Ann. Rev. Psychol. **42**, 135-159.

Milbrandt, J. C., Albin, R. L., and Caspary, D. M. C. (**1994**). "Age-related decrease in gabab receptor binding in the fischer 344 rat inferior colliculus," Neurobiol. Aging **15**, 699-703.

Miller, R. L., Schilling, J. R., Franck, K. R., and Young, E. D. (**1997**). "Effects of acoustic trauma on the representation of the vowel /ε/ in cat auditory nerve fibers," J. Acoust. Soc. Am. **101**, 3602-3616.

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

Møller, A. R. (**1999**). "Review of the roles of temporal and place coding of frequency in speech discrimination," Acta Otolaryngol. **119**, 424-430.

Møller, A. R. (**2007**). "Neural generators for auditory brainstem evoked potentials (abr)," in *Auditory evoked potentials: Basic principles and clinical application*, edited by R. F. Burkard, M. Don, and J. J. Eggermont (Lippincott Williams and Wilkins, Baltimore, MD), pp. 336-354.

Møller, A. R., and Jannetta, P. J. (**1981**). "Compound action potentials recorded intracranially from the auditory nerve in man," Exp. Neurol. **74**, 862-874.

Møller, A. R., and Jannetta, P. J. (**1982**). "Evoked potentials from the inferior colliculus in man," Electroen. Clin. Neuro. **53**, 612-620.

Møller, A. R., and Jannetta, P. J. (**1983**). "Auditory evoked potentials recorded from the cochlear nucleus and its vicinity in man," J. Neurosurg. **59**, 1013-1018.

Møller, A. R., Jannetta, P. J., and Jho, H. D. (**1994**). "Click-evoked responses from the cochlear nucleus: A study in human," Electroen. Clin. Neuro. **92**, 215-224.

Monaghan, J. J. M., Krumbholz, K., and Seeber, B. U. (**2013**). "Factors affecting the use of envelope interaural time differences in reverberationa)," J. Acoust. Soc. Am. **133**, 2288-2300.

Moore, B. C. J. (**1973a**). "Frequency difference limens for narrow bands of noise," J. Acoust. Soc. Am. **54**, 888-896.

Moore, B. C. J. (**1973b**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610-619.

Moore, B. C. J. (**2007**). *Cochlear hearing loss: Physiological, psychological and technical issues, second edition* (Wiley, Chichester), pp. 79-83.

Moore, B. C. J. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," J. Assoc. Res. Otolaryngol. **9**, 399-406.

Moore, B. C. J. (**2012**). "The importance of temporal fine structure for the intelligibility of speech in complex backgrounds.," in *Speech perception and auditory disorders*, edited by T. Dau, M. L. Jepsen, T. Poulsen, and J. C. Dalsgaard (The Danavox Jubilee Foundation, Ballerup, Denmark), pp. 21-32.

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., and Glasberg, B. R. (**1987a**). "Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes," J. Acoust. Soc. Am. **82**, 69-79.

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

Moore, B. C. J., and Glasberg, B. R. (**1993**). "Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech," J. Acoust. Soc. Am. **94**, 2050-2062.

Moore, B. C. J., and Glasberg, B. R. (**1998**). "Use of a loudness model for hearing-aid fitting. I. Linear hearing aids," Br. J. Audiol. **32**, 317-335.

Moore, B. C. J., Glasberg, B. R., Flanagan, H. J., and Adams, J. (**2006a**). "Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure," J. Acoust. Soc. Am. **119**, 480-490.

Moore, B. C. J., Glasberg, B. R., and Hopkins, K. (**2006b**). "Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure," Hear. Res. **222**, 16-27.

Moore, B. C. J., Glasberg, B. R., Stoev, M., Fullgrabe, C., and Hopkins, K. (**2012a**). "The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (l)," J. Acoust. Soc. Am. **131**, 1003-1006.

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántara, J. I. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205-224.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1992**). "Detection of temporal gaps in sinusoids by elderly subjects with and without hearing loss," J. Acoust. Soc. Am. **92**, 1923-1932.

Moore, B. C. J., and Sek, A. (**1995**). "Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (am vs fm)," J. Acoust. Soc. Am. **97**, 2468-2478.

Moore, B. C. J., and Sek, A. (**1996**). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," J. Acoust. Soc. Am. **100**, 2320-2331.

Moore, B. C. J., and Sek, A. (**2009a**). "Development of a fast method for determining sensitivity to temporal fine structure," Int. J. Audiol. **48**, 161-171.

Moore, B. C. J., and Sek, A. (**2009b**). "Sensitivity of the human auditory system to temporal fine structure at high frequencies," J. Acoust. Soc. Am. **125**, 3186-3193.

Moore, B. C. J., and Skrodzka, E. (**2002**). "Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation," J. Acoust. Soc. Am. **111**, 327-335.

Moore, B. C. J., Vickers, D. A., and Mehta, A. (**2012b**). "The effects of age on temporal fine structure sensitivity in monaural and binaural conditions," Int. J. Audiol. **51**, 715-721.

Moore, B. C. J., Wojtczak, M., and Vickers, D. A. (**1996**). "Effect of loudness recruitment on the perception of amplitude modulation," J. Acoust. Soc. Am. **100**, 481-489.

Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., and Fozard, J. L. (**1996**). "Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level," J. Acoust. Soc. Am. **100**, 1949-1967.

Moushegian, G., Rupert, A. L., and Stillman, R. D. (**1973**). "Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range," Electroen. Clin. Neuro. **35**, 665-667.

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

Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Petersen, A. S., and Schaik, A. (**2009**). "Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition," Int. J. Audiol. **48**, 758-774.

Neher, T., Laugesen, S., Søgaard Jensen, N., and Kragelund, L. (**2011**). "Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities?," J. Acoust. Soc. Am. **130**, 1542-1558.

Neher, T., Lunner, T., Hopkins, K., and Moore, B. C. J. (**2012**). "Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (l)," J. Acoust. Soc. Am. **131**, 2561-2564.

Nielsen, J. B., and Dau, T. (**2009**). "Development of a danish speech intelligibility test," Int. J. Audiol. **48**, 729-741.

Noble, W., Byrne, D., and Lepage, B. (**1994**). "Effects on sound localization of configuration and type of hearing impairment," J. Acoust. Soc. Am. **95**, 992-1005.

Oxenham, A. J., Micheyl, C., and Keebler, M. V. (**2009**). "Can temporal fine structure represent the fundamental frequency of unresolved harmonics?," J. Acoust. Soc. Am. **125**, 2189-2199.

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

Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (**2012**). "Musical experience offsets age-related delays in neural timing," Neurobiol. Aging **33**, 1483e1481-1483e1484.

Parbery-Clark, A., Skoe, E., and Kraus, N. (**2009**). "Musical experience limits the degradative effects of background noise on the neural processing of sound," J. Neurosci. **29**, 14100-14107.

Parthasarathy, A., and Bartlett, E. (**2012**). "Two-channel recording of auditory-evoked potentials to detect age-related deficits in temporal processing," Hear. Res. **289**, 52-62.

Parthasarathy, A., Datta, J., Torres, J. A. L., Hopkins, C., and Bartlett, E. L. (**2014**). "Age-related changes in the relationship between auditory brainstem responses and envelope-following responses," J. Assoc. Res. Otolaryngol. **15**, 649-661.

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

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788-1803.

Peissig, J., and Kollmeier, J. (**1997**). "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," J. Acoust. Soc. Am. **101**, 1660-1670.

Peters, A. (**2002**). "The effects of normal aging on myelin and nerve fibers: A review," J. Neurocytol. **31**, 581-593.

Peters, A., and Sethares, C. (**2003**). "Is there remyelination during aging of the primate central nervous system?," J. Comp. Neurol. **460**, 238-254.

Pichora-Fuller, M. K., and Schneider, B. A. (**1992**). "The effect of interaural delay of the masker on masking-level differences in young and old subjects," J. Acoust. Soc. Am. **91**, 2129-2135.

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593-608.

Pichora-Fuller, M. K., Schneider, B. A., Macdonald, E., Pass, H. E., and Brown, S. (**2007**). "Temporal jitter disrupts speech intelligibility: A simulation of auditory aging," Hear. Res. **223**, 114-121.

Picton, T. W., Hillyard, S. A., Krausz, H. I., and Galambos, R. (**1974**). "Human auditory evoked potentials. I: Evaluation of components," Electroen. Clin. Neuro. **36**, 179-190.

Plomp, R. (**1978**). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am. **63**, 533-549.

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

Plyler, P. N., and Ananthanarayan, A. K. (**2001**). "Human frequency-following responses: Representation of second formant transitions in normal-hearing and hearing-impaired listeners," J. Am. Acad. Audiol. **12**, 523-533.

Purcell, D. W., John, S. M., Schneider, B. A., and Picton, T. W. (**2004**). "Human temporal auditory acuity as assessed by envelope following responses," J. Acoust. Soc. Am. **116**, 3581-3593.

Qin, M. K., and Oxenham, A. J. (**2003**). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," J. Acoust. Soc. Am. **114**, 446-454.

Rakerd, B., and Hartmann, W. M. (**2010**). "Localization of sound in rooms. V. Binaural coherence and human sensitivity to interaural time differences in noise," J. Acoust. Soc. Am. **128**, 3052-3063.

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

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305-1352.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769-793.

Rose, J. E., Gross, N. B., Geisler, C. D., and Hind, J. E. (**1966**). "Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source," J. Neurophysiol. **29**, 288-314.

Rose, J. E., Hind, J. E., Anderson, D. J., and Brugge, J. F. (**1971**). "Some effects of the stimulus intensity on response of auditory nerve fibers in the squirrel monkey," J. Neurophysiol. **34**, 685-699.

Ross, B. (**2008**). "A novel type of auditory responses: Temporal dynamics of 40-hz steady-state responses induced by changes in sound localization," J. Neurophysiol. **100**, 1265-1277.

Ross, B., Fujioka, T., Tremblay, K. L., and Picton, T. W. (**2007**). "Aging in binaural hearing begins in mid-life: Evidence from cortical auditory-evoked responses to changes in interaural phase," J. Neurosci. **27**, 11172-11178.

Rowan, D., and Lutman, M. E. (**2006**). "Learning to discriminate interaural time differences: An exploratory study with amplitude-modulated stimuli," Int. J. Audiol. **45**, 513-520.

Rowe III, M. J. (**1978**). "Normal variability of the brain-stem auditory evoked response in young and old adult subjects," Electroen. Clin. Neuro. **44**, 459-470.

Ruggero, M. A. (**1994**). "Cochlear delays and traveling waves: Comments on 'experimental look at cochlear mechanics'," Audiology **33**, 131-142.

Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (**2012**). "Why middle-aged listeners have trouble hearing in everyday settings," Curr. Biol. **22**, 1417-1422.

Ruggles, D., and Shinn-Cunningham, B. (**2011**). "Spatial selective auditory attention in the presence of reverberant energy: Individual differences in normal-hearing listeners," J. Assoc. Res. Otolaryngol. **12**, 395-405.

Sachs, M. B., F., V. H., and D., Y. E. (**1983**). "Auditory nerve representation of vowels in background noise," J. Neurophysiol. **50**, 27-45.

Sachs, M. B., and Young, E. D. (**1979**). "Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate," J. Acoust. Soc. Am. **66**, 470-479.

Schatteman, T. A., Hughes, L. F., and Caspary, D. M. (**2008**). "Aged-related loss of temporal processing: Altered responses to amplitude modulated tones in rat dorsal cochlear nucleus," Neuroscience **154**, 329-337.

Scherg, M., and Brinkmann, R. D. (**1979**). "Least-square-fit technique applied to the frequency following potential: A method to determine components, latencies and amplitudes," Scand. Audiol. Suppl., 197-203.

Scherg, M., and Von Cramon, D. (**1985**). "A new interpretation of the generators of baep waves i–v: Results of a spatio-temporal dipole model," Electroen. Clin. Neuro. **62**, 290-299.

Schmiedt, R. A. (**1996**). "Effects of aging on potassium homeostasis and the endocochlear potential in the gerbil cochlea," Hear. Res. **102**, 125-132.

Schmiedt, R. A. (**2010**). "The physiology of cochlear presbycusis," in *The aging auditory system*, edited by S. Gordon-Salant, Frisina, R. D., Popper, A. N. and Fay, R. R. (Springer, New York).

Schmiedt, R. A., H., M. J., and A., B. F. (**1996**). "Age-related loss of activity of auditory-nerve fibers," J. Neurophysiol. **76**, 2799-2803.

Schneider, B. (**1997**). "Psychoacoustics and aging: Implications for everyday listening," J. Speech-Lang. Path. Aud. **21**, 111-124.

Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., and Lamb, M. (**1994**). "Gap detection and the precedence effect in young and old adults," J. Acoust. Soc. Am. **95**, 980-991.

Schrott, A., Stephan, K., and Spoendlin, H. (**1989**). "Hearing with selective inner hair cell loss," Hear. Res. **40**, 213-219.

Schubert, E. D., and Schultz, M. C. (**1962**). "Some aspects of binaural signal selection," J. Acoust. Soc. Am. **34**, 844-849.

Schuknecht, H. F. (**1955**). "Presbycusis," Laryngoscope **65**, 402-419.

Schuknecht, H. F. (**1964**). "Further observations on the pathology of presbycusis," Arch. Otolaryngol. **80**, 369-382.

Schuknecht, H. F., and Woellner, R. C. (**1955**). "An experimental and clinical study of deafness from lesions of the cochlear nerve," J. Laryngol. Otol. **69**, 75-97.

Sek, A., and Moore, B. C. J. (**1995**). "Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am. **97**, 2479-2486.

Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (**2013**). "Age-related cochlear synaptopathy: An early-onset contributor to auditory functional decline," J. Neurosci. **33**, 13686-13694.

Shamma, S., and Lorenzi, C. (**2013**). "On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system," J. Acoust. Soc. Am. **133**, 2818-2833.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303-304.

Shinn-Cunningham, B. G., and Best, V. (**2008**). "Selective attention in normal and impaired hearing," Trends Amplif. **12**, 283-299.

Smith, J. C., Marsh, J. T., and Brown, W. S. (**1975**). "Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources," Electroen. Clin. Neuro. **39**, 465-472.

Smoski, W. J., and Trahiotis, C. (**1986**). "Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss," J. Acoust. Soc. Am. **79**, 1541-1547.

Snell, K. B. (**1997**). "Age related changes in temporal gap detection," J. Acoust. Soc. Am. **101**, 2214-2220.

Sohmer, H., and Pratt, H. (**1977**). "Identification and separation of acoustic frequency following responses (ffrs) in man," Electroen. Clin. Neuro. **42**, 493-500.

Sommers, M. S., Hale, S., Myerson, J., Rose, N., Tye-Murray, N., and Spehar, B. (**2011**). "Listening comprehension across the adult lifespan," Ear Hear. **32**, 775-781.

Sommers, M. S., and Humes, L. E. (**1993**). "Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners," J. Acoust. Soc. Am. **93**, 2903-2914.

Soucek, S., and Michaels, L. (**1990**). *Hearing loss in the elderly: Audiometric, electrophysiological, and histopathological aspects* (Springer-Verlag).

Souza, P. E., Boike, K. T., Witherell, K., and Tremblay, K. (**2007**). "Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise," J. Am. Acad. Audiol. **18**, 54-65.

Spoendlin, H. (**1970**). "Structural basis of peripheral frequency analysis," in *Frequency analysis and periodicity detection in hearing*, edited by R. Plomp, and G. F. Smoorenburg (Sijthoff, Leiden), pp. 2-40.

Spoendlin, H., and Schrott, A. (**1989**). "Analysis of the human auditory nerve," Hear. Res. **43**, 25-38.

Srulovicz, P., and Goldstein, J. L. (**1983**). "A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," J. Acoust. Soc. Am. **73**, 1266-1276.

Stillman, R. D., Crow, G., and Moushegian, G. (**1978**). "Components of the frequency-following potential in man," Electroen. Clin. Neuro. **44**, 438-446.

Stillman, R. D., Moushegian, G., and Rupert, A. L. (**1976**). "Early tone-evoked responses in normal and hearing-impaired subjects," Audiology **15**, 10-22.

Stone, M. A., and Moore, B. C. J. (**2003**). "Tolerable hearing-aid delays. Iii. Effects on speech production and perception of across-frequency variation in delay," Ear Hear. **24**, 175-183.

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., and Kaplan, G. A. (**2000**). "Negative consequences of hearing impairment in old age: A longitudinal analysis," Gerontologist **40**, 320-326.

Strelcyk, O., and Dau, T. (**2009**). "Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing," J. Acoust. Soc. Am. **125**, 3328-3345.

Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (**1998**). "Temporal processing in the aging auditory system," J. Acoust. Soc. Am. **104**, 2385-2399.

Swets, J. A., Shipley, E. F., McKey, M. J., and Green, D. M. (**1959**). "Multiple observations of signals in noise," J. Acoust. Soc. Am. **31**, 514-521.

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

Terkildsen, K., Osterhammel, P., and in't Veld, F. H. (**1974**). "Far field electrocochleography, electrode positions," Scand. Audiol. **3**, 123-129.

Turner, C. W., Burns, E. M., and Nelson, D. A. (**1983**). "Pure tone pitch perception and low-frequency hearing loss," J. Acoust. Soc. Am. **73**, 966-975.

Turner, C. W., Humes, L. E., Bentler, R. A., and Cox, R. M. (**1996**). "A review of the literature on hearing aid benefit as a function of time," Ear Hear. **17**, 14S-28S.

Tyler, R. S., Summerfield, A. Q., Wood, E. J., and Fernandes, M. A. (**1982**). "Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners," J. Acoust. Soc. Am. **72**, 740-752.

van Bergeijk, W. A. (**1962**). "Variation on a theme of békésy: A model of binaural interaction," J. Acoust. Soc. Am. **34**, 1431-1437.

Van de Bogaert, T., Wouters, J., Klasen, T. J., and Moonen, M. (**2005**). "Distortion of interaural time cues by directional noise reduction systems in modern digital hearing aids," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005.* (New Paltz, New York, USA), pp. 57-60.

Van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," J. Acoust. Soc. Am. **101**, 1671-1680.

Vander Werff, K. R., and Burns, K. S. (**2011**). "Brain stem responses to speech in younger and older adults," Ear Hear. **32**, 168-180.

von Békésy, G. (**1960**). *Experiments in hearing* (McGraw-Hill, New York).

Wagener, K., Josvassen, J. L., and Ardenkjær, R. (**2003**). "Design, optimization and evaluation of a danish sentence test in noise: Diseño, optimización y evaluación de la prueba danesa de frases en ruido," Int. J. Audiol. **42**, 10-17.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). "The precedence effect in sound localization," Am. J. Psychol. **62**, 315-336.

Walton, J. P., Simon, H., and Frisina, R. D. (**2002**). "Age-related alterations in the neural coding of envelope periodicities," J. Neurophysiol. **88**, 565-578.

Warren, L. R., Wagener, J. W., and Herman, G. E. (**1978**). "Binaural analysis in the aging auditory system," J. Gerontol. **33**, 731-736.

Webster, F. A. (**1951**). "The influence of interaural phase on masked thresholds: The role of interaural time deviation," J. Acoust. Soc. Am. **23**, 452-462.

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.

Winter, I. M., and Palmer, A. R. (**1990**). "Responses of single units in the anteroventral cochlear nucleus of the guinea pig," Hear. Res. **44**, 161-178.

Woolf, N. K., Ryan, A. F., and Bone, R. C. (**1981**). "Neural phase-locking properties in the absence of outer hair cells," Hear. Res. **4**, 335-346.

Worden, F. G., and Marsh, J. T. (**1968**). "Frequency-following (microphonic-like) neural responses evoked by sound," Electroen. Clin. Neuro. **25**, 42-52.

Working Group on Speech Understanding and Aging (**1988**). "Speech understanding and aging," J. Acoust. Soc. Am. **83**, 859-895.

Wright, A., Davis, A., Bredberg, G., Ulehlova, L., and Spencer, H. (**1987**). "Hair cell distributions in the normal human cochlea," Acta Otolaryngol. **Suppl. 444**, 1-48.

Wright, B. A., and Fitzgerald, M. B. (**2001**). "Different patterns of human discrimination learning for two interaural cues to sound-source location," Proc. Natl. Acad. Sci. U. S. A. **98**, 12307-12312.

Yost, W. A. (**1974**). "Discrimination of interaural phase differences," J. Acoust. Soc. Am. **55**, 1299-1303.

Young, E. D., and Sachs, M. B. (**1979**). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibres," J. Acoust. Soc. Am. **66**, 1381-1403.

Zurek, P. M., and Delhorne, L. A. (**1987**). "Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment," J. Acoust. Soc. Am. **82**, 1548-1559.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J. Acoust. Soc. Am. **28**, 860-864.