Multi-frequency Electrophysiological Estimates of Auditory Temporal Acuity

Negar Ahzan
The University of Western Ontario

Supervisor
Purcell, David W.
The University of Western Ontario

Graduate Program in Health and Rehabilitation Sciences
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science
© Negar Ahzan 2019

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons, and the Speech and Hearing Science Commons

Recommended Citation
https://ir.lib.uwo.ca/etd/6297

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.
Abstract

Auditory temporal acuity, a listener’s ability to discriminate rapid changes in the envelope of an auditory signal over time, is crucial for understanding speech. Electrophysiological measurement of auditory temporal acuity is beneficial when we cannot achieve reliable behavioural responses. The envelope following response (EFR) evoked by a changing (swept) amplitude-modulated (AM) stimulus is significantly correlated with behavioural measures of temporal acuity in humans. Previous research using AM broadband noise carriers may have been affected by the cancellation of evoked potentials at the measurement electrodes due to out-of-phase interference of parallel responses initiated at different times due to cochlear travelling wave timing differences. This study aimed to examine the possibility of using narrow-band noise carriers with different center frequencies, which enables recording of EFR from low, mid and high-frequency cochlear regions individually, while the modulation frequency gradually varied over time, and to determine whether we can improve the previously proposed objective method in three groups of normal-hearing participants.

Keywords

Electrophysiological measurements, temporal acuity, envelope following response (EFR), narrow-band noise carrier, normal hearing adults
Summary for Lay Audience

Temporal auditory acuity refers to the sensitivity of the human auditory system to fluctuations in the loudness of a sound over time, which is essential for speech perception. The envelope following response (EFR) is brain activity that can be recorded while special sounds are presented to the listener. It is measured from surface electrodes placed on the human scalp. In our study, special noise sounds that fluctuated in loudness were used to elicit EFRs. The maximum rate at which the sounds could fluctuate and still caused a detectable EFR was compared with people’s ability to notice the fluctuations behaviorally. The purpose of the study was to investigate how the EFR varies with the frequency content of sound, and whether some frequencies obtain a better correlation with behavioral measurements. This research contributes to our understanding of how sound is processed in the brain over time.
Acknowledgments

First of all, I would like to sincerely thank my supervisor Dr. David Purcell for his exceptional support, patience, motivation, aspiring guidance and constructive advice through all stages of my master’s thesis. He has always encouraged me to work hard, but also made sure I had a healthy work-life balance. I could not have imagined having a better and kinder supervisor and mentor. Dr. Purcell: thank you for giving me this wonderful learning opportunity.

I would also like to thank my advisory committee members, Dr. Prudence Allen and Dr. Vijay Parsa for their invaluable advice during my project. My special thanks goes to Dr. Andrew Johnson for his valuable suggestions with regard to the statistical analyses of this thesis, and Dr. Chris Allan for her willingness to help recruiting SLP students for my data collection.

I would like to thank all my study volunteers who devoted their time to participate in this research study.

I am also thankful to postdoctoral scholars of the National Centre for Audiology: Jordana Soares, Maaike Van Eeckoutte and Sangam Veeranna. Sangam and Jordana: thank you for sharing your knowledge of psychophysical measures, for countless discussions and your invaluable suggestions. Maaike: thank you for being my mentor in learning RStudio, for guiding me through my funny codes, for reading my thesis, and for being such a supportive friend. I would also like to thank Dr. Erfan Aref Eshghi for his generous help with R and sharing his enthusiasm for science.

Thank you to my lab mates, Adrienne, Ashira, Caitlin, Matt, Max, Monisha and Rebecca for all the moral supports, wonderful conversations, all the fun we had in the last two years and more importantly being such amazing friends. Special thanks to Monisha for helping us create our stimuli with her excellent auditory temporal acuity, and for being the best distraction in the SAFER Lab.

Finally, I am grateful to my loving family: my parents, my sister and brother, and my wonderful friends in Canada for providing me with their endless support and continues encouragement along the way.

This research project was supported by the Natural Sciences and Engineering Research Council of Canada and Ontario Research Fund.
# Table of Contents

Abstract ........................................................................................................................................... ii

Summary for Lay Audience .............................................................................................................. iii

Acknowledgments .............................................................................................................................. iv

Table of Contents ............................................................................................................................. v

List of Tables ..................................................................................................................................... viii

List of Figures ................................................................................................................................... ix

List of Appendices ............................................................................................................................. xi

List of Abbreviations ........................................................................................................................... xii

Chapter 1 ........................................................................................................................................... 1

1 Introduction ................................................................................................................................... 1

1.1 Tools for assessing temporal acuity ............................................................................................ 3

1.2 Behavioral measures of amplitude modulation detection thresholds .................................. 4

1.3 Electrophysiological measures of the temporal modulation transfer function (TMTF) ................................................................................................................................. 7

1.3.1 Definition .............................................................................................................................. 7

1.3.2 Stimulus parameters ............................................................................................................ 9

1.3.3 Recording ............................................................................................................................ 9

1.3.4 ASSRs to amplitude-modulated stimuli with fixed AM rate .............................................. 10

1.3.5 EFRs to amplitude-modulated stimuli with swept AM rate .............................................. 11

1.3.6 Neural Generators for EFR ................................................................................................ 14

1.4 EFR phase interactions ................................................................................................................ 17

1.5 Purpose of thesis ......................................................................................................................... 20

Chapter 2 ...................................................................................................................................... 22

2 General method ............................................................................................................................... 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.2 Effects of carrier frequency on EFR thresholds</td>
<td>55</td>
</tr>
<tr>
<td>4.2.3 Phase interactions</td>
<td>55</td>
</tr>
<tr>
<td>4.3 Behavioral outcomes</td>
<td>56</td>
</tr>
<tr>
<td>4.3.1 Effects of carrier frequency on behavioural AM detection thresholds</td>
<td>56</td>
</tr>
<tr>
<td>4.4 Correlation between EFR and behavioural thresholds</td>
<td>58</td>
</tr>
<tr>
<td>4.4.1 Reliability of behavioural outcomes</td>
<td>60</td>
</tr>
<tr>
<td>4.4.2 Reliability of EFR outcomes</td>
<td>65</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>68</td>
</tr>
<tr>
<td>5 Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>References</td>
<td>69</td>
</tr>
<tr>
<td>Appendices</td>
<td>79</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>88</td>
</tr>
</tbody>
</table>
List of Tables

Table 2.1: Summary of stimuli parameters.............................................................. 26

Table 3.1: Summary of age, pure-tone hearing thresholds, behavioural AM detection thresholds and EFR maximum significant frequency across three groups.........................36

Table 3.2: Results from Kruskal-Wallis Rank-Sum test that evaluates whether a significant difference exists between NBN groups in terms of EFR amplitudes across different modulation frequency bands........................................................................................................42

Table 3.3: Results from Wilcoxon Rank-Sum test that identified whether the low, mid and high-frequency NBN groups were different in terms of the EFR amplitudes across the modulation frequency band 100-200 Hz.................................................................42

Table 3.4: Results from Kruskal-Wallis rank sum test that evaluates whether the low, mid and high-frequency NBN groups were different in terms of EFR maximum frequency..................................................................................................................44

Table 3.5: Results from Kruskal-Wallis Rank-Sum test that evaluates whether the low, mid and high-frequency NBN groups were different in terms of modulation detection threshold (Hz)..................................................................................................................46

Table 3.6: Results from Wilcoxon Rank-Sum test that identified whether the low, mid and high-frequency NBN were different in terms of modulation detection threshold.............46

Table 3.7: Output from the Poisson regression model determining the non-significant difference between different types of carriers in terms of the number of valleys in the EFR modulation transfer function........................................................................................................50

Table 3.8: Results from the Wilcoxon Rank-Sum test indicated whether th NBN groups were different than the BBN groups in terms of the number of valleys in the EFR modulation transfer function. The low p-value indicates a high probability that the test results are correct..................................................................................................................52
List of Figures

Figure 1.1: TMTF for broadband noise carrier.................................................................6

Figure 1.2: The 40-Hz AM stimulus (top) and the evoked auditory steady-state response and in time (bottom left) and frequency domain (bottom right). The steady-state auditory response is evoked at the frequency of modulation and follows the periodicity of the AM stimulus……8

Figure 2.1: Study protocol ..........................................................................................22

Figure 3.1: EFRs as a function of AM rates from 80-600 Hz for all participants for low, mid and high-frequency NBN groups........................................................................38

Figure 3.2: EFRs and corresponding EEG background noise mean amplitude and standard error of the mean for different NBN carriers.................................................................39

Figure 3.3: Grand average of EFRs and standard error of the mean for all retained participants for low, mid and high-frequency NBN carriers in the range of 80 to 600 Hz.................41

Figure 3.4: Grand average of EEG background noise and standard error of the mean for all retained participants for low, mid and high-frequency NBN carriers in the range of 80 to 600 Hz.................................................................41

Figure 3.5: Boxplots of EFR thresholds across different NBN carriers. No statistically significant difference was found between different types of carriers.................................43

Figure 3.6: Boxplots of behavioural AM threshold across different carriers. A Statistically significant difference was found between different types of carriers.................................45

Figure 3.7: Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for low-frequency NBN carrier.................................................................47

Figure 3.8: Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for mid-frequency NBN carrier.................................................................48
Figure 3.9: Scatterplots are showing the correlation between the EFR and behavioural thresholds (Hz) for high-frequency NBN carriers………………………………………………………….48

Figure 3.10: Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for all subjects………………………………………………………………………………………………..49

Figure 3.11: Histograms of valleys in EFRs modulation transfer function for different stimuli across different studies in different group of individuals……………………………………..51

Figure 3.12: Boxplots of valley counts across different types of carriers. The NBN carriers had a lower number of valleys or fine structure in comparison with studies that used BBN carriers…………………………………………………………………………………………………………………53

Figure 4.1: Discrepancy between behavioural and EFR maximum frequency (EFR threshold). A) Behavioural TMTF for a sample subject. Behavioural AM threshold was 176 Hz for this listener. B) EFR modulation transfer function and corresponding EEG noise-estimate for the same listener in Figure (A). EFR maximum frequency for this subject was 513 Hz …………………………………………………………………………………………………….63

Figure 4.2: Multiple attempts of the behavioral AM detection task for a sample subject. The first to third attempts show the result of behavioural AM detection threshold obtained in the two-hour TMTF/EFR study session. This participant was asked to repeat the behavioural test in a second session. The plots of the fifth to eighth attempts were obtained in the second session of the behavioural test………………………………………………………………………………….65
List of Appendices

Appendix A: Approval for Research Involving Human Participants ........................................... 82

Appendix B: Letter of Information and Consent Form ................................................................. 84
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Evoked Potentials</td>
<td>AEPs</td>
</tr>
<tr>
<td>American Speech-Language-Hearing Association</td>
<td>ASHA</td>
</tr>
<tr>
<td>Amplitude Modulation</td>
<td>AM</td>
</tr>
<tr>
<td>American Academy of Audiology</td>
<td>AAA</td>
</tr>
<tr>
<td>Auditory Neuropathy Spectrum Disorder</td>
<td>ANSD</td>
</tr>
<tr>
<td>Auditory Processing Disorder</td>
<td>APD</td>
</tr>
<tr>
<td>Auditory Steady-State Responses</td>
<td>ASSRs</td>
</tr>
<tr>
<td>Broadband Noise</td>
<td>BBN</td>
</tr>
<tr>
<td>Carrier Signal</td>
<td>$f_c$</td>
</tr>
<tr>
<td>Cochlear Nucleus</td>
<td>CN</td>
</tr>
<tr>
<td>Decible</td>
<td>dB</td>
</tr>
<tr>
<td>Discrete Fourier Transform</td>
<td>DFT</td>
</tr>
<tr>
<td>Dynamic Range</td>
<td>DR</td>
</tr>
<tr>
<td>Electroencephalogram</td>
<td>EEG</td>
</tr>
<tr>
<td>Envelope Following Responses</td>
<td>EFRs</td>
</tr>
<tr>
<td>Fourier Analyser</td>
<td>FA</td>
</tr>
<tr>
<td>Frequency Following Response</td>
<td>FFRs</td>
</tr>
<tr>
<td>Fundamental Frequency</td>
<td>$F_0$</td>
</tr>
<tr>
<td>Inferior Colliculus</td>
<td>IC</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>IQR</td>
</tr>
<tr>
<td>Hertz</td>
<td>Hz</td>
</tr>
<tr>
<td>Hearing Level</td>
<td>HL</td>
</tr>
<tr>
<td>Magnetoencephalography</td>
<td>MEG</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Medial Geniculate Body</td>
<td>MGB</td>
</tr>
<tr>
<td>Modulation Depth</td>
<td>m</td>
</tr>
<tr>
<td>Modulation Frequency</td>
<td>$f_m$</td>
</tr>
<tr>
<td>Narrow-Band Noise</td>
<td>NBN</td>
</tr>
<tr>
<td>Parameter Setting by Sequential Estimation</td>
<td>PEST</td>
</tr>
<tr>
<td>Post-Stimulus Time histogram</td>
<td>PST</td>
</tr>
<tr>
<td>Quantile-Quantile</td>
<td>Q-Q</td>
</tr>
<tr>
<td>Quasi-Frequency Modulated</td>
<td>QFM</td>
</tr>
<tr>
<td>Signal to Noise Ratio</td>
<td>SNR</td>
</tr>
<tr>
<td>Sensation Level</td>
<td>SL</td>
</tr>
<tr>
<td>Sound Pressure Level</td>
<td>SPL</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>SD</td>
</tr>
<tr>
<td>Standard Error of Measurement</td>
<td>SEM</td>
</tr>
<tr>
<td>Superior Olivary Complex</td>
<td>SOC</td>
</tr>
<tr>
<td>Tucker Davis Technologies</td>
<td>TDT</td>
</tr>
<tr>
<td>Temporal Modulation Transfer Function</td>
<td>TMTF</td>
</tr>
<tr>
<td>Three-Interval Two-Alternative Forced-Choice</td>
<td>3I-2AFC</td>
</tr>
<tr>
<td>Two-Alternative Forced-Choice</td>
<td>2AFC</td>
</tr>
<tr>
<td>Vertex</td>
<td>Cz</td>
</tr>
</tbody>
</table>
Chapter 1

1 Introduction

Hearing loss is a global health issue, and is the fourth major contributor to years lived with disability (Wilson, Tucci, Merson, & O’Donoghue, 2017). More than 5 percent of the world’s population (around 466 million people) suffer from disabling hearing loss, and 34 million of these are children (World Health Organization, 2018). Unaddressed hearing loss can affect how we learn spoken language, how we learn to read and write, as well as development of emotional status and learning of social skills. Hearing loss can be caused by deficits at any level of the auditory system, from transmission of the acoustic signal through the external and middle ear, or the absence of proper transduction where sound vibrations are converted to electrical impulses by the hair cells of the cochlea in the inner ear. Besides, hearing loss manifests as speech perception difficulties which could be caused by a dysfunction in neural coding of the auditory nerve, encoding of sound at the synaptic level or in the auditory central nervous system. Speech perception deficits can be accompanied by elevated pure-tone audiometric thresholds or in the absence of peripheral hearing loss.

The speech signal is not a constant waveform; instead, it shows different kinds of variations over time. Having good temporal processing abilities is essential for speech intelligibility. According to Rosen (1992), the temporal information of a speech signal can be divided into the temporal fine structure, the fast temporal changes in amplitude and frequency of the carrier sound over time, the temporal envelope, the slower changes in overall peak amplitude superimposed on the carrier sound over time, and periodicity. We will focus on the more gradual changes in the temporal envelope of the signal over time. Many studies have shown that much of the information needed to understand spoken language is carried within the temporal envelope, and even in the absence of spectral cues, we can understand speech by relying on the temporal envelope (Rosen, 1992; Shannon, Zeng, Kamath, Wygonski & Ekelid, 1995). Also, envelope cues are enough to identify different groups of consonants (Van Tasell, Soli, Kirby, & Widin, 1987). In continuous speech, envelope fluctuations range from 2 to 50 Hz which conveys segmental cues to the manner of articulation, voicing, vowel identity, and prosodic cues (Rosen, 1992). Also, low-frequency modulations of 4 and 20 Hz correspond to the processing rate of syllables and phonemes in speech, respectively,
which are essential for speech intelligibility (Shannon et al., 1995). Auditory temporal acuity, the ability of the auditory system to follow fluctuations in the amplitude envelope of an acoustic signal over time, is one of the fundamental aspects of temporal processing, and is crucial for speech understanding.

Temporal processing is impaired in various clinical populations. First, reduced ability in following the speech envelope is one reason for hearing problems in older adults, which can affect the way they perceive speech and their benefit from amplification (Schneider, Daneman, & Pichora-Fuller, 2002; Schneider, Speranza, & Pichora-Fuller, 1998). A smaller number of phase-locked evoked responses to amplitude-modulated stimuli in older listeners can be interpreted as an objective correlate of their poor word recognition ability in difficult hearing situations (Leigh-Paffenroth & Fowler, 2006). Second, auditory neuropathy spectrum disorder (ANSD) is a disorder of the afferent auditory nervous system but with normal and near normal cochlear and outer hair cell function. Although ANSD could be accompanied by different degrees of hearing loss, speech perception abilities do not correspond with their loss of audibility. Clinically, ANSD is accompanied by disordered processing of essential temporal features of the auditory signal for sound localization, speech discrimination, and identification of sound in background noise (Starr et al., 1991; Zeng, Oba, Garde, Sininger, & Starr, 1999). Individuals with ANSD have more difficulty performing time-based discrimination tasks compared to frequency discrimination (Starr et al., 1991). Narne (2013) investigated the sensitivity of individuals with ANSD to both temporal envelope and fine structure cues. Results indicated that impaired ability to follow temporal envelopes in listeners with ANSD is responsible for their reduced speech perception abilities in quiet environments. An impaired ability to follow both envelope and fine structure cues of the speech signal is the reason underlying their poor speech perception in the presence of noise. Third, temporal processing problems often occur in children with developmental dyslexia, which is a specific learning disability in reading. Reading difficulties in dyslexic children can be caused by temporal processing problems (Tallal, 1980). In individuals with dyslexia, the neural representation of speech sounds is defective, and as a result, they cannot relate speech sounds to alphabetic symbols. In children with dyslexia, an elevated modulation detection threshold has been seen for modulation frequencies between 4-1024 Hz with the most significant difference around 4 Hz corresponding to the syllabic rate of speech (Poelmans et al., 2011; Rocheron,
Lorenzi, Fullgrabe, & Dumont, 2002). Adults with dyslexia have an elevated amplitude modulation detection threshold over a broad range of modulation frequencies between 10 and 320 Hz. This elevated threshold is indicative of amplitude modulation processing difficulties for modulation frequencies associated with the phonemic rate of speech and for higher modulation frequencies that are less important for speech envelope processing (McAnally & Stein, 1997; Menell, McAnally, & Stein, 1999). Adults with dyslexia have been shown to have smaller auditory evoked responses to AM assessed by the steady-state auditory response (Menell et al., 1999). Finally, auditory processing disorder (APD) is characterized by deficits in the central nervous system for auditory perceptual abilities including sound localization, auditory discrimination, auditory pattern recognition, temporal aspects of audition, and reduced auditory performance in the presence of background noise [American Speech-Language-Hearing Associations (ASHA), 2005]. Children with APD usually have listening difficulties despite having normal audiometric thresholds [American Academy of Audiology (AAA), 2010] and the listening difficulties can be accompanied by speech and language difficulties (ASHA, 2005). The results of one study on temporal acuity of children with APD, compared to normally developing children, showed that gap detection thresholds were significantly poorer (longer) in APD group (Philips, Comeau, and Andrus, 2010).

Given the importance of the temporal envelope in speech perception, assessment of temporal auditory acuity as a clinical tool for diagnosing and monitoring temporal processing disorders in the auditory system is essential.

### 1.1 Tools for assessing temporal acuity

Two commonly used behavioral methods to measure auditory temporal acuity are: 1) amplitude modulation detection, which assesses the ability of the listener to behaviorally discriminate similar auditory signals that are distinct in their temporal envelope; and 2) gap detection threshold, which is the ability of the listener to detect a gap in a continuous noise or tone. Behavioural measures of auditory temporal acuity correspond well with speech understanding abilities (Narne, 2013; Schneider et al., 2002). However, outcomes of behavioural measurements can be affected by the reliability of the responses because behavioural tests do require memory, attention, and
active cooperation of the listener. It is essential to validate alternative methods to evaluate temporal acuity objectively.

Electrophysiological measurement of auditory temporal acuity would be beneficial when the accuracy of behavioural measurements is affected by poor reliability of the responses as these neural responses could be obtained objectively without the active contribution of the listener. It would also be useful in special situations such as evaluating difficult-to-test populations, including individuals with cognitive disabilities, or infants and young children. Studies have shown that there is a good correlation between subjective and objective measures of temporal acuity (Purcell, John, Schneider, & Picton, 2004).

Sections 1.2 and 1.3 focus on behavioural and electrophysiological measures of AM detection.

1.2 Behavioral measures of amplitude modulation detection thresholds

Auditory temporal acuity can be assessed using psychophysical tests which rely on the listener’s behavioral responses. A common psychophysical method for assessing temporal resolution is to examine sensitivity to amplitude modulation (AM) through an amplitude modulation detection task. The amplitude modulation detection task is beneficial compared to other measures of temporal acuity, such as gap detection, as it reflects the effects of intensity resolution and temporal resolution separately, therefore it is a better measure of temporal acuity (Strickland & Viemeister, 1997). Moreover, the amplitude modulated stimulus varies in temporal envelope over time, which has envelope fluctuations similar to speech and could provide information about how a listener perceives changes in a speech signal over time.

The envelope fluctuations can be created by modulating a sound. The process of manipulating some characteristics of a carrier signal is known as modulation. In amplitude modulation (AM), the overall amplitude of the carrier signal is varied systematically over time. The AM sound is produced by multiplying two sine waves: a carrier \( f_c \) and a modulator \( f_m \). A tonal or noise stimulus is often used as the carrier sound. The carrier signal determines the specific place on the basilar membrane to which hair cells are activated in response to the AM stimulus. The modulation process
changes the spectral characteristics of the sound. The AM sound does not have energy at $f_m$; instead, the spectrum of the sinusoidal AM sound contains spectral energy at the $f_c$ and two sideband frequencies ($f_c \pm f_m$) on either side of the carrier which are separated by intervals equal to $f_m$. The amplitude of the sidebands is half that of $f_c$. For a more complex stimulus, the modulation process adds sidebands for each frequency component of the carrier.

It is common to refer to the modulation frequency ($f_m$) of the modulator. The modulation depth ($m$) represents the ratio of the change in the amplitude of the AM carrier relative to the unmodulated condition and represents the amount of modulation. Modulation depth is expressed in percentage or a dB value as $20 \log (m)$ and ranges from 0 or 0% (unmodulated signal) to 1.0 or 100% modulation.

In the psychophysical task of amplitude modulation detection, the listener is asked to discriminate amplitude-modulated from unmodulated sounds. In this test, at a fixed modulation frequency, the modulation detection threshold is obtained by adjusting the depth of modulation systematically using adaptive procedures (Findlay, 1978; Levitt, 1971; Taylor & Creelman, 1967) to find the minimum modulation depth that a listener is able to detect. In this adaptive procedure, by using stepping rule criteria, the modulation depth is adjusted based on the response of the listener to the previous depth. The level of difficulty increases after a specific number of correct responses and reduces after a specific number of incorrect responses. The test continues until a preset percentage of correct responses is obtained. The same procedure is repeated for a broad range of modulation frequencies. The adaptive procedure will be investigated in detail in chapter 2 of this thesis. The output of this task is plotted in the form of a graph called the temporal modulation transfer function (TMTF; Figure 1.1). TMTF is a plot that relates amplitude modulation detection thresholds across modulation frequencies.

Viemeister (1979) performed one of the earliest psychophysical measurements of the TMTF. According to Viemeister (1979), the auditory system has a limited capacity in following amplitude modulation. In young normal-hearing listeners, the behavioral TMTF for a sinusoidally AM broadband noise (BBN) carrier has a low-pass filter pattern with a 3-dB cut-off frequency of about 55 Hz, which means that the amplitude modulation detection threshold remains relatively constant in the range of 2 to 55 Hz, and beyond this range, the sensitivity to the AM carrier declines. A normal-hearing
A listener can discriminate amplitude modulation of 25 percent up to about 500 Hz (Viemeister, 1979). At higher modulation depths, the detection cut-off frequency increases because discrimination is easier at larger depths.

**Figure 1.1: TMTF for broadband noise. Figure extracted from Viemeister (1979)**

In addition to broadband noise as a carrier sound, psychophysical studies have examined AM detection thresholds with frequency-specific carriers such as pure-tones (Kohlrausch, Fassel, & Dau, 2000; Moore & Glasberg, 2001), and narrow-band noise carriers (Dau, Kollmeier, & Kohlrausch, 1997; Eddins, 1999; Eddins, 1993; Formby & Muir, 1988; Strickland, 2000; Strickland & Viemeister, 1997). A narrow-band noise (NBN) is characterized by a center frequency and bandwidth and its energy is distributed over a relatively small region of the cochlea, whereas a broadband noise carrier has a broader spectrum and activates the whole cochlea. These carriers have different modulation detection thresholds. Using pure-tone carriers, AM threshold decreases (improves) at higher modulation frequencies which is in contrast with the low-pass filter shape of TMTF for a broadband carrier. One likely reason for this improved sensitivity at higher modulation frequencies is that the sideband frequencies produced by amplitude modulation can be resolved. Sideband frequencies can be heard as separate tones if they are separated by at least one auditory critical band (Viemeister, 1979). Therefore, recognition of spectral cues, not temporal cues, help the listener to recognize the amplitude modulation. A group of researchers (Dau et al., 1997; Eddins, 1999; Eddins, 1993; Formby & Muir, 1988; Strickland, 2000; Strickland & Viemeister, 1997) investigated the influence of narrow-band noise bandwidth and frequency region on the AM detection threshold. In these studies, different techniques were applied to
limit the availability of spectral cues. In the absence of spectral cues, the TMTF had a low-pass filter characteristic, and AM threshold increased with an increase in modulation frequency, which is consistent with TMTF for broadband noise carriers. These studies assessed the influence of narrow-band noise bandwidth and frequency region on the AM detection threshold. Sensitivity to AM improves as the bandwidth of the AM stimulus becomes wider by activating a broader range of the cochlea. Also, the TMTF cut-off frequency is lower for low frequency narrow-band noises compared to high frequency narrow-band noises with the same bandwidth at low stimulus spectrum levels. However, the same effect has not been observed at stimulus spectrum levels of 40 dB SL or higher in the literature (Eddins, 1993, 1999; Due et al., 1997; Strickland & Viemeister, 1997).

1.3 Electrophysiological measures of the temporal modulation transfer function (TMTF)

Auditory temporal acuity can be studied by recording auditory evoked potentials (AEPs) to amplitude modulated sounds. In this section, electrophysiological estimates of temporal acuity are discussed. First, a definition of auditory evoked potentials is provided which is followed by an introduction to auditory steady-state and envelope following responses, their stimulus parameters, and recording procedures. This section ends with a review of the literature.

1.3.1 Definition

Auditory evoked potentials (AEPs) are extracted from scalp-recordings of the brain’s electrical activity (the electroencephalogram or EEG) generated by synchronous neural activity in the auditory system in response to an auditory stimulus. AEPs have been proven to be an effective and non-invasive objective method of measuring neural responses to auditory stimuli and testing the integrity of the central auditory pathway (Picton, 2011). As it has been suggested in a review by Picton (2013), AEPs can be classified into different categories based on their evoking auditory stimuli and how they respond over time. First, transient responses are evoked by the rapid changes in an auditory stimulus, and the response to the first stimulus ends before the second acoustic stimulus occurs. Second, sustained responses are evoked by the continuity of the
stimulus, and they last through the duration of the acoustic stimulus. Third, the following responses are between transient and sustained responses, and they are evoked by repetitive changes in the stimulus. The rate of stimulus change is too fast for the response to resolve, therefore the response to the first acoustic stimulus overlaps with responses to subsequent stimuli (Picton, John, Dimitrijevic, & Purcell, 2003). Following responses can either follow the fine structure of the carrier stimulus which are called frequency following responses (FFRs), or the envelope of the acoustic stimulus which are called envelope following responses (EFRs). EFRs are elicited by a stimulus that continuously changes in the envelope over time. Since the stimulus is not constant, the amplitude and the phase of brain responses also change correspondingly (Aiken & Picton, 2008; Purcell et al., 2004). If the change in the envelope of the acoustic stimulus is periodic with a fixed rate, the following response is called the auditory steady-state response (ASSR; Figure 1.2).

![Stimulus and Response Diagram](Figure 1.2: The 40-Hz AM stimulus (top) and the evoked auditory steady-state response and in time (bottom left) and frequency domain (bottom right). The steady-state auditory response is evoked at the frequency of modulation and follows the periodicity of the AM stimulus. Figure extracted form Burkard, Don, Eggermont, 2007, Ch21, Fig 21.1.)
1.3.2 Stimulus parameters

A variety of stimuli have been used for eliciting ASSRs including broadband stimuli such as noise (John, Lins, Boucher, & Picton, 1998), or frequency-specific stimuli, such as amplitude-modulated tones (John et al., 1998). The most commonly-used stimulus for recording ASSRs are AM pure tones (Picton, John, Purcell, & Plourde, 2003). The ASSR follows the envelope of the AM tone rather than the higher frequency of the carrier tone. In the time domain, the response demonstrates the periodicity of the AM signal, and in the frequency domain, it appears as a peak at the frequency of modulation (Picton, John, Purcell, et al., 2003). ASSRs can be recorded at a wide range of modulation frequencies. A critical characteristic of the ASSR is that the amplitude of the response reduces with increasing modulation frequency. However, background ongoing EEG and myogenic noise also decrease with increasing frequency which improves the signal-to-noise ratio at higher modulation frequencies (Picton, John, Purcell, et al., 2003). Studies have shown that the modulation frequency of 40-Hz evokes the largest ASSR in terms of its amplitude in awake adults and with a smaller peak at around 90 Hz. (Galambos, Makeig, & Talmachoff, 1981; Rees, Green, & Kay, 1986). The amplitude of the ASSR increases as the depth of modulation increases and modulation depths of 50% to 100% evoke the largest responses; however, the amplitudes saturate as the depth of modulation reaches 50% (Lins & Picton, 1995). By definition, ASSRs are evoked by a constant amplitude-modulated stimulus, and the phase and the amplitude of the evoked response are constant throughout the measurement time. In the case of recording EFRs, the evoking stimulus is not constant. Instead, it changes in the envelope over time. EFRs have been recorded in response to a wide range of sounds including speech stimuli (Aiken & Picton, 2008), an amplitude modulated white noise carrier with swept modulation frequency (Purcell et al., 2004), and an amplitude modulated white noise carrier with a swept modulation depth (Dimitrijevic et al., 2016).

1.3.3 Recording

Recordings should be performed while participants sit in a comfortable chair in a sound-attenuated room. The ASSR or EFR are measured by recording EEG with electrodes placed at specific sites on the scalp during stimulus presentation. The anatomic structures generating evoked responses to auditory stimuli are located at some distance
from the measurement electrodes; therefore, the response is mediated through body tissues, fluid, and skin to reach the measurement electrodes which reduces the amplitude of the evoked responses. The electrical activity at the electrodes is amplified to strengthen the response to the signal of interest and then filtered to remove the electrical activity that is not related to the response. These responses are small in amplitude, and it is carried with other brain activity including EEG background activity, electrical signals from other sources in the sound booth, and myogenic noises from jaw or neck movement. To improve the signal-to-noise ratio, hundreds of stimuli are presented, and the collected EEG signal is represented in the format of short time intervals of 1.024 second duration called epochs. As the recording continues, epochs are added together to make larger segments called sweeps. Then, brain activity is synchronously averaged across sweeps. The auditory evoked response will follow the same pattern in all stimulus repetitions so averaging will retain the response. However, the unrelated EEG noise does not have the same pattern in all stimulus presentations and therefore averages toward zero. The averaged time domain response is converted to the frequency domain by a discrete Fourier Transform algorithm. Finally, a statistical analysis determines whether or not the energy of the envelope following response at a specific modulation frequency has a statistically higher amplitude than the EEG background noise (John and Purcell, 2008).

1.3.4 ASSRs to amplitude-modulated stimuli with fixed AM rate

Multiple research studies have investigated associations between ASSRs and EFRs with modulation perception abilities. To estimate the modulation transfer function, separate ASSRs should be recorded at a broad range of modulation frequencies to determine electrophysiological representations of the TMTF. The magnitude of the ASSR represents sensitivity at the modulation frequency, and the depth at which the response can first be detected can serve as an estimate of the modulation detection threshold. Rees, Green, & Kay (1986) recorded ASSRs to sinusoidal AM pure-tones and broadband noise at modulation frequencies of 2-400 Hz in normal hearing individuals. Their plot of ASSR amplitude as a function of AM rate resembled a psychophysically measured TMTF with a 3 to 6 dB low-pass filter shape and a cut-off frequency of 40-50 Hz for both broadband noise and pure-tone carriers. The result of the study by Picton, Skinner, Champagne, Kellett, & Maiste (1987) showed that ASSRs could be recorded at modulation depths relatively similar to psychophysically measured
amplitude modulation thresholds in the frequency range of 30 to 50 Hz. In one magnetoencephalographic study, Roß, Borgmann, Draganova, Roberts, & Pantev (2000) recorded ASSRs to an AM 250 Hz pure tone carrier in the range of 10-98 Hz. ASSR modulation transfer function had a low-pass filter characteristic with an upper cut-off frequency near 50 Hz. This pattern is similar to the previously reported psychophysical estimates of TMTF when AM detection was based on the temporal changes of the envelope without spectral cues for tonal carriers (Viemeister, 1979).

1.3.5 EFRs to amplitude-modulated stimuli with swept AM rate

Recording ASSRs to a broad range of modulation frequencies is too time consuming for participants to realistically complete. A more efficient way of objectively obtaining TMTF is by sweeping through the modulation frequencies. The sweep technique was first introduced by Regan (1966, 1989) for visually evoked responses. In this technique, instead of recording multiple responses to different values of the same parameter, a particular parameter of interest changes continuously in a preset range and the evoked responses are analyzed as a function of that parameter (Picton, John, Dimitrijevic, et al., 2003). Since the evoking stimulus is not constant and changes in frequency and phase continuously, the evoked responses are no longer a steady-state response and have been called envelope following responses (EFRs). In hearing research, the sweep technique has previously been used by changing modulation rate or intensity (Linden, Campbell, Hamel, & Picton, 1985). Using the sweep technique, Purcell, John, Schneider, & Picton (2004) assessed auditory temporal acuity in humans objectively.

Purcell and colleagues (2004) recorded EFRs evoked by a white noise carrier amplitude modulated at a fixed modulation depth of 25 percent with the modulation rate swept over time from 20-600 Hz. Duration of each sweep of the stimulus was 30 seconds. In the first half of each sweep, the modulation rate increased linearly from minimum to maximum value, and in the second half of each sweep, the modulation frequency decreased to reach the minimum value. The first and second halves of each sweep were then averaged. A total of 50 to 100 sweeps of the swept stimulus were averaged in order to improve the signal-to-noise ratio. The highest modulation frequency at which the EFR was statistically significantly different from the EEG background noise was considered to be the objective threshold. According to their results, EFRs were significant until 485 Hz and 235 Hz for the young normal hearing and older adults,
respectively. Remarkable valleys and peaks were observed for both groups including a peak near 40 Hz, which is similar to what is often demonstrated for “standard” ASSRs, a peak near 80 Hz, as well as a minimum amplitude point near 70 Hz. Also, psychophysical measures of amplitude modulation and gap detection were carried out. Purcell and colleagues (2004) found a statistically significant correlation between the maximum modulation frequency evoking a significant EFR and the psychophysical measures of the gap (r = -0.43) and modulation detection (r = 0.72) in the combined group of young and older individuals. The result of this study confirmed that a swept stimulus elicits responses similar to those evoked using multiple discrete modulation rates in the same frequency range at both low and high modulation frequencies.

In a study by Dimitrijevic et al. (2016), EFRs were recorded to an amplitude modulated white noise carrier at the fixed modulation frequency of 41 Hz in which the amplitude modulation depth was gradually swept from 2 to 100 percent. A significant correlation was observed between the behavioural threshold for minimum AM depth detection and minimum AM depth that can evoke a detectable EFR. Moreover, EFR amplitudes elicited by swept AM depth were similar to ASSRs evoked by the same fixed modulation depth in normal hearing individuals and two groups of older adults with normal hearing and mild hearing losses. Given the absence of a significant increase in AM detection threshold with age, it is suggested that EFR amplitude does not change with age. Instead, aging affects other variables in the relationship between EFR and AM depth, such as degree of linearity, plateau level, dynamic range (DR) and slope of the EFR versus AM depth function. In the group of young adults, EFR amplitude as a function of AM depth showed a linear pattern with a shallower slope and less saturation. In contrast, in the group of older adults with puretone average of 35 dB HL, the function showed non-linearity at high modulation depths. This means that amplitude changes rapidly with a steeper slope and reaches the saturation level at lower AM depth. As a result, the dynamic range of the EFR amplitude is smaller in the group of older adults. Based on literature in aging, the reduced dynamic range is a sign of reduced neural encoding at the suprathreshold level, which could reduce speech intelligibility. Furthermore, the phase of the EFR as a function of AM depth was constant for the younger group, but it was negative for older groups for AM depths higher than 40 percent. The older adults with hearing loss had high-frequency hearing losses; this reduction in the phase was suggested as being related to neural encoding of the high-
frequency portion of the white noise carrier. However, this study suggested that future research should focus on the influence and contribution of low and high frequency carriers on the aforementioned phase-slope difference. The sweep technique used in this study provided important information about the plateau level, the degree of linearity, slope, and dynamic range at suprathreshold levels which helps to determine fine variability in the EFR response in different listeners.

In the third chapter of an unpublished dissertation by Alsamri (2017, p.40), EFRs to AM broadband noise were recorded with two conditions in four groups of subjects including a group of individuals (aged 10-66 years old) with ANSD and hearing loss where pure tone average thresholds exceeded 20 dB HL, and three non-ANSD control groups including a group of young normal hearing subjects (18-28 years), one group of older adults (41 to 62 years old) with hearing loss with pure tone averages of 30 dB HL, and another group of older adults (67 to 82 years old) with hearing loss with pure tone averages of 49 dB HL.

In condition one, at the fixed modulation frequency of 41-Hz, AM depth was swept continuously between 2% to 100% for a white noise carrier. In the second condition, at a fixed AM depth of 100%, the AM rates were continuously swept from 2-59 Hz and 62-300 Hz for the white noise carrier, and EFRs were analyzed as a function of the frequency of modulation and depth of modulation for both conditions. Additionally, psychophysical measurement of the modulation detection threshold was carried out to obtain the minimum amplitude modulation depth needed to detect a modulated white noise carrier at 41 Hz. Results of the behavioural test indicated that participants with ANSD showed elevated AM detection thresholds for 41 Hz AM stimuli regardless of their degree of hearing loss in comparison with the control groups in this study. Furthermore, for individuals with ANSD, larger AM depths were needed for EFRs to become significant and the ANSD group had significantly smaller EFR amplitudes compared to their non-ANSD counterparts. The findings of this study suggest that in all subjects, there was a significant correlation between behavioral amplitude modulation detection thresholds and the smallest depth at which the EFR was statistically detectable in the amplitude modulation depth condition ($r = .89$). In the sweep AM rate condition, the amplitudes of EFRs were smaller in participants with ANSD than in normal groups at AM rates of 2-59 Hz. The ANSD group not only
produced smaller EFR amplitudes at AM rates of 62-300 Hz, but also the highest AM rates at which EFRs were detected was shifted to lower AM rates.

The results of these studies suggest that EFRs to swept stimuli have the potential to be used as an effective and valid tool for objective measurement of temporal acuity due to the strong and significant correlation of this method with psychophysical measures.

1.3.6 Neural Generators for EFR

In the auditory system, neurons phase-lock their firing spikes to the temporal properties of the acoustic stimulus by firing at a specific phase of the stimulus envelope (Joris, Schreiner & Rees, 2004). In the auditory system, multiple neurons in different regions can phase lock to the same stimulus frequency, and as we ascend the auditory pathway, the upper cut-off limit of phase locking decreases. Identifying the isolated contribution of each generator is complicated since AEPs are recorded from electrodes on the surface of the scalp that are far from the possible generators. These far-field potentials are volume conducted to the scalp and represent the summed activity of groups of nuclei at different levels of the auditory pathway. Attributing this summed activity to individual anatomical origins is difficult. Magnetoencephalography (MEG) and scalp-recorded EEG studies have shown that the distinctive characteristic of these following potentials is that the relative contribution of individual sources changes depending on the stimulus frequency (Herdman, Picton, & Stapells, 2002; Joris, Schreiner & Rees 2004; Kuwada et al., 2002).

Much of the knowledge that we have about the neurophysiologic mechanism underlying amplitude modulation encoding in the auditory system comes from direct recordings in animals. Kuwada et al. (2002) recorded ASSRs to sinusoidally AM tones from the surface of the brain in unanesthetized rabbits and from scalp recorded surface electrodes in humans. In the same study, multiple local recordings were carried out from subcortical nuclei including the superior olivary complex (SOC), inferior colliculus (IC), and the primary and non-primary auditory cortex. Their study showed that the ASSR modulation transfer functions as a function of modulation frequency in rabbits and humans were similar in magnitude and latency. The TMTFs, the amplitude of ASSRs as a function of modulation frequency, showed a series of peaks and valleys in both rabbits and humans. TMTF in rabbits showed two large peaks, with one peak
around 25-35 Hz and a larger peak at 62 Hz. At higher modulation frequencies, there were an additional two smaller peaks. Similarly in humans, a peak at low modulation frequencies had the largest amplitude (around 40 Hz), and there were two smaller peaks at higher modulation frequencies (with the first dominant peak at around 80-100 Hz). The phases of the TMTF around the peak frequencies were fitted with linear regression lines and the slopes were used for estimating neural delays. The regions at the first peak (below 46 Hz), the second peak (80-100 Hz), and the third peak (160-260 Hz) had neural delays of 27, 12, and 8 ms, respectively. These peaks and valleys reflect the contribution of more than one generator contributing to the response at each modulation frequency, in which peaks are frequencies for which the responses of different sources are in phase, and valleys are a result of their out-of-phase interference. In the same study, local recordings from different areas of the brain in rabbits showed that the peak amplitude shifted from a high modulation frequency of about 250 Hz at the SOC, to 90 Hz at the IC, and 20 Hz at the auditory cortex. Since the peak amplitude of surface-recorded ASSRs in rabbits was not observed at the same modulation frequency as the locally-recorded ASSRs from the auditory cortex or IC, this supports the hypothesis that more than one generator contributed to surface-recorded ASSRs. Kuwada et al. (2002) suggested that the cortex is the primary neural generator for modulation frequencies lower than 80 Hz and has a minimal contribution for modulation frequencies higher than 150 Hz, while subcortical areas are generating responses to higher modulation frequencies.

Furthermore, ASSRs to low modulation frequencies with cortical origin were influenced by behavioural stimulation and pharmacological manipulation, while ASSRs to high modulation frequencies with subcortical origin remained almost unaffected or were influenced only slightly. In another study, the intracerebral generators for ASSRs in humans were investigated by recording ASSRs in response to a 1 kHz tone which was amplitude modulated at 12, 39, and 88 Hz (Herdman et al., 2002). ASSRs were recorded from 46 electrodes placed on the head. Six dipoles were modelled for each modulation frequency where the first two asymmetrical sources were fitted to two brainstem sources, two at the ipsilateral auditory cortex and two at the contralateral auditory cortex. Herdman et al. (2002) found that the brainstem contributes to all the tested modulation rates and had a larger response at 39 Hz compared to 12 Hz, and is the principal generator for 88 Hz. Instead, the cortical sources
contributed more to low modulation frequencies, and less than brainstem to high modulation rates. Also, both cortical and brainstem sources simultaneously contributed to generating the response at 39 Hz. This study estimated a latency of 19 ms for brainstem sources at 88 Hz which is consistent with previously reported latencies for the brainstem in the literature (John and Picton, 2000). However, further studies have suggested an apparent latency of 8 to 10 ms is more representative of neural generators in the brainstem.

The morphology of EFRs to swept AM rate stimuli gave rise to a hypothetical model for EFR generation with cortical and brainstem sources with different response amplitudes and latencies (Purcell et al., 2004). In this model, the cortical source has a latency of 29 ms and an amplitude of 85 nV up to 50 Hz, which decreases to 0 nV at 95 Hz. The brainstem source has constant amplitude of 35 nV up to a modulation frequency of 100 Hz and latency of 7.3 ms. The amplitude then declines linearly to 0 nV for modulation frequencies of 100-500 Hz. In this hypothetical model, the recorded response from scalp-recorded electrodes at any modulation frequency represents the summed response of these two sources. Keeping the latencies constant, amplitudes and relative phases of these sources change with the modulation frequency. As a result, an in-phase summation of the responses of these sources produced a peak around 40 Hz, and an out-of-phase interference of these sources created a null point around 70 Hz. Apparent latency of the summed response of the two sources depended on the relative amplitudes of cortical and brainstem sources. The summation of responses led to longer apparent latencies when the cortical source dominated and shorter apparent latencies when the brainstem source had larger amplitudes.

Adapting from FFR studies employing source modelling techniques, Bidelman, Jennings, & Strickland (2015) recorded 64 channel speech-evoked FFR to localize the generators for speech-evoked FFR. The results of this study were consistent with a midbrain (IC) origin for frequencies above 80 Hz. However, one study suggested that human auditory cortex might contribute to the high modulation frequencies when FFRs are recorded with magnetoencephalography and elicited at voice pitch with the fundamental frequency of 100 Hz (Coffey, Colagrosso, Lehmann, Schönwiesner & Zatorre, 2016). Their results suggested that the contribution of subcortical nuclei including CN, IC and MGB accounted for only 10% of the response variance (Coffey
et al., 2016). However, a recent multichannel study of FFRs to speech via EEG recording confirmed the presence of multiple FFR generators that include bilateral AN, brainstem IC, and bilateral primary auditory cortex, and showed that the relative contribution of these neural sources varies with the stimulus frequency (Bidelman, 2018). The AN and brainstem sources produce strong FFRs up to the sixth or seventh harmonics of speech around 600 to 700 Hz, whereas primary auditory cortex accounted for encoding stimulus frequency up to the speech fundamental frequency ($F_0$) around 100 Hz and showed dramatically weaker phase-locking beyond $F_0$. In the case that speech contains low-pitch energy around 100 Hz, only 10% of scalp-recorded FFRs is due to cortical phase-locked activity and more than 60% of the response is dominated by phase-locked responses of subcortical origins (Bidelman, 2018).

In summary, ASSRs, and more generally EFRs, are a series of overlapping responses from multiple neural generators along the auditory pathway. The contribution of the auditory cortex is limited to stimuli with low modulation frequencies (<80 Hz) with an approximate latency of 29 ms, while subcortical generators, especially the IC, are responsive to stimuli with high modulation frequencies (>150 Hz) with an approximate latency of 7.3 ms. Each neural generator has a latency that increases as we ascend the auditory pathway. The relative phase of phase-locked responses elicited from these neural generators determines if these responses destructively or constructively interfere to create the aggregate response at the scalp. The outcome of these interactions correspond with response peaks of higher amplitude and valleys with lower amplitude.

### 1.4 EFR phase interactions

The human auditory system has a tonotopic organization (Békésy, 1949). This fundamental coding characteristic of the auditory system starts from the cochlea and is preserved up to the auditory cortex. At the level of the cochlea, mechanical properties of the basilar membrane vary along the length of the cochlea, which is the basis for basilar membrane frequency selectivity. The base of the basilar membrane with greater stiffness and narrower width is most sensitive to high frequencies, and the apex with less stiffness and greater width is most sensitive to low rates. The peripheral auditory system has been visualized as having a series of overlapping band-pass filters with different center frequencies and level dependent bandwidths. When a complex sound like a broadband stimulus enters the cochlea, the tonotopic representation of the basilar
membrane separates the frequency components of the sound. Each peripheral filter on the basilar membrane responds to a portion of the sound that falls within its critical band in a way that high-frequency components activate auditory filters at the base and low-frequency components activate auditory filters at the apex. Because of this tonotopic arrangement of the basilar membrane, more time is needed for low-frequency components of a sound to reach their place of maximal activation at the apex. Therefore, there is a phase delay from the base to the apex across frequencies. The basilar membrane traveling wave latency across different frequency bands is estimated to be approximately 3 ms from 10-kHz to 1-kHz, and 5 ms to travel from 1 kHz to 250 Hz (Eggermont, 1979; Schoonhoven et al., 2001).

A review of the literature on electrophysiological measurements of auditory temporal acuity reveals that studies mostly have used a broadband noise carrier because it has a broader spectrum, which activates more auditory nerve fibres over a broad region of the basilar membrane. As a result, a broadband noise carrier generally elicits larger responses compared to tonal stimuli (Picton, John, Dimitrijevic and Purcell, 2003; John, Lins, Boucher and Picton, 1998; John et al., 2003). However, EFR characteristics might be influenced by the bandwidth characteristics of the broadband stimulus. In support of this hypothesis, the steady-state responses evoked by BBN stimuli have been shown to be lower in amplitude than the arithmetic sum of responses evoked by low-pass noise and high-pass noise stimuli in the same frequency range as the BBN stimulus (John et al., 2003). Zhu, Bhardwaj, Xia, and Shinn-Cunningham (2013) reported similar findings, in which EFRs were recorded in response to four different harmonic complex tones comprised of resolved and unresolved harmonics as well as a broadband stimulus simultaneously containing the same resolved and unresolved harmonics at a single fundamental frequency \( f_0 \). The broadband stimulus was expected to activate cochlear channels tuned to low, mid and high frequency components of the harmonic complex, and consequently the total EFR might be similar to the sum of responses to individual narrowband harmonic complex tones in terms of its amplitude. However, EFRs to the broadband stimulus were lower in amplitude than the sum of responses to individual low, mid and high-frequency harmonic complex tones. For a complex signal, the low-frequency harmonics of the fundamental frequency are resolved as they are processed within individual auditory filters, and high-frequency harmonics are considered unresolved because multiple harmonics are closely spaced and processed within wider
auditory channels where the output of each auditory channel reflects interaction of multiple high-frequency harmonics (Moore, 2003). Moreover, EFRs evoked by natural vowels differed significantly in amplitude across different vowels in the absence of intensity or bandwidth differences (Aiken and Picton, 2006; Choi, Purcell, Coyne and Aiken, 2013). These outcomes could be explained by phase interactions between neural responses evoked from low and high-frequency components within the broadband noise and complex tones, or low and high-frequency formants of speech vowels. Scalp-recorded EFRs to a broadband sound are the sum of all neural responses elicited at the same envelope frequency from stimulated auditory channels with their varying cochlear response delays. Across-channel delay can cause delayed onset of responses between auditory channels and as a result, lead to concurrent responses with different phases at the measurement electrodes. Therefore, the magnitude of the overall EFR can vary depending on the constructive or destructive superposition of responses initiated from different frequency regions with different phases. Applying a phase delay to the high frequency components of a broadband stimulus might compensate for the cochlear traveling delay, leading to a more synchronized initiation of neural responses from different frequency bands and constructive interference of these concurrent responses at the level of the scalp electrodes. In accordance with this idea, a recent study evoked EFRs with two tone pairs that represented the first and second formant frequencies of vowels. By varying the relative phase between first and second formants, the amplitude of the EFR at the fundamental frequency changed. The highest amplitude EFR was obtained when the envelope phase delay imposed on the second formant tone pair compensated for the phase difference between concurrent EFRs initiated from the first and second formant tone pairs (Easwar, Banyard, Aiken, & Purcell, 2018a). The influence of phase interactions between EFRs initiated by broadband vowels was further investigated and confirmed the importance of formant frequency characteristics (Easwar, Banyard, Aiken, & Purcell, 2018b).

According to the result of Easwar et al. (2018b), broadband stimuli like vowels may produce small EFRs at the measurement electrodes because of the out-of-phase interaction of EFRs elicited by the multiple frequency components of a broadband stimulus. Therefore, at the level of the measurement electrodes, responses initiated from one region of the tonotopic cochlea can be cancelled out by responses initiated from another frequency region because of the difference in travelling wave delay that leads
to fibers in the two regions firing out of phase. Previous research using amplitude modulated broadband noise (Purcell et al., 2004) may have been affected by cancellations of evoked potentials at the measurement electrodes due to interference from multiple parallel responses initiated along the length of the tonotopic cochlea.

These out-of-phase interactions manifest in the EFR modulation transfer function as frequency regions with low amplitudes that can be described as valleys, dips, or fine structure. The presence of valleys, dips or fine structure in the EFR modulation transfer function has been reported in previous studies in some individuals and for specific stimulus conditions using broadband noise carriers (Purcell et al., 2004). Thus we hypothesized that employing more frequency specific carriers would confine response initiation to frequency-specific regions of the cochlea and consequently reduce the degree of destructive interactions between concurrent EFRs at the level of the scalp.

1.5 Purpose of thesis

This thesis examines the possibility of using NBN carriers to investigate the individualized responses evoked from different frequency regions of the basilar membrane and to determine how these individual regions might affect the amplitude modulation threshold. Several behavioural studies have examined thresholds for amplitude modulation detection with frequency-specific carriers such as pure tones and NBN carriers (Dau, Kollmeier, & Kohlrausch, 1997b; Eddins, 1999; Eddins, 1993; Strickland, 2000; Strickland & Viemeister, 1997). Narrow band noise is characterized by a center frequency and a bandwidth and has its energy distributed over a relatively small region of the cochlea. The NBN carriers with different center frequencies enable recording of EFRs from low, mid, and high-frequency regions. The central hypothesis of this thesis is that multi-frequency NBN stimuli that are amplitude modulated will elicit EFRs at rates that are similar to those detectable behaviorally.

This thesis project aims to record envelope following responses (EFRs) to amplitude modulated NBN carriers with different center frequencies. The modulation frequency will be gradually changed over time and the measured electrophysiological response will be compared with behavioural measures in normal hearing participants. The research questions that are addressed in this study are:
- Do EFR amplitude versus modulation frequency patterns differ for NBN carriers as compared to broadband noise?

- Can EFR patterns evoked by NBN carriers be optimized to serve as a neural correlate of behavioural temporal acuity?

To answer these research questions, the specific approaches used here are: 1) To record EFRs to amplitude modulated NBN carriers with low, mid, and high center frequencies; 2) To compare EFRs elicited by amplitude-modulated NBN carriers with different center frequencies with one another; 3) To determine the correlation between behavioural and EFR measurements of temporal acuity; 4) To compare the number of valleys, dips, or fine structure across the EFR modulation transfer function evoked by NBN carriers with previous studies using broadband noise.
Chapter 2

2 General method

This study used a mixed design. All participants were enrolled in a behavioural procedure to find their amplitude modulation detection threshold and an electrophysiological procedure to measure the envelope following response (EFR). Figure 1.2 shows a summary of this study protocol.

![Figure 2.1: Study protocol]

2.1 Ethics approval

Approval of this project was secured from the University of Western Ontario, Office of Research Ethics (Appendix A). Participants were asked to read the study Letter of Information and sign a consent form for study participation before their enrolment in the study (Appendix B).
We placed the approved recruitment poster on the Western University campus. Also, we used existing departmental bulk email lists (by departmental permission) and the Department of Psychology summer participant pool.

2.2 Participants

A total of 67 adults (56 female, 11 male) between the ages of 18 to 33 years (mean (SD) = 24.4 (3.51) years) were recruited for this study. An upper age limit of 35 was chosen to exclude adults who may have experienced small decreases in auditory processing ability that are known to be associated with normal aging and would be undetectable by pure-tone audiometry (Ross, Fujioka, Tremblay, & Picton, 2007).

The following sequence was used for each participant after informed consent had been obtained. A brief questionnaire was filled out by the participant to record information about age, handedness, as well as self-reporting of any known hearing, vision, speech and language, or neurological problems. Otoscopy then was performed to visually determine that the ear canal and middle ear were normal. Using a ten dB-down, five dB-up bracketing technique, pure tone audiometric thresholds were obtained bilaterally at the octave and inter-octave frequencies between 250 and 6000 Hz with a MADSEN ITERA clinical audiometer. Signals from ITERA were presented through TDH39 audiometric headphones placed on the ears. Individuals who did not have normal hearing thresholds (≤20 dB HL) as assessed with the pure tone audiometry, were excluded. If any individuals were identified as potentially having a hearing impairment, they were encouraged in an appropriately compassionate manner to seek a professional assessment from audiologist. Information was provided about obtaining an evaluation at Western’s H.A. Leeper Speech & Hearing Clinic in the same building as the study (Elborn College). Three individuals were identified as having impacted cerumen impact and were encouraged to visit a registered audiologist or family physician. Participants were separated into three randomized groups: Low-frequency NBN (n = 22), mid frequency NBN (n = 22), high-frequency NBN (n = 23). This grouping allowed us to evaluate the effect of stimulus characteristics on EFRs while keeping an individual’s EFR measurement time under two hours.
2.3 Stimuli

We used amplitude modulated NBN of low, mid, and high frequency relative to the speech range in each group of participants to determine whether our participants had similar psychophysical performance for different NBN stimuli and whether different NBN stimuli elicit similar EFRs. Each group had a specific NBN carrier unique to that group, with common amplitude modulation, depth of modulation, and stimulus level presentation across groups. For every group, the modulation frequency ranged from 80 to 600 Hz at a fixed modulation depth of 50 percent and the stimuli were presented at the level of conversational speech at 60 dB SPL. The low-frequency NBN carrier was centred at 750 Hz with the lower spectral edge, $f_l$, at 150 Hz, and the higher spectral edge, $f_h$, at 1350 Hz. The mid-frequency NBN carrier was centred at 2200 Hz with the lower spectral edge, $f_l$, at 1600 Hz, and the higher spectral edge, $f_h$, at 2800 Hz. Finally, the high-frequency NBN carrier was centred at 3700 Hz with the lower spectral edge, $f_l$, at 3100 Hz, and the higher spectral edge, $f_h$, at 4300 Hz. The bandwidth for all low, mid, and high-frequency NBN was 1200 Hz. The selection of the NBN parameters was challenging because the behavioural listening tasks for estimating temporal acuity can be influenced by the spectral properties of the signal. Therefore, to obtain a true TMTF, reflecting sensitivity to the temporal envelope, discrimination of amplitude modulation should be made based on temporal cues, and not on any other cues in the amplitude-modulated stimulus.

During stimulus development, we discovered participants had inconsistently better sensitivity at higher modulation frequencies for amplitude modulated pure tones and NBN carriers compared to broadband noise. A likely cause for this problem was that participants were detecting the sideband frequencies that are produced by amplitude modulation. Different methods of stimulus generation have been used in the literature to reduce the influence of spectral cues that stem from resolving of sideband frequencies above and below the spectral edges of the carrier frequency. Carrier bandwidth must be selected to be at least twice the highest modulation frequency in order to eliminate the possibility of sideband frequencies being perceived as a cue by the listener. The modulation transfer function has low-pass filter characteristics in the absence of sideband cues (Dau et al., 1997; Eddins, 1993). We set the bandwidth of NBN carriers at 1200 Hz, which is the minimum bandwidth required at our highest rate of 600 Hz. Our maximum modulation rate was chosen to avoid further behavioural complexities.
that can occur at very high rates due to intensity cues. We performed multiple listening experiments to adjust our stimulus parameters. In our first listening experiment, we filtered the unmodulated broadband noise carrier (BBN) to make the unmodulated NBN and then applied amplitude modulation with a presentation level of 60 dB SPL. Some participants were able to use increasing sideband width as a spectral cue with an increase in modulation rate. One suggested solution is a filtering-after-modulation technique (Dau et al., 1997; Eddins, 1993). However, the technique of filtering-after modulation reduces the effective modulation depth substantially, having the most significant effect for high-modulation frequencies and NBN with narrower bandwidths (Eddins, 1999). In our experiment, even after employing filtering-after-modulation technique some participants still detected a changing combination tone as modulation frequency increased. A solution to reduce detection of combination tones is to mask distortion components by adding a low-pass noise to the modulated stimuli (Strickland, 2000; Strickland & Viemeister, 1997). By considering this solution, we evaluated filter-before-modulation stimuli while adding a low-pass masking noise of 50 dB SPL; however, some participants were still able to use a high-frequency sideband cue. Previous studies showed that adding low-pass and high-pass masking noise limits the listening regions below and above the test stimuli (Strickland, 2000; Strickland & Viemeister, 1997). Therefore, we added both low-pass and high-pass masking noise commencing 50 Hz beyond the sideband limits and evaluated the effect of high-pass masker bandwidth. A narrow high pass masker of 1200 Hz bandwidth was found to be most effective at eliminating upper spectral sidebands cues because the masking energy was concentrated near the upper sideband rather than spreading to very high frequencies.

To sum up, in the current study the technique of modulation after filtering was applied to make AM NBN stimuli, a notched-noise masker was used at 50 dB SPL, and the bandwidth of the high-frequency masker was 1200 Hz. For the EFR stimulus, the notched-noise masker was deemed not necessary since behavioural responses were not obtained during the EFR recording and combination tones were considered too low level to impact EFR initiation. Table 2.1 summarizes our final stimulus parameters.
Table 2.1: Summary of stimuli parameters

<table>
<thead>
<tr>
<th>NBN</th>
<th>Center Frequency (Hz)</th>
<th>Bandwidth (Hz)</th>
<th>Low Edge (Hz)</th>
<th>High Edge (Hz)</th>
<th>Stimulus Level (dB SPL)</th>
<th>Modulation Frequency (Hz)</th>
<th>Notched noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Low Band (Hz)</td>
</tr>
<tr>
<td>Low</td>
<td>750</td>
<td>1200</td>
<td>150</td>
<td>1350</td>
<td>60</td>
<td>80-600</td>
<td>10-100</td>
</tr>
<tr>
<td>Mid</td>
<td>2200</td>
<td>1200</td>
<td>1600</td>
<td>2800</td>
<td>60</td>
<td>80-600</td>
<td>10-1550</td>
</tr>
<tr>
<td>High</td>
<td>3700</td>
<td>1200</td>
<td>3100</td>
<td>4300</td>
<td>60</td>
<td>80-600</td>
<td>10-3050</td>
</tr>
</tbody>
</table>

2.4 Procedure

All testing was completed at the Speech, Auditory Feedback, and Electrophysiology Research (SAFER) Laboratory of the National Centre for Audiology in Elborn College at Western University. Participants attended for a maximum of 2 hours on a single day. At the completion of testing, each participant was compensated at the rate of $5 per half hour or part thereof as a thank you for their time.

2.4.1 Behavioural measurement protocol

2.4.1.1 Signal Generation

A type 4157, Bruel and Kjaer ear simulator and Bruel and Kjaer Type 2250 sound-level meter were used to calibrate the stimuli. Sounds were presented monaurally to a randomly selected ear via an Etymotic Research ER-2 insert earphone, which was secured in the ear with a foam ear tip. Signals were presented at 60 dB SPL. The auditory stimuli were stored as wave files, which were played through the sound card of an IBM ThinkPad computer.

2.4.1.2 Behavioural response recording

Following these initial procedures, a modulation detection threshold task was carried out to evaluate temporal acuity behaviorally in a sound attenuated booth. This test usually took about 20 minutes to complete and was done before the electrophysiological
measurement because it required attention. In this study, the auditory stimuli were presented through a forced choice paradigm, and the target stimulus value was selected adaptively. In psychophysical adaptive techniques, the level of the signal’s property of interest in each trial is determined by the response of the listener in the previous trial. Thresholds were estimated using an adaptive three-interval two-alternative forced-choice (3I-2AFC) AXBD procedure with feedback. In this procedure, three intervals were presented, where the middle interval (X) was the standard stimulus, and it matched with the stimulus in either the first (A) or the last (B) interval on a random basis. The participant’s task was to judge whether (A) or (B) contained the stimulus that is different from (X). “D” in the AXBD method refers to the selection of the stimulus different from the middle interval rather than the same sound as the middle interval. The order of the target (the different) stimulus was selected randomly. The AXB design eliminates some of the disadvantages of other psychophysical approaches and guides the listener’s attention to the property on which A and B are different. The AX design is a difficult task to perform as listeners often confused about what property of a sound they are supposed to listen in the discrimination task. This confusion increases the bias in comparing two sounds with same or different quality without having a reference to be compared to (Mengler, Hogben, Michie, & Bishop, 2005).

In contrast, the AXB design is not affected by the listener’s overall bias in making a same-different comparison, which occurs for the 2AFC AX designs according to signal detection theory (Macmillan, Kaplan, & Creelman, 1977). Also, the AXB design has an advantage over other 3-interval methods (3AFC oddball) of reducing the need to rely on short-term memory because only neighbouring intervals are compared as the stimulus varies at either the A–X edge or the X–B edge. In this study, each trial involved the presentation of three intervals including one target and two standard signals that contained 500 ms NBN carrier with a 500 ms inter-stimulus interval (gap between intervals within each trial). The standard stimulus was an unmodulated NBN, which was matched with either the first or the last interval, and the target stimulus was a 50% amplitude-modulated NBN stimulus. Either the first or the third interval was randomly selected to be amplitude-modulated. Both modulated and unmodulated stimuli had equal duration and power. The listener was seated in front of the computer screen and was instructed to identify whether either the first or the last interval contained the sound (target) that was different from the interval in the middle. As each sound was heard, a
coloured box was highlighted as an animated dinosaur jumped on it. The listener responded by clicking the mouse on the box that the listener thought was the target stimulus (Sutcliffe & Bishop, 2005). A correct response was rewarded by having an image pop up in the left corner of the screen. All participants received training trials before the actual testing started. In the training trials, the modulation frequency adaptively varied from 20 Hz. For training, we selected lower modulation frequencies compared to the main study because these are easier to identify and gives the participant a better understanding of the sound parameter they are looking to identify. Participants were asked to guess when they were not able to identify the difference. Participants could take a break at any time. There was an option of restarting a trial if a listener reported that he or she needed to repeat that trial due to an unexpected redirection of attention (e.g., a cough).

In this study, the Virulent PEST (Parameter Setting by Sequential Estimation) (Findlay, 1978) adaptive psychophysical procedure was employed for targeting the 79% correct point on the psychometric function. PEST (Taylor & Creelman, 1967) is the set of rules that are used for adjusting the difficulty level of the discrimination to lead to a threshold value. Unlike the adaptive staircase technique (Levitt, 1971), PEST uses different step sizes as the test proceeds. The target stimulus has the same properties as the standard stimulus except in one parameter. Initially, the target stimulus has a substantial difference from the standard stimulus in a parameter, and as it is manipulated systematically, this difference progressively decreases until an incorrect response is made. At this point, the discrimination is made easier. After a certain number of correct responses which is determined by the expected percent correct value, the step size is halved. Reversals in adaptive methods are points at which an increase in signal level is followed by a decrease (lower turning point), or a decrease in signal level is followed by an increase (Levitt, 1971). The step size is fixed during the first few reversals and changes after the predetermined number of reversals. Testing continues with the difficulty level of the signal parameter stepping up or down depending on the listener’s response, until a threshold (in this study 79% correct) is reached. The Virulent PEST algorithm stops when the number of trials exceeds the maximum number of trials as specified by the investigator, or when a specific number of reversals had occurred, whichever occurs first.
In this study, the initial interval started at the modulation frequency of 80 Hz (the lower the modulation frequency, the easier the discrimination will be). If the listener was not able to identify the target stimulus at the lowest level of difficulty (80-Hz), the target signal was presented at the same difficulty level until a correct response was seen. Initially, using a one up two down (2-step) rule, the level of difficulty increased for every two successive correct responses and decreased per erroneous response or a sequence of one correct response followed by an incorrect response. Beyond the third reversal, the rule was one up three down (3-step), where the level of difficulty increased for every three successive correct responses and decreased for each incorrect response or a sequence of zero, one, or two correct responses followed by an incorrect response. A maximum step size of 10 (100 Hz) and minimum step size of one (10 Hz) were used, and the maximum number of trials was 80. The initial step size was 10 (100 Hz). Beyond the fifth reversal, the step size reduced to five (50 Hz). Beyond the sixth reversal, the step size reduced to two (20 Hz) up to the seventh reversal at which the step size reduced to one (10 Hz). If the listener was able to identify all the presented intervals at the highest level of difficulty (600 Hz), the target signal was presented at the same difficulty level until an error occurred. Bracketing continued after the seventh reversal until there were three successive correct responses. The third correct response is called the eighth reversal, and threshold estimation stopped. There was a total of eight reversals for each threshold estimation. Data from trials before the fourth reversal were discarded, and an estimate of modulation detection threshold was determined as the average of the levels from the fourth reversal onwards. In the classic PEST method, tracks terminate when the step size reached a predetermined minimum size, and the estimate of threshold was the last presented level, which meant the complete history of the run did not need to be evaluated or stored (Taylor & Creelman, 1967). This method is slightly less precise than the virulent PEST’s averaging method. In our study, precision was deemed necessary, so we used the averaging method.

2.4.2 Electrophysiological measurement protocol

2.4.2.1 Signal generation

In this study, the methodology for stimulus generation and recording was previously developed and implemented by Purcell et al. (2004). The swept AM stimuli varied in modulation frequency from 80 to 600 Hz across each stimulus sweep. Each stimulus
sweep consisted of 30 epochs of 1.024 seconds within a total duration of 60 minutes. During the first half of each sweep, the modulation rate was swept linearly upward from 80 Hz to reach the maximum preset modulation frequency of 600 Hz. In the second half of the sweep, the modulation frequency swept downward to reach the starting point (80 Hz). Sweeps were linked together so that they repeated without discontinuity (Purcell et al. 2004). Acoustic stimuli were generated by a custom program developed using LabVIEW (Version 8.5; National Instruments, Austin, TX). During EFR testing, digital-to-analog conversion of the stimulus and analog-to-digital conversion of the EEG were executed using a National Instruments PCI-6289 M-series acquisition card at 4,000 samples per second with 16-bit resolution for digital-to-analog conversion of the stimulus, and 18-bit resolution for analog-to-digital conversion of the output signal. The intensity of the stimuli was adjusted to 60 dB SPL by a Tucker-Davis Technologies PA5 attenuator and an SA1 power amplifier. Calibration of the EFR stimulus level was performed with a Brüel and Kjær Type 2250 sound level meter in a flat-weighted $L_{eq}$ as the stimulus was presented to a Type 4157 ear simulator.

### 2.4.2.2 EFR recording

The electroencephalogram (EEG) was recorded from a single-channel by placing three disposable Medi-Trace Ag/AgCl electrodes on the participant at vertex (Cz) as the non-inverting electrode, just below the hairline at the posterior midline of the neck as the inverting electrode, and one of the collarbones as the ground electrode. The skin at the place of the electrode sites was first cleaned with an alcohol prep pad. Then, each electrode site was scrubbed with Nuprep gel and GRASS EC2 conductive cream was applied to the underside of each electrode to improve electrical contact with the skin. For each participant, electrode impedances were measured using an F-EZM5 GRASS impedance meter at 30 Hz, before and after the EEG measurements. Impedances were required to be under 5 kΩ and with inter-electrode differences of less than 2 kΩ. In the case that we did not meet the criteria beforehand, we re-cleaned the electrode sites to reach the optimal values. The electrophysiological measurement was also performed in the electromagnetically shielded and sound attenuated booth. During the experiment, participants sat in a comfortable chair that could be reclined. A rolled towel was placed under the neck of the participant to reduce muscle activity. Also, participants were provided with a blanket for more comfort. Subjects were encouraged to sleep during the test and were instructed to ignore the stimulus. The light was switched off during
the test. During the recording, the inverting, non-inverting and ground electrodes were connected to a Grass LP511 AC amplifier with the band-pass filter setting between 3 and 1000 Hz and a gain of 50,000. An additional gain of 2 was applied by the PCI-6289 card for a total gain of 100,000. The recorded EEG was shown to the operator so she could ascertain data quality as a time-series voltage-signal in 1.024 s windows, as well as a chart displaying a noise metric and the 60 Hz component. The AM NBN stimulus was presented monaurally (to the same ear as the behavioural measurement) via an electromagnetically shielded Etymotic Research ER-2 insert earphone which was sealed securely in the ear canal with a disposable foam tip.

2.4.2.3 Response detection and analysis
Real-time monitoring of the EEG measurement was possible during the data collection, but data analysis was performed offline after the recordings were completed. EFR offline analysis was conducted by using a custom LabVIEW 8.5 program with a similar method as Purcell et al. (2004). In the offline analysis, sweeps were averaged synchronously. First, a noise metric criteria was checked for all recorded epochs. Each epoch’s noise metric was calculated by determining the averaged amplitude of EEG and myogenic activity within the response frequency band between 80 to 115 Hz. Then, a mean and standard deviation (SD) of the noise metrics were calculated for all epochs. Any epoch with a noise metric higher than +2 SDs of the mean were excluded from further analysis. Epochs which passed the noise rejection criteria were averaged synchronously in time to create an averaged sweep that nominally contained EEG during 117 repetitions of the stimulus. To decrease the effect of myogenic noise on group EFR estimates as much as possible, EFR data from participants with average noise metrics plus two standard deviations of the noise metric > 1,000 nV were excluded from further analyses (7 subjects for High-NBN group), as their values were deemed highly influenced by myogenic artifacts. This rejection threshold was conservative; noise metrics were generally much lower (averaged across participants in each group: f_c= Low-NBN, mean = 219.62 nV, SD = 168.70; f_c= Mid-NBN, mean = 211.95 nV, SD = 154.79; f_c= High-NBN, mean = 204.44 nV, SD = 149.37).

The amplitude and the phase of the EFR at each modulation frequency were estimated using a Fourier analyzer (Regan, 1989; Purcell et al., 2004). In this method, sine and cosine sinusoids generated from the instantaneous modulation frequency of the stimuli were used as a reference signal. The averaged sweep was corrected by 10 ms for the
estimate of brainstem processing delay to align the response and the reference sinusoids (Purcell et al., 2004). The average EEG sweep was multiplied with the references to produce real and imaginary components of the EFR which were low-pass filtered with two rectangular window (1.024 s) moving averages in series. As each stimulus sweep compromised two halves with reversed modulation rate, the second half of the complex response sweep was folded, and vector averaged with the first half. To determine whether EFR amplitude was statistically greater than a background EEG noise estimate, the amplitude of the EFR and noise estimate were compared at each modulation frequency using an F-ratio. The noise estimate was calculated from the response sweep folded and averaged in the time domain using a discrete Fourier transform (DFT). The noise estimate was calculated by considering +/- 60 DFT frequency bins above and below the EFR frequency. The EFRs were considered significantly different from the background noise estimate when the amplitude of the EFRs was 1.75 times greater than the amplitude of the noise using an F statistic with degrees of freedom of 2 and 240 at the p < .05 level. The highest modulation frequency at which EFR amplitude was significantly different from the EEG noise estimate was estimated for comparison with behavioural thresholds and was considered as the EFR threshold.

This EFR threshold, or EFR maximum significant frequency, was determined similarly to Purcell et al. (2004). A phase-weighted amplitude approach was used as the measured EFR amplitude never reaches zero due to noise passing through the FA, whereas the phase-weighted amplitude is driven toward 0 nV. For each subject, a reference modulation frequency band was selected from regions that the EFR amplitude was statistically different from noise (p < .05) using the F-test. The mean response phase slope in the reference frequency band was then determined. The response phase slopes of neighbouring modulation frequencies were also calculated. Those neighboring areas with phase slopes within 2 SD of the reference frequency band were also included in the phase analysis. A single line was fitted to the phase analysis band to estimate the EFR source delay by assuming a single source dominated the measured EFR. The model phase was subtracted from the EFR phase data to flatten the phase response. Then, the circular mean of the flattened phase was used as an expected phase. The phase-weighted amplitude was calculated by projecting the output of the FA onto the expected phase. The EFR amplitude at each modulation frequency was multiplied with the cosine of the difference between the expected phase and the phase of EFR to
calculate the phase-weighted amplitude. The highest modulation frequency at which the EFR could be detected was taken as the first modulation frequency beyond the phase analysis band where the phase-weighted amplitude was not significantly different from zero as evaluated with a t-test.

2.4.2.4 Stimulus artifact check

The possibility of stimulus artifact was evaluated by recording EEG with the same recording conditions as the EFR measurement with the electrodes placed on a participant while the stimulus was presented to a Zwislocki coupler. From each group of low, mid and high-frequency narrow-band noise carriers, we recruited one participant to complete the stimulus artifact check. Additionally, we used a phantom head by placing electrodes in water while the stimulus was presented again to the Zwislocki coupler. In both conditions, EEG was analyzed offline with the same methodology as the main experiment. In all conditions, a significant response detection was considered as a false positive response. In the human participant condition, the maximum amplitude at the response frequency in the range 80-600 Hz was 16.96, 13.44, and 14.06 nV for low, mid and high-frequency NBN respectively. The prevalence of false positive detections were 4.58, 7.5, and 6.66 percent for low, mid and high-frequency NBN, respectively. In the phantom head condition, the maximum amplitude of the responses were 2.43, 2.73, and 2.80 nV for low, mid and high-frequency NBN, respectively. For low-frequency NBN, none of the recorded responses reached the statistical significance level. The rate of false positive response occurrence was 1.66 % and 5.8 % for the mid and high-frequency NBN carriers, respectively. The observed false positive rates in humans are similar to the expected 5%, and the maximum amplitudes are typical of background myogenic and EEG noise in the average sweep. The amplitudes observed in the phantom conditions were very low. It is therefore unlikely that stimulus artifact had a significant influence on the EFR amplitudes for the different NBN carriers.

2.4.2.5 Fine structure selection

As the central research question of this thesis project, we hypothesized that employing NBN carriers would confine response initiation to frequency-specific regions within the cochlea and consequently reduce the degree of destructive interactions between concurrent EFRs at the level of the scalp. To examine this hypothesis, the number of
valleys across the EFR amplitude modulation spectrum evoked by our NBN stimuli were compared with previous studies using broadband noise.

Our analysis counted the total number of dips or fine structure in the EFR modulation transfer function that had a pattern suggestive of out-of-phase interactions between EFRs. An out-of-phase interaction manifests as a relatively deep valley in EFR amplitude as modulation frequency increases. These valleys are distinct from the gradual decrease in EFR amplitude that occurs with increasing modulation frequency (Purcell and Dajani, 2008). A MATLAB algorithm was created to objectively select the dips and the modulation frequency of their occurrence in the EFR modulation transfer function. Only dips that occurred with the following conditions were deemed indicative of out-of-phase interactions. First, the minimum distance between neighbouring peaks and valleys should be at least 20-Hz. Second, a 50-Hz frequency separation should exist between two consecutive valleys or consecutive peaks. Finally, the depth of a valley relative to its neighboring peaks must be a minimum of 5 nV. This objective selection methodology was used to determine the number of valleys in the EFR modulation transfer function for the multi-frequency AM NBN stimuli in young normal-hearing adults in this study, an AM BBN carrier in older and younger adults in Purcell et al. (2004) and an AM BBN stimulus in older subjects in one unpublished study.

2.5 Statistical analysis

Statistical analyses were performed using Rstudio (version 1.1.463). Using correlation analyses, the non-parametric Spearman’s rank correlation coefficient was computed to determine if the electrophysiological measurement as assessed by EFR to AM NBN carriers with different center frequencies could serve as a predictive index of the behavioural responses. A Kruskal-Wallis test was conducted to examine any differences between EFR thresholds across different NBN stimuli. Similar non-parametric tests were used to determine the differences between behavioural thresholds across different AM NBN stimuli. These tests were used to determine whether stimulus frequency might affect the amplitude modulation threshold either behaviorally or electrophysiologically. A Wilcoxon Rank-Sum test for pairwise comparison identified groups with a significant difference. A Poisson regression model was used to assess if the type of NBN stimulus can explain the number of valleys in the fine structure of the envelope following responses. Details are described in the result section.
Chapter 3

3 Results

3.1 EFRs

Data from seven participants in the high-frequency NBN group were excluded from our analysis because of artifact rejection threshold > 1000 nV. One participant from mid-frequency, two participants from high-frequency, and three participants from low-frequency NBN groups were excluded because only 25% or less of their EFRs were significantly different from noise over the modulation frequencies of 80-600 Hz. Also, data from one participant from the low-frequency NBN group was excluded with no detectable EFRs across the entire modulation frequency range. In total, data from 22 participants in the low-frequency NBN group, 22 participants in the mid-frequency NBN group, and 23 participants in the high-frequency NBN group were included in the study analysis. Table 3.1 summarizes the average data for age, pure-tone hearing threshold of the tested ear at octave and inter-octave frequencies between 250 and 6000 Hz, behavioural AM detection threshold, and EFR thresholds across different NBN groups.

Table 3.1: Summary of age, pure-tone hearing thresholds, psychophysical AM detection thresholds and EFR maximum significant frequency across three groups

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Age</th>
<th>Pure-Tone Threshold (dB HL)</th>
<th>AM Detection Threshold (Hz)</th>
<th>EFR Maximum Significant Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low - NBN Group</td>
<td>22</td>
<td>23.64 (3.58)</td>
<td>4.67 (3.41)</td>
<td>204.18 (78.4)</td>
<td>342.82 (162.08)</td>
</tr>
<tr>
<td>Group Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid - NBN Group</td>
<td>22</td>
<td>25.05 (2.8)</td>
<td>4.37 (3.5)</td>
<td>351.18 (152.54)</td>
<td>432.82 (119.87)</td>
</tr>
<tr>
<td>Group Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High - NBN Group</td>
<td>23</td>
<td>24.52 (4.04)</td>
<td>3.86 (2.5)</td>
<td>293 (152.45)</td>
<td>407.39 (147.57)</td>
</tr>
<tr>
<td>Group Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2 Assumption check

Different statistical tests applied to our data. To determine the appropriate statistical test the probability that our sample came from a population that is approximately normally distributed was evaluated. Quantile-quantile (Q-Q) plots were used to plot the quantiles of the cumulative probability of the results of each NBN group against a normal distribution, and histograms were created. Inspecting the Q-Q plots, histograms, and the result of Shapiro-Wilk test (p < .05) showed that the data violated the assumption of normality. Moreover, the homogeneity of variances was tested with Levene’s test to determine if data from different groups had the same variance. As normality and equal variance were not observed for our data, non-parametric tests were used as our data violated these assumptions needed for employing parametric tests.

A Kruskal-Wallis H test was assessed to determine if the hearing threshold and age were different for three groups of participants. No significant difference was found between three NBN groups in terms of hearing thresholds, H (2) = 1.44, p = .485, df = 2, or age, H (2) = 2.96, p = .227, df = 2. Across all 67 subjects, the mean hearing thresholds and the associated standard deviations were 4.26 ± 3.08 HL for the right ear and 4.34 ± 3.24 HL for the left ear. Results of pairwise comparisons using the Wilcoxon Rank-Sum test showed that the left-right ear differences did not significantly affect pure-tone hearing thresholds, W = 584.5, p = .74, behavioral modulation detection thresholds, W = 594.5, p = .65, and the highest frequency at which the EFR was detected, W = 573, p = .855.

3.3 Effect of swept-frequency AM on EFRs amplitude

Results of this study showed that with young normal hearing individuals, the AM NBN of low (centred at 750 Hz), mid (centred at 2200 Hz), and high frequency (centred at 4300 Hz) at a modulation depth of 50% with modulation rate being swept linearly between 80-600 Hz can produce robust EFR responses that are significantly different from background EEG and myogenic noise. A peak amplitude of around 80-100 Hz has been described previously (Picton, John, Purcell, et al., 2003; Purcell et al., 2004). For some individuals, there were multiple non-significant areas in the EFR modulation transfer function which were followed by regions of significant response at higher frequency regions. Figure 3.1 shows the EFRs and their corresponding EEG-noise
estimates as a function of AM rates for all individuals per each NBN condition (n = 22 for low-frequency NBN group, n= 22 for mid-frequency NBN group, n= 23 for high-frequency NBN group). As can be seen there is quite a lot of individual variability. EFRs and EEG noise across different NBN conditions were averaged separately and are shown in Figure 3.2.

Figure 3.1: EFRs as a function of AM rates from 80-600 Hz for all participants for low, mid and high-frequency NBN groups.
Figure 3.2: EFRs and corresponding EEG background noise mean amplitude and standard error of the mean for different NBN carriers.
3.4 Effect of the different carrier frequencies on the EFR amplitudes

The EFRs mean amplitude across modulation rates of 80 to 600 Hz for low, mid and high-frequency NBN groups are plotted in Figure 3.3, and the corresponding EEG background noise estimate for low, mid and high-frequency NBN groups are plotted in Figure 3.4. In order to have a more detailed look at the effect of different types of AM NBN stimuli on EFR amplitudes, comparisons were made within multiple frequency bands ranging from 80-100 Hz, 100-200 Hz, 200-300 Hz, 300-400 Hz, 400-500 Hz, and 500-600 Hz. To do so, responses were averaged for each participant in the mentioned modulation frequency bands. The mean data from all participants were compared between different NBN groups across each frequency band using Kruskal-Wallis tests. A post-hoc analysis was performed using the non-parametric Wilcoxon Rank-Sum test to determine which NBN groups differed. Differences were considered as significant if the p-value < .05 for both Kruskal-Wallis and post-hoc tests after applying a Bonferroni correction.

The result of this statistical analysis showed EFR amplitudes were significantly affected by the type of NBN carriers only in the range of 100-200 Hz, H (2) = 11.432, p = .003, df = 2. Pairwise comparisons of the mean ranks between groups using the Wilcoxon rank-sum test showed that EFR amplitudes were significantly lower for the low-frequency NBN group (Mdn = 23.97) compared to the mid-Frequency NBN group (Mdn = 29.28), W = 114, p = .002, r = -.48. A similar significant difference was observed for EFR amplitudes of the low-frequency NBN group (Mdn = 23.97) compared to the high-frequency NBN group (Mdn = 31.42) where EFR amplitudes in the range of 100-200 Hz for the low-frequency NBN group were again significantly lower, W = 132, p = .005, r = -.43. However, EFR amplitudes for mid-frequency NBN group were not significantly different from the high-frequency NBN group (p = .45).

In all cases, the critical p level for significance (alpha = .05) was corrected for the number of comparison tests performed to control the familywise error rate using a Bonferroni correction. The type of NBN carrier did not significantly affect the amplitudes of EFRs for other modulation frequency bands (all p-values >.05). Table 3.2 and 3.3 summarise all the significant and non-significant results.
Figure 3.3: Grand average of EFRs and standard error of the mean for all retained participants for low, mid and high-frequency NBN carriers in the range of 80 to 600 Hz.

Figure 3.4: Grand average of EEG background noise and standard error of the mean for all retained participants for low, mid and high-frequency NBN carriers in the range of 80 to 600 Hz.
Table 3.2: Results from Kruskal-Wallis Rank-Sum test that evaluates whether a significant difference exists between NBN groups in terms of EFR amplitudes across different modulation frequency bands

<table>
<thead>
<tr>
<th>Subgroups being compared</th>
<th>Parameter being used</th>
<th>N</th>
<th>80-100-Hz Mean Amplitude (nV)</th>
<th>100-200Hz Mean Amplitude (nV)</th>
<th>200-300Hz Mean Amplitude (nV)</th>
<th>300-400Hz Mean Amplitude (nV)</th>
<th>400-500Hz Mean Amplitude (nV)</th>
<th>500-600Hz Mean Amplitude (nV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-NBN Group Mean (SD)</td>
<td></td>
<td>22</td>
<td>45.03 (16.69)</td>
<td>25.02 (8.27)</td>
<td>14.89 (7.48)</td>
<td>10.04 (4.57)</td>
<td>6.23 (3.39)</td>
<td>3.3 (1.34)</td>
</tr>
<tr>
<td>Mid-NBN Group Mean (SD)</td>
<td></td>
<td>22</td>
<td>41.32 (16.01)</td>
<td>34.25 (10.26)</td>
<td>17.18 (6.61)</td>
<td>9.66 (4.6)</td>
<td>6.41 (3.4)</td>
<td>2.97 (1.24)</td>
</tr>
<tr>
<td>High-NBN Group Mean (SD)</td>
<td></td>
<td>23</td>
<td>38.61 (16.06)</td>
<td>31.85 (7.25)</td>
<td>16.43 (5.69)</td>
<td>10.21 (4.35)</td>
<td>6.89 (3.27)</td>
<td>3.32 (1.58)</td>
</tr>
<tr>
<td>Low-NBN Group Median</td>
<td></td>
<td>22</td>
<td>39.77</td>
<td>23.97</td>
<td>14.12</td>
<td>9.58</td>
<td>5.41</td>
<td>3.16</td>
</tr>
<tr>
<td>Mid-NBN Group Median</td>
<td></td>
<td>22</td>
<td>39.92</td>
<td>29.28</td>
<td>16.68</td>
<td>9.88</td>
<td>5.89</td>
<td>2.96</td>
</tr>
<tr>
<td>High-NBN Group Median</td>
<td></td>
<td>23</td>
<td>35.71</td>
<td>31.42</td>
<td>16.46</td>
<td>9.4</td>
<td>5.7</td>
<td>3.18</td>
</tr>
</tbody>
</table>

H Statistics: 1.48  11.43  1.15  .12  .59  .45  
Significance: .477  .003**  .561  .941  .743  .797

* p<.05.
** p<.01.
*** p<.001.

Table 3.3: Results from Wilcoxon Rank-Sum test that identified whether the low, mid and high-frequency NBN groups were different in terms of the EFR amplitudes across the modulation frequency band 100-200 Hz.

<table>
<thead>
<tr>
<th>Subgroups being compared</th>
<th>Parameter being used</th>
<th>N</th>
<th>Wilcoxon test statistics (W)</th>
<th>Significance</th>
<th>Effect size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LowNBN-MidNBN</td>
<td>100-200-Hz mean amplitude (nV)</td>
<td>44</td>
<td>114</td>
<td>0.002**</td>
<td>-0.48</td>
</tr>
<tr>
<td>LowNBN-HighNBN</td>
<td>100-200-Hz mean amplitude (nV)</td>
<td>45</td>
<td>132</td>
<td>0.005**</td>
<td>-0.43</td>
</tr>
<tr>
<td>MidNBN-HighNBN</td>
<td>100-200-Hz mean amplitude (nV)</td>
<td>44</td>
<td>287</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>

* p<.05.
** p<.01.
*** p<.001.
3.5 Effect of different NBN carriers on AM detection threshold as assessed by EFRs

The purpose of this analysis was to determine the effect of carrier type on the highest frequency at which the EFR was no longer significantly different from the EEG-noise estimate which we have defined as the objective AM detection threshold or EFR threshold. According to the results from the Kruskal-Wallis rank sum test, no significant difference was found across different NBN carriers in terms of EFR maximum significant frequency, $H = 4.35$, $p = .113$, $df = 2$. The boxplots in Figure 3.5 summarize the result. The descriptive statistics for EFR threshold frequency are shown in Table 3.4.

![Boxplots of EFR thresholds across different NBN carriers. No statistically significant difference was found between different types of carriers.](image)

The boxplots in Figure 3.5 show the distribution of EFR thresholds across all participants for low, mid, and high-frequency NBN carriers. The boxes range from the 25th percentile in the bottom, and the 75th percentile at the top. The area inside the boxes represents the interquartile range (IQR; 50% of the data set are located in this range). The medians are shown by a line across the boxes and are the middle of the data sets. The bottom and top whiskers are extended from the 25th percentile and 75th percentile.
respectively to the maximum and minimum values in the data sets (which is less than the standard whisker range of 1.5 times the IQR).

Table 3.4: Results from Kruskal-Wallis rank sum test that evaluates whether the low, mid, and high-frequency NBN groups were different in terms of EFR maximum frequency.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>EFR maximum frequency Mean(SD)</th>
<th>Median</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-NBN</td>
<td>22</td>
<td>342.82 (162.08)</td>
<td>297.5</td>
<td>265</td>
</tr>
<tr>
<td>Mid-NBN</td>
<td>22</td>
<td>432.82 (119.87)</td>
<td>474.5</td>
<td>213</td>
</tr>
<tr>
<td>High-NBN</td>
<td>23</td>
<td>407.39 (147.57)</td>
<td>403</td>
<td>235</td>
</tr>
</tbody>
</table>

Kruskal-Wallis H Statistics 4.35
Significance .113

*p < .05.
**p < .01.
***p < .001.

3.6 Effect of different NBN carriers on behavioural AM detection threshold

Using the Kruskal-Wallis Rank-Sum test, the influence of carrier type on the behavioural AM detection threshold was assessed. The type of NBN carrier was significant at the highest detectable modulation frequency, $H = 8.91$, $p = .011$, $df = 2$. Pairwise comparisons of the mean ranks between groups using the Wilcoxon Rank-Sum test showed that behavioural thresholds were significantly lower for the low-frequency NBN carrier compared to the mid-frequency NBN carrier, $W = 114$, $p = .002$, $r = -0.47$. However, behavioural AM detection threshold did not differ significantly between the low-frequency and high-frequency NBN carriers, $W = 178$, $p = .09$, or between the mid-frequency and high-frequency NBN carriers, $W = 306$, $p = .233$. The boxplots in Figure 3.6 show graphical depictions of these differences. Tables 3.5 and 3.6 show significant and non-significant results.
Figure 3.6: Boxplots of behavioural AM threshold across different carriers. A statistically significant difference was found between different types of carriers.

The boxplots in Figure 3.6 show the distribution of behavioural AM thresholds across all participants for low, mid, and high-frequency NBN carriers. The boxes range from the 25th percentile in the bottom to the 75th percentile at the top. The area inside the boxes represents the interquartile range (IQR; 50% of the psychophysical thresholds are located in this range). The medians are shown by a line across the boxes and are the middle of the data sets. The bottom and top whiskers are extended from the 25th percentile and 75th percentile respectively to the maximum and minimum values in the data sets (which is less than the standard whisker range of 1.5 times the IQR).
Table 3.5: Results from Kruskal-Wallis Rank-Sum test that evaluates whether the low, mid, and high-frequency NBN groups were different in terms of modulation detection threshold (Hz).

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Behavioural Threshold Mean(SD)</th>
<th>Median</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-NBN</td>
<td>22</td>
<td>204.18(78.4)</td>
<td>211</td>
<td>107</td>
</tr>
<tr>
<td>Mid-NBN</td>
<td>22</td>
<td>351.18(152.64)</td>
<td>370</td>
<td>266</td>
</tr>
<tr>
<td>High-NBN</td>
<td>23</td>
<td>293(152.45)</td>
<td>220</td>
<td>274</td>
</tr>
</tbody>
</table>

Kruskal-Wallis Chi-Squared 8.91
Significance .011*

* p < .05.
** p < .01.
*** p < .001.

Table 3.6. Results from Wilcoxon Rank-Sum test that identified whether the low, mid, and high-frequency NBN were different in terms of modulation detection threshold.

<table>
<thead>
<tr>
<th>Subgroups Being Compared</th>
<th>Parameter Being Used</th>
<th>N</th>
<th>Wilcoxon Test Statistics (W)</th>
<th>Significance</th>
<th>Effect Size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LowNBN-MidNBN</td>
<td>Modulation Detection (Hz)</td>
<td>44</td>
<td>114</td>
<td>.002**</td>
<td>-.47</td>
</tr>
<tr>
<td>LowNBN-HighNBN</td>
<td>Modulation Detection (Hz)</td>
<td>45</td>
<td>178</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>MidNBN-HighNBN</td>
<td>Modulation Detection (Hz)</td>
<td>44</td>
<td>306</td>
<td>.233</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05.
** p < .01.
*** p < .001.

3.7 Association between EFR AM detection thresholds and behavioural AM detection thresholds

A correlation analysis was computed to assess the association between the psychophysical threshold for detecting 50% amplitude modulation of NBN carriers, and the electrophysiological threshold as the highest frequency at which the phase-
weighted EFR amplitude was no longer significantly different from noise, i.e. the EFR maximum frequency. A non-parametric Spearman rank-order correlation coefficient was calculated for low, mid and high-frequency NBNs separately. The result revealed that with young normal-hearing participants, no significant correlation was observed between behavioural and objective measures within each NBN group separately when the critical p level for significance was corrected for familywise error rate (low-NBN, rho=.46, p=.029; mid-NBN, rho=.1, p=.63; High-NBN, rho=.46, p=.025). The scatterplots in Figures 3.7, 3.8 and 3.9 summarize the correlation results.

![Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for low-frequency NBN carrier](image)

**Figure 3.7:** Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for low-frequency NBN carrier

![Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for mid-frequency NBN carrier](image)

**Figure 3.8:** Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for mid-frequency NBN carrier
Figure 3.9: Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for high-frequency NBN carrier

Grouping individuals across all three NBN noise carriers, there was a significant correlation between psychophysical and electrophysiological measures of modulation sensitivity (rho=.39, p=.001; Figure 3.10).

Figure 3.10: Scatterplot showing the correlation between the EFR and behavioural thresholds (Hz) for all subjects.
3.8 Valleys in the EFR modulation transfer function

3.8.1 Comparison between different types of narrow-band noise

Employing the fine structure selection criteria, fine structure valleys were not observed in the EFR amplitude modulation spectrum of 24 subjects out of a total of 67 participants whose data were analyzed in this study.

A Poisson regression model was fitted to assess if the type of NBN stimulus can explain the number of valleys in the fine structure of the envelope following responses. In this model, carrier type was a categorical predictor with the three levels of low-frequency (n = 22, Average counts = 1), mid-frequency (n = 22, Average counts =.81), and high-frequency (n= 23, Average counts =.78) NBN carriers. The low-NBN stimulus was selected as a reference category. According to this model, there does not seem to be a significant difference in the number of valleys between mid and high NBN stimuli with the low-frequency NBN stimulus as a reference variable according to the z-value statistics for each row in Table 3.7. Therefore, type of NBN carrier does not have a significant effect on the number of valleys in the EFR modulation transfer function.

Table 3.7: Output from the Poisson regression model determining the non-significant difference between different types of carriers in terms of the number of valleys in the EFR modulation transfer function

| Coefficient | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------|----------|------------|---------|---------|
| Intercept   | .000     | .213       | 0       | 1       |
| Mid-NBN     | -.201    | .318       | -.631   | .528    |
| High-NBN    | -.245    | .318       | -.771   | .441    |
3.8.2 Comparison to BBN studies

In the second part of this analysis, the number of valleys in the EFR modulation transfer functions evoked by NBN stimuli in the current study was compared with Purcell et al. (2004) where EFR were evoked by a 25% AM white noise carrier and the modulation frequency was varied with a sweep of 80-600 Hz in young normal-hearing individuals (n = 15, Average valley count = 2.73) and older individuals (n = 13, Average valley count = 2.92). Also, data from the present multifrequency NBN study were compared with an unpublished study where EFR were evoked by a 25% AM white noise carrier at modulation frequencies of 71-600 Hz in older individuals (n = 43, Average valley counts = 1.63). The rate of change of modulation frequency for the rising or falling modulation sweep was 33.73 Hz per second over 15.36 seconds for this current study, 32.42 Hz per second over 15.36 seconds for Purcell et al. (2004), and 32.3 Hz per second over 16.38 seconds for the unpublished study.

Figure 3.11 shows the histograms of valley counts for EFR modulation transfer functions across different studies.

Figure 3.11: Histograms of valleys in EFRs modulation transfer function for different stimuli across different studies in different group of individuals
The Wilcoxon Rank-Sum test was conducted to assess the difference among studies in terms of valleys in the EFR amplitude modulation spectrum. A Wilcoxon Rank-Sum test indicated that the number of valleys was significantly lower for NBN stimuli than BBN carriers in Purcell et al. (2004) in young normal hearing and older adults, and the unpublished study in older subjects. Table 3.10 shows significant results. The boxplots in Figure 3.12 summarizes this significant outcome.

Table 3.8: Results from the Wilcoxon Rank-Sum test indicated whether the NBN groups were different than the BBN groups in terms of the number of valleys in the EFR modulation transfer function. The low p-value indicates a high probability that the test results are correct.

<table>
<thead>
<tr>
<th>Subgroups Being Compared</th>
<th>Parameter Being Used</th>
<th>N</th>
<th>Wilcoxon Test Statistics (W)</th>
<th>Significance</th>
<th>Effect Size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multifrequency-NBN &amp; BBN - Older Adults (Unpublished study)</td>
<td>Number of Valleys</td>
<td>110</td>
<td>883.5</td>
<td>0.000338***</td>
<td>-.34</td>
</tr>
<tr>
<td>Multifrequency-NBN &amp; BBN - Younger Adults (Purcell et al. (2004))</td>
<td>Number of Valleys</td>
<td>82</td>
<td>103</td>
<td>0.0000051 ***</td>
<td>-.51</td>
</tr>
<tr>
<td>Multifrequency-NBN &amp; BBN - Older Adults (Purcell et al. (2004))</td>
<td>Number of Valleys</td>
<td>80</td>
<td>82.5</td>
<td>0.0000001***</td>
<td>-.6</td>
</tr>
</tbody>
</table>

*p< .05.  
**p< .01.  
***p< .001.
Figure 3.12: Boxplots of valley counts across different types of carriers. The NBN carriers had a lower number of valleys or fine structure in comparison with studies that used BBN carriers.

The boxplots in Figure 3.12 show the distribution of the number of valleys across different studies. The boxes range from the 25th percentile in the bottom, and the 75th percentile at the top. The area inside the boxes represents the interquartile range (IQR) (50% of the behavioural thresholds are located in this range). The medians are shown as the line that divides the boxes into two parts. The black dots indicate the means. The bottom and top whiskers are extended from the 25th percentile and 75th percentile respectively to 1.5 times the IQR or the height of the boxes. The endpoints of whiskers indicated with a longitudinal line represent the maximum and a minimum number of valleys in the data sets. The points above the endpoints are outliers which are defined as values that do not lie inside the whiskers.
Chapter 4

4 Discussion

4.1 Summary of results
Recent studies have suggested that broadband carriers can elicit small EFRs at the measurement electrodes due to the out-of-phase superposition of multiple parallel independent EFRs initiated from different regions of the tonotopic cochlea (Easwar, Banyard, Aiken, & Purcell, 2018a). These out-of-phase interactions could manifest in the EFR modulation transfer function as frequency regions with low amplitudes that can be described as valleys, dips, or fine structure. This project was motivated with aim of improving a previously proposed electrophysiological measures of AM detection (Purcell et al., 2004). We investigated the possibility of confining such out-of-phase interactions in EFRs by using NBN carriers.

To examine this hypothesis, EFRs were evoked by different AM NBN of low (centred at 750 Hz), mid (centred at 2200 Hz), or high frequency (centred at 3700 Hz) at a modulation depth of 50 percent where modulation rates \( f_m \) were swept linearly between 80-600 Hz over 15.36 s. In this study, a sweeping technique similar to Purcell et al. (2004) was used for changing the modulation rate. In the sweeping technique, neighbouring response estimates are not independent, but it does provide a smoothed-across-modulation-rates estimate of the electrophysiological modulation transfer function. Purcell et al. (2004) had also shown that the sweep EFR amplitudes were very similar to fixed modulation rate methods. In one-hour of EFR recording time, a small set of fixed modulation frequencies could be evaluated with acceptable SNR given typical EEG background noise as well as myogenic noise. In addition, recording times longer than 1 hour could become uncomfortable for participants and lead to higher artifact caused by myogenic activity.

In this thesis, we found that there were no significant differences between low, mid, and high-frequency NBN carriers in terms of EFR maximum significant frequency. Behaviorally, AM detection threshold was significantly lower for low-frequency NBN carrier compared to mid-frequency NBN. The behavioural AM detection threshold did not differ significantly for low-frequency NBN compared to high-frequency NBN carrier, or for mid-frequency NBN compared to high-frequency NBN carrier.
Additionally, no correlation could be found between psychophysical and electrophysiological measures. Moreover, the results of this study showed that the number of valleys was significantly lower for NBN stimuli than studies in which BBN was used as a carrier frequency.

4.2 EFR outcomes

4.2.1 Effects of carrier frequency on EFR amplitudes

The result of the present study showed that swept-frequency amplitude-modulated narrow band noise carriers could produce robust EFRs. The shape of EFR amplitude as a function of modulation frequency followed a general pattern of amplitude reduction as the AM rate increased from 80-600 Hz, which was similar to the low-pass filter characteristics of EFRs evoked by the AM broadband noise carrier in the literature (Purcell et al., 2004). We evaluated the effects of carrier frequency on EFRs. The amplitude of EFRs was significantly lower for low-frequency NBN carrier as compared to mid and high-frequency NBN carriers at modulation rates of 100-200 Hz. Such a carrier-frequency effect was not observed with modulation rates in the range of 80-100, 200-300, 300-400, 400-500 and 500-600 Hz. Similar outcomes have been previously reported in the literature (Lins & Picton, 1995; John & Picton, 2002; Zhu, Bharadwaj, Xia & Shinn-Cunningham, 2013; Sturzabecher et al., 2006). The 80 Hz ASSR is nearly identical in terms of amplitude for different carrier frequencies (Lins & Picton, 1995). In single stimulus recording, response amplitudes tend to be larger for 1 and 2 kHz than lower or higher carrier frequencies at low or mid intensities, but the amplitude differences were not statistically significant. John and Picton (2000) found a significant effect of modulation frequency and carrier frequency on the amplitude of ASSRs for modulation rates of 150-190 Hz. This significant effect was seen for tonal carriers where frequencies of 750 Hz and 6 kHz had lower amplitudes ($F = 8.04$, df $=3, 21$, $p < 0.05$) compared to 1.5 and 3 kHz (modulation rates were 158.7, 178.2, 168.5, and 178.2 Hz, respectively). In the same study, there was no significant effect on ASSR amplitude for the same carrier frequencies in the modulation range of 80-100 Hz (for ascending carrier frequency, modulation rates were 80.078, 80.322, 80.566, and 80.811 Hz, respectively). The carriers of 750 Hz and 6 kHz elicited lower amplitudes, but this difference was not statistically significant.
The observed influence of carrier frequency on the amplitude of EFRs has been explained by different phenomena in the auditory system in different publications. First, interactions between responses evoked from different neural generators in the ascending auditory nervous system at the same modulation rate have been suggested as a contributing factor (Picton, 2011). As the scalp-recorded EFR is the sum of multiple overlapping responses from different neural generators with different onset latencies, the interaction of these responses might be different depending upon the type of carrier frequency. The low amplitude of EFRs for the low-frequency NBN centred at 750 Hz in the modulation range of 100-200 Hz could have been caused by the out-of-phase interactions of neural generators in the upper and lower brainstem to the same modulation rates. Furthermore, the lower velocity of the travelling wave in the low-frequency region of the cochlea in the apex as compared to high-frequency regions in the base leads to longer delays and consequently poor synchronization of auditory nerve fibres. The poor synchrony has been indicated as a factor contributing to the lower amplitude of EFRs at low carrier frequencies (Zhu et al., 2013; Sturzabecher et al., 2006). Zhu et al. (2013) showed that the strength of EFRs (specifically, the phase-locking value) evoked at the fundamental frequency of 100 Hz was significantly weaker/smaller for a low-frequency complex tone. This low-frequency complex tone was comprised of the 1st to 5th harmonics of a 100 Hz tone compared to their high-frequency stimulus comprised of the 12th to 16th harmonics as well as a broadband stimulus containing the first 20 harmonics of the 100 Hz tone.

On the other hand, another possibility might be related to the characteristics of the auditory filters. The low-frequency harmonics located at the apex are relatively well resolved on the basilar membrane as individual auditory filters process these harmonics individually. Instead, high-frequency harmonics are not resolved, and multiple neighbouring harmonics are processed within a single auditory filter. These differences in response initiation may manifest as differences at the scalp for some modulation rates. Both mechanisms discussed above, resolvability and poor synchrony, could contribute to lower EFR amplitudes for responses initiated from the apical region by our low-frequency NBN carrier, however the reasons why the effect is limited to modulation rates between 100 and 200 Hz would require further study.
4.2.2 Effects of carrier frequency on EFR thresholds

Furthermore, we evaluated the effect of carrier frequency on the highest frequency at which the EFR could be detected (EFR threshold). Our finding showed that there were no significant differences between low, mid, and high-frequency NBN carriers in terms of EFR maximum significant frequency. As can be seen from boxplots in Figure 3.4, EFR maximum significant frequency tended to be lower for the low-frequency NBN carrier. The median EFR maximum significant frequency was numerically lower (297.5 Hz) for the low-frequency carrier than for the mid (474.5 Hz), and high frequency (403Hz) NBN carriers, though this difference failed to reach statistical significance.

4.2.3 Phase interactions

To investigate the possibility of out-of-phase interactions between independent EFRs initiated from different cochlear frequency regions, the number of valleys in the EFR amplitude modulation spectrum evoked by our NBN stimuli was determined and compared with previous studies using broadband noise. The histograms showing the number of valleys across different stimuli depicted in Figure 3.11 support the statistics indicating that the number of valleys was significantly lower for NBN stimuli than studies in which BBN was used as a carrier frequency. This present finding supports the hypothesis that phase interactions exist at the level of the scalp electrodes across independent responses initiated by different frequency components of broadband stimuli. This is consistent with the findings of Easwar et al. (2018b) who found the amplitude of the scalp-recorded EFR changes by systematically delaying the envelope phase of the second formant relative to the first format of two naturally spoken vowels. Easwar et al. (2018b) indicated that the amplitude of the EFR is maximum when the applied envelope phase delay compensates the cochlear delay across lower and higher frequency formant bands of the vowel stimuli, which leads to the constructive interference of the evoked responses at the level of the scalp. Previous studies reported similar results where the amplitude of EFRs change with the spectral characteristics of the vowel, or vowel identity (Aiken and Picton, 2006; Choi, Purcell, Coyne, and Aiken, 2013).

Considering that in the present study, the number of valleys in the EFR modulation transfer function has been compared between AM narrow-band noise stimuli with 50%
depth of modulation and previous studies using AM broadband noise stimuli with 25% depth of modulation, a more direct comparison should be made using the same modulation depth in both conditions. Therefore, future studies needed to determine the prevalence of valleys in the EFR modulation transfer function using AM BBN stimulus with a 50% depth of modulation.

4.3 Behavioral outcomes

4.3.1 Effects of carrier frequency on behavioural AM detection thresholds

Using psychophysical measures, the influence of carrier type on the behavioural AM detection threshold was investigated to determine whether the perceptual ability of listeners to track time-varying changes in the amplitude envelope varies depending on the frequency region of the carrier signal. Results of our study showed that the type of NBN carrier significantly affected the highest detectable modulation frequency. The behavioural threshold was significantly lower for low-frequency NBN carrier as compared to mid-frequency NBN carrier (p=.002). However, the behavioural AM detection threshold did not differ significantly for low-frequency NBN compared to high-frequency NBN carrier, or for mid-frequency NBN compared to high-frequency NBN carrier.

The effect of carrier cochlear frequency region on temporal envelope processing has been studied previously. In most psychophysical studies, the effect of carrier frequency has been studied in terms of TMTF cut-off frequency and sensitivity. TMTF is measured by determining the thresholds (minimum modulation depth) for detecting AM as a function of modulation frequency. For a broadband noise carrier, the AM detection threshold at low modulation frequencies is low (i.e., shallow modulation depth) and remains nearly constant until it begins to roll off at the cut-off frequency, similar to a low-pass filter function (Viemeister, 1979). TMTF sensitivity corresponds to the absolute values of the AM thresholds in the plateau of the low-pass filter function (Strickland, 2000). The TMTF cut-off frequency is the rate of modulation at which the AM threshold begins to increase by 3-4 dB per octave [converted from modulation depth (m) by 20*log (m)] relative to the low-modulation rate part of the function (Viemeister, 1979).
Findings from several psychophysical studies have shown that as the upper spectral edge of the modulated carrier increases in frequency, the temporal resolution improves (Viemeister, 1979; Bacon & Viemeister, 1985; Formby & Muir, 1988). Viemeister (1979) showed that as the center frequency of band-limited noise was reduced from 10 to 0.2 kHz, the TMTF cut-off frequency decreased correspondingly, which suggests reduced temporal acuity at lower center frequencies. Formby and Muir (1988) compared TMTFs obtained using broadband noise, high-pass filtered-noise with cut-off 4 kHz, and low-pass filtered-noises with cut-off frequencies of 1, 2 and 4 kHz. Their result showed that modulation detection thresholds were similar for both broadband noise and high-pass filtered-noise, which suggests that stimuli with wider spectra and higher frequency content improve AM detection thresholds. Accordingly, the low-pass filtered-noise with the lowest cut-off frequency had the most elevated thresholds. However, further studies pointed out that in both Viemeister (1979) and Formby and Muir (1988), that the bandwidth of the carrier signal increased as higher frequencies were included. Eddins (1993) investigated the influence of both frequency region and bandwidth on amplitude modulation detection. In this study stimulus bandwidths of 200, 400, 800 and 1600 Hz, and high frequency cut offs of 600, 2200 and 4400 Hz were used. Eddins (1993) showed that AM detection threshold increases by increasing the bandwidth of the NBN stimuli. Eddins (1993) and Due et al. (1997) found no effect of frequency region when the bandwidth was kept constant, but in both these studies, modulation was applied before filtering. Eddins (1999) showed that the technique of modulating after filtering reduces the effective modulation depth substantially, and has the largest effect on high-modulation frequencies and narrow-noise carrier bandwidths. The results of previous studies by Eddins (1993; 1999) showed that temporal acuity, as assessed by detection of amplitude modulation, does not depend on the frequency region of the carrier frequency over the range of 600 to 12800 Hz. Also, Strickland and Viemeister (1997) found no effect of frequency region using AM and QFM noises filtered and then modulated with the upper spectral edge of 1.2 kHz and above.

According to Strickland (2000), at a low stimulus spectrum level of 10 dB SL, the TMTF cutoff frequency is lower for bandlimited noise of 400-Hz bandwidth with an upper carrier spectral edge of 600 Hz compared to bandlimited noise of the same bandwidth with an upper spectral edge of 1200 Hz. With a 1600-Hz bandwidth, the cutoff frequency changed less between upper spectral edges of 2400 and 4800 Hz. The
same study showed at a stimulus spectrum level of 40 dB SL, that the frequency region did not show any effect on TMTF cut-off frequency, and at frequency regions above 2.4 kHz, TMTF cut-off frequency remained unchanged with changing the level. This is consistent with results of other studies that found no effect of frequency region on the TMTF cutoff frequency at a stimulus spectrum level of 40 dB SL or higher (Eddins, 1993, 1999; Due et al., 1997; Strickland and Viemeister, 1997).

On the other hand, Strickland (2000) showed that by keeping the frequency region constant, sensitivity (but not cut-off frequency) increased as the bandwidth of the stimulus increased, and this effect is independent of stimulus spectrum level. Similar findings have been suggested previously with bandlimited stimuli at high stimulus spectrum levels (Eddins, 1999; Strickland & Viemeister, 1997). Therefore, findings of previous studies suggest that increasing the bandwidth of bandlimited stimuli improves the AM threshold. Also, the TMTF cut-off frequency is influenced by the frequency region and the level of the signal.

Determining the highest modulation frequency at which AM was detectable (our behavioural AM threshold), and the effect of carrier frequency region was not the focus of the above prior studies. This study differed from prior studies in terms of the adaptive procedure for measuring the AM detection threshold, total number of trials and reversals, number of reversals that used to calculate the threshold, number of behavioral runs taken as the threshold estimate, as well as number of training trials provided to the participants, and the sample size. Therefore results of this study are not directly comparable with the results of previous studies in the literature.

Overall, the results of our study were consistent with some previous studies which suggested that a stimulus with wider spectral bandwidth, containing higher frequency regions, is a better signal for detecting amplitude modulation. However, other factors including the behavioural performance of the listeners would have been a contributing factor for our lower behavioural threshold for low-frequency NBN carrier as compared to mid-frequency NBN carrier, which will be explained later in the discussion.

4.4 Correlation between EFR and behavioural thresholds

One central aim of this study was to investigate whether a relationship exists between the objective measures of auditory temporal acuity as assessed by envelope following
responses and behavioural measures of amplitude modulation detection in normal hearing individuals. Our study showed no significant correlation between behavioural and objective measures within each NBN group separately when the critical p level for significance was corrected for familywise error rate (low-NBN, rho=.46, p=.029; mid-NBN, rho=.1, p=.63; High-NBN, rho=.46, p=.025). Grouping individuals across all three NBN noise carriers, there was a significant correlation between behavioural and objective measures of modulation sensitivity (rho=.39, p=.001).

The observed weak correlation for normal hearing adults can be influenced by the lack of between-subject variability of scores across participants. The EFR recorded in the present study provides an objective measure of auditory temporal acuity at the level of the brainstem. However, further data is required to assess the clinical utility of the measure. All of the tested subjects had normal hearing with no concern or difficulties regarding speech comprehension. It is well accepted that the value of a correlation coefficient is greater if the variability between observations is higher. With other factors equal across observations, the greater the range of observations, the greater the correlation between two variables (Goodwin and Leech, 2006).

In support of this hypothesis, similar weak correlation results were reported in Purcell et al. (2004) when correlation analysis was carried out within each group of normal hearing or older adults individually. The lack of variation across participants within each group and small sample size were deemed to contribute to the absence of strong significant correlation within each group in Purcell et al. (2004).

Moreover, the results of an unpublished dissertation by Alsamri (2017, p.40) showed that in different groups of individuals including young normal hearing adults, older adults with different types of hearing loss, and individuals with ANSD, the EFR evoked by an AM broadband noise carrier with depth of modulation swept from 2 to 100% at a fixed modulation frequency of 41 Hz was strongly correlated (r=.89) with behavioral AM detection thresholds. In the same study, EFRs evoked by the AM white noise carrier at a fixed modulation depth of 100% with the modulation frequency swept from 62 to 300 Hz were lower in amplitude for individuals diagnosed with ANSD. The highest modulation frequency at which the EFR was detectable was also lower (120 Hz) than would be expected from normal hearing individuals (300 Hz), indicating relatively poor temporal envelope processing in ANSD group. Therefore, including
populations with known temporal processing disorders, including older adults with hearing impairment, children diagnosed with auditory processing disorder (APD), individuals diagnosed with auditory neuropathy spectrum disorder (ANSD), and dyslexic children and adults, would increase the range of observed AM thresholds and consequently might improve the correlation between the objective and behavioral methods.

Furthermore, the lack of significant relationship between behavioural and EFR measures can be explained by the observed discrepancy between the EFR and behavioural thresholds. As can be seen from scatterplots (Figures 3.6 to 3.8), for some participants EFRs were detectable at higher AM rates, but participants were only able to detect AM at lower rates. Figure 4.1 shows an example of this disagreement for a sample subject.

### 4.4.1 Reliability of behavioural outcomes

The utility of the correlation analysis between behavioural and objective measures, as well as the effect of carrier frequency on the highest perceptually detectable modulation rate, might be influenced by the reliability of the behavioural measures. This dataset has been collected from student and non-student communities in order to ensure our data is as representative as possible of the normal hearing population. Our participants had no self-reported concerns or risk factors related to auditory processing difficulties. Even though careful and proper data collection was employed in all stages of data collection, and all participants were instructed and trained carefully, the behavioural test was challenging for some participants that had never performed the test before.

If we classify the participants of this study based on their behavioural performance, we observed three different groups of individuals. For some participants, the psychoacoustic adaptive 2AFC amplitude modulation detection task was easy to perform, and they achieved reliable thresholds with only one behavioural attempt.

On the other hand, for some individuals, one run of the behavioural test was not enough, and they required further runs to achieve a more reliable pattern from which we could confidently obtain a threshold. Those who needed more than one run of the behavioural test might achieve reliable thresholds with two or more runs; some individuals might need many runs to achieve reliable performance, or may never be able to. For some
participants, there was a learning element as their behavioural performance improve significantly after multiple runs. Figure 4.2 shows the learning curve for a sample subject where after eight runs they could achieve ceiling performance.

![Behavioral TMTF for a sample subject](image)

Figure 4.1: Discrepancy between behavioural and EFR maximum frequency (EFR threshold). A) Behavioural TMTF for a sample subject. Behavioural AM threshold was 176 Hz for this listener. B) EFR modulation transfer function and corresponding EEG noise-estimate for the same listener in Figure (A). EFR maximum frequency for this subject was 513 Hz.

Furthermore, a high degree of variability was observed in recorded behavioural thresholds. For some participants, a reliable response pattern gave a behavioural threshold that was not repeatable; another repetition of the behavioural test could yield an unreliable response pattern indicating an elevated threshold. Lack of attention and fatigue could contribute to increased variability and poor performance after multiple
runs of the behavioural AM detection task. According to Moore (2006), cognitive processes including memory, attention, and motivation play an essential role (to some extent) in any sensory or motor function. A significant number of participants in the present study demonstrated difficulty with performing the behavioural task, while they were able to perform well on the easy audiometry task. In our statistical analysis, the first attempt was used for the behavioural AM detection threshold for all participants.

In behavioural tests, some degree of variability is expected, but the least amount of variability is desirable for considering a test as a reliable clinical measure. Modifying the behavioural method might improve correlation between behavioural and EFR thresholds. In this study, a three-interval two-alternative forced-choice AXBD procedure was used, and behavioural thresholds were estimated using Virulent PEST (Findlay, 1978). The test stopped when the maximum number of 80 trials or the eighth reversal occurred, whichever happened first. However, for all participants in this study, the test stopped on the eighth reversal. The limit of 80 trials was never reached. On average, the behavioural thresholds of all participants in the low, mid and high-frequency NBN groups were estimated from 34.5 trials (SD = 8.2).

In adaptive psychophysical studies, a greater number of trials result in more accurate threshold estimation (Levitt, 1971). Amitay, Irwin, Hawkey, Cowan, & Moore, (2006) determined that an optimal adaptive procedure for efficient and precise threshold estimation in frequency discrimination and backward-masking tasks requires 30 trials for obtaining a reliable threshold estimate when inexperienced listeners with interpersonal variability are being tested. Also, as suggested, between-track variability might be controlled by combining the thresholds from multiple tracks with fewer trials rather than using long tracks for estimating thresholds. Therefore, having fewer trials (but not less than 30 trials) in each track might allow for the collection of more than one behavioural threshold estimate with less probability that attention or fatigue influence the listener's performance.
Figure 4.2: Multiple attempts of the behavioral AM detection task for a sample subject. The first to third attempts show the result of behavioural AM detection threshold obtained in the two-hour TMTF/EFR study session. This participant was asked to repeat the behavioural test in a second session. The plots of the fifth to eighth attempts were obtained in the second session of the behavioural test.
An alternative method could be a staircase procedure with a three-interval three-alternative forced-choice paradigm (Levitt, 1971). In this procedure, three intervals are presented to the listener, and the target stimulus can be in any of the three intervals. Each run of the adaptive staircase procedure must have a minimum of two blocks of trials, and the average of thresholds from two blocks is used as a threshold. If the first two thresholds differ by a predetermined value, a third estimate should be obtained. Using the latter procedure might help to yield more consistent results in a naive listener.

In psychophysical studies of AM detection, the modulation frequency is normally held constant while the depth of modulation is varied adaptively to determine the minimum depth of modulation needed to discriminate the AM stimulus. In contrast, in this study, the depth of modulation was held constant at 50% while the modulation frequency was varied adaptively. There is a possibility that a 50% depth of modulation was not deep enough for some listeners to perceptually contrast with unmodulated stimuli. In general, for a suprathreshold condition, the target signal is highly detectable and noticeable by the listener. In an AM detection task, as the modulation depth increases, the detection of AM gets easier. Therefore, listeners with a behavioural AM detection threshold around 50% depth may need larger AM depths as a reference to form their perceptual strategies and cues for the detection task.

An alternative approach that could yield reliable behavioural thresholds and consequently better insight about the correlation between objective and behavioural measures is to obtain standard temporal modulation transfer functions (i.e. for fixed rates, determine modulation depth at threshold similar to Viemeister (1979) at specific modulation rates of interest.

A more recent study has shown that EFRs can be recorded using a swept modulation depth technique (Dimitrijevic et al., 2016). According to the result of Dimitrijevic et al. (2016), the minimum depth of modulation evoking a significant EFR using an AM white noise carrier modulated at 41 Hz is correlated (r=.48) with behavioural AM detection thresholds in their sample of young normal-hearing and hearing-impaired older adults. A future study could use a similar swept AM depth approach at higher modulation frequencies (e.g. > 90 Hz) to give further insight about brainstem AM processing.
4.4.2 Reliability of EFR outcomes

EFR measures should also demonstrate adequate reliability to be established as an accurate clinical tool. For a test to be reliable, it should yield the same outcome at different times for a given set of measurement parameters. Minimal test-retest variability is desirable across repeated measurements. We determined that reliability of our EFR TMTF measure would need to be evaluated in future studies due to masters degree timing constraints. However, there is some literature that can inform us about EFR reliability. The test-retest reliability of 80-Hz ASSR (as a type of envelope following response) has been studied in previous literature in terms of the measured ASSR thresholds (D’haenens et al., 2008; Kaf et al., 2006; Luts and Wouters, 2005), and response amplitude (D’haenens et al., 2008; Wilding, McKay, Baker, & Kluk, 2011).

Luts and Wouters, (2005) reported a high test-retest reliability of 80-Hz ASSR for both MASTER and AUDERA systems for normal-hearing and hearing-impaired participants as assessed by determining the change in the mean difference between behavioural and ASSR thresholds between two sessions. Kaf et al. (2006) found a moderate to strong correlation (r= .74 to .93) between sessions for 80-Hz ASSR thresholds for different degrees of simulated sensory neural hearing loss for carrier frequencies of 500 to 4000 Hz. The 500-Hz carrier frequency had the weakest correlation coefficient compared to other carrier frequencies. Using a larger sample size than previous studies, as well as employing a more comprehensive statistical analysis, D’haenens et al. (2008) investigated the test-retest reliability of 80-Hz multiple-frequency ASSRs in normal-hearing adults. Absolute test-retest variability was evaluated by calculating the two standard error of measurement (±2 SEM; the 95% confidence interval of SEM was calculated as 1.96*SEM) in normal-hearing adults. Repeated 80-Hz ASSR thresholds were within ± 17, ± 12.3, ± 10.6, and ± 11.3 dB for carrier frequencies of 500, 1000, 2000, and 4000 Hz, respectively. This test-retest variability was considered possible due to their long test duration of 1 hour and 20 minutes. In the same study, the degree of occurrence of significant ASSRs at different intensities and different carrier frequencies were determined, in which at suprathreshold levels (50, 40, and 30 dB HL), the percentage of significant ASSRs were relatively similar between sessions for all carrier frequencies between 500 to 4000 Hz. Also, the test-retest reliability for ASSR thresholds was assessed with correlation analysis
between two sessions, and it was reported to be poor \((r = .34)\) for 500-Hz carrier frequency and moderate \((r = .55)\) for 1000, 2000, and 4000 Hz carrier frequencies respectively. D’haenens et al. (2008) reported the lack of between-subject variation in normal-hearing individuals as a contributing factor for the poor to moderate correlation, which may not be true for groups of individuals with hearing impairment.

The between-session correlation of response amplitude across different intensity levels for AM carrier frequencies of 500 to 4000 Hz were 0.91 for 50 dB HL, 0.82 for 30 dB HL, 0.69 for 20 dB HL, and 0.57 for 10 dB HL (D’haenens et al., 2008). Furthermore, the variability in response amplitude as measured by ±2 SEM was ± 6 to ±10 nV for 0 to 20 dB HL, and ±11 to ±15 nV for 30 to 50 dB HL in response to all AM carrier frequencies (pooled frequencies). In the same study, no significant difference was found between two sessions for ASSRs response amplitude to AM tones (ANOVA analysis). Wilding et al. (2011) investigated the repeatability of 80-Hz ASSR amplitudes at suprathreshold levels for an AM 2000-Hz carrier frequency in hearing-impaired and normal-hearing subjects by estimating the repeatability coefficient (which represents the 95% confidence interval of the test-retest mean difference). The repeatability coefficient for ASSR response amplitude for an AM 2000-Hz tone at 50 dB HL was estimated to be 29 nV for the normal hearing, and 57 nV for hearing-impaired individuals. However, test-retest variability in Wilding et al. (2011) was higher than test-retest variability in D’haenens et al. (2008) for the same stimulus at the same intensity for normal-hearing individuals. In Wilding et al. (2011), the mean artifact rejection levels were 31 μV (SD =11), and 23 μV (SD =4) for normal hearing and hearing impaired groups, respectively. As mentioned in Wilding et al. (2011), their higher artifact rejection threshold as compared to D’haenens et al. (2008) could be the reason for their higher test-retest variability.

ASSR (EFR) recordings can be influenced by within-subject variability in the level of background EEG-noise in the recording. This variability in noise can be caused by variability in myogenic activity or the ability of the listeners to sleep across ASSR testing sessions. The between-session correlation coefficient between noise levels was reported to be moderate for AM tones presented at 0 to 20 dB HL, and strong for 30 to 50 dB HL in D’haenens et al. (2008) study. In the same study, the means and standard deviations of the noise and ±2 SEM were quite small and diminished as intensity
reduced to 30 dB HL. They increased as intensity was reduced further from 20 to 0 dB HL. The within-subject variability in noise levels was low ranging from ±3.39 nV at 50 dB HL, ±1.39 nV at 30 dB HL, and 3.42 nV at 10 dB HL. Considering this low test-retest variability in noise estimates compared to ASSR amplitude variability, the test-retest variability of ASSR thresholds could not be explained by noise variability. Other factors might have contributed to variability in the response amplitude. Wilding et al. (2011) showed that the repeatability coefficient of EEG noise amplitude between two sessions was 10 and 22 nV for normal hearing and hearing impaired individuals, respectively. Amplitude variability could result from small changes in electrode placement, the electrode montage, and the change in impedance between the scalp-recording electrodes and volume conductors like skin and bone.

Further studies are needed to evaluate the test-retest reliability of the EFR in measures of auditory temporal acuity to determine whether the objective AM threshold is consistently repeatable across multiple recordings.
Chapter 5

5 Conclusion

This thesis study provides evidence regarding the presence of out-of-phase interactions at the level of the scalp among multiple, parallel EFRs initiated from different frequency regions of the cochlea by different spectral components of BBN stimuli. Using NBN carriers, the number of valleys in the EFR modulation transfer function was smaller compared to previous studies using BBN stimuli. This finding suggests that NBN carriers limit the initiation of out-of-phase responses and consequently reduces the degree of destructive interference of EFRs.

In the current study, EFRs were evoked by 50% AM of low, mid, and high-frequency NBN carriers, while the previous studies used BBN stimuli with 25% depth of modulation. Future studies are needed to replicate this work with similar stimulus parameter using a BBN carrier in normal hearing individuals to have a more equivalent comparison between NBN and BBN carriers in terms of the number valleys in the EFR modulation transfer function and have a better understanding of the nature of phase interactions across EFRs.

This thesis project aimed to improve a previously proposed electrophysiological measure of AM detection assessed by EFR. We demonstrated that the swept modulation rate technique can evoke robust EFRs in response to AM NBN carriers. The lack of correlation between psychophysical AM detection and EFR thresholds observed in this study could be caused by the lack of variability across thresholds in our sample of normal hearing individuals without known temporal processing disorders.

Including older adults with or without hearing impairment and clinical populations with a known temporal processing disorder such as individuals with ANSD and children with APD could improve the correlation of EFR thresholds with behavioural measures of AM detection, and thus improve the clinical applicability of the EFR as an objective clinical method. However, the complexity of these disorders can vary, which might lead to a different outcome of this objective measure in populations with the aforementioned disorders. With further research, objective measures could give insight into the integrity of the auditory nervous system, as well as the presence of abnormalities in neural processes underlying observed disordered behaviours.
References


Strickland, E. A. (2000). The effects of frequency region and level on the temporal


Appendices
Appendix A: Approval for Research Involving Human Participants
Date: 5 July 2018
To: Dr. David Perrelli
Project ID: 110308

Study Title: Maintenance Electrophysiological Measures of Temporal Acuity

Application Type: HSREB Initial Application
Review Type: Disagreed
Full Board Reporting Date: July 17, 2018
Date Approved Issued: 03/Jul/2018
REB Approval Expiry Date: 03/Jul/2019

Dear Dr. David Perrelli,

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above mentioned study as described in the HSREB application form as of the HSREB Initial Approval Date noted above. This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

<table>
<thead>
<tr>
<th>Document Name</th>
<th>Document Type</th>
<th>Document Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email Advertisement - 08 05 2018</td>
<td>Email Script</td>
<td>08 Jan 2018</td>
</tr>
<tr>
<td>Email Script_ImpContact - 08 05 2018</td>
<td>Email Script</td>
<td>08 Jan 2018</td>
</tr>
<tr>
<td>Consent and Consent - 11 05 2018</td>
<td>Written Consent Assent</td>
<td>13 Jan 2018</td>
</tr>
<tr>
<td>Questionnaire - 05 02 2018</td>
<td>Paper Survey</td>
<td>05 Feb 2018</td>
</tr>
<tr>
<td>Recruitment material - 04 06 2018</td>
<td>Recruitment Material</td>
<td>08 Jan 2018</td>
</tr>
<tr>
<td>SONA Advertisement_PsychologySONA system - 08 05 2018</td>
<td>Recruitment Materials</td>
<td>08 Jan 2018</td>
</tr>
<tr>
<td>Study_Protocol - 01 05 2018</td>
<td>Protocol</td>
<td>01 May 2018</td>
</tr>
<tr>
<td>SubjectCollectionForm</td>
<td>Other Data Collection Instruments</td>
<td>01 May 2018</td>
</tr>
</tbody>
</table>

No documents from or changes to the protocol or WEARE application should be initiated without prior written approval of an appropriate member from Western HSREB, except when necessary to obtain immediate written consent of study participants or when the change(s) involve only administrative or logistical aspects of the trial.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University HSREB operates in compliance with, and is committed in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2), the International Conference on Harmonization: Good Clinical Practice Consolidated guideline (ICH GCP), Part C, Division 1 of the Food and Drug Regulations, Part 5 of the General Health Products Regulations, Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number DR 00000940.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Karan Gopani, Ethics Officer on behalf of Dr. Joseph Gilbert, HSREB Chair

Note: This correspondence includes an electronic signature validation and approval via an online system that is compliant with all regulations.
Appendix B: Letter of Information and Consent
Letter of Information and Consent

Project Title:
Multifrequency Electrophysiological Estimate of Temporal Acuity

Principal Investigator:
David Purcell, Ph.D., Associate Professor, National Centre for Audiology, School of Communication Sciences and Disorders, Western University

LETTER OF INFORMATION

Introduction
You are being invited to participate in a research study investigating how to assess human hearing when listening to rapidly changing sounds. We will record electrical brain activity related to sound processing known as the envelope following response, and how this electrophysiological response is related to behavioral performance in listening tasks. You are being recruited for this study because you have normal hearing and our measurements will help us understand temporal processing in the normal hearing brain which is crucial for auditory and speech perception.

This letter is intended to provide you with the information you require to make an informed decision on participating in this research. Please take the time to read this information and feel free to ask questions if there is anything unclear to you. You will be offered a copy of this letter for your records.

Background and purpose of the research
Temporal auditory acuity refers to the sensitivity of the human auditory system to fluctuations in the loudness of a sound over time, which is essential for speech perception. The envelope following response (EFR) is brain activity that can be recorded while sounds that contain loudness fluctuations are presented to the listener. It is measured from surface electrodes placed on the human scalp. The maximum rate at which a sound can fluctuate and still cause a detectable EFR will be compared with people's ability to notice the fluctuations. The purpose of the study is to investigate how the EFR varies with the frequency content of sound, and whether some frequencies obtain a better correlation with behavioral measurement. This research will contribute to our understanding of how sound is processed in the brain over time.

Number of Participants
This study will include up to a total of 100 individuals.
Estimated Time
The total time required for the study is approximately 120 minutes.

Location of the study
The study will be completed at the National Centre for Audiology, Elborn College room 1207, Western University, London, ON, Canada.

Study procedures

- Questionnaire and Hearing Assessment
  If you agree to participate in the study, you will take part in a brief questionnaire and a brief assessment of your hearing. This will be followed by the main experiment, which will be conducted over one testing session. The questionnaire will ask you to report your age and handedness, and any known neurological, speech and language, vision and hearing problems. You may choose to omit a response to a specific question on the questionnaire without any penalty.

  The hearing assessment will be a visual examination of your ear canals and a measurement called a pure-tone audiogram which takes about 12 minutes to complete. You will hear tones one at a time through headphones, and you will signal when you detect each tone. The tones will progressively become quieter until you are no longer able to hear them. This procedure is repeated for several different pitches and for each ear.

- Behavioral measurement
  In this test your ability to discriminate sounds with different characteristics is tested. We will run a computer-based program for you, and your task is to use a computer mouse to choose the sound that is different from the others. The task becomes harder after correct answers and becomes simpler after incorrect answers. This test will take 10 to 20 minutes depending on your performance.

- Envelope Following Response (EFR)
  An electrical measurement of your brain's response to sound will be recorded. This requires the placement of an earphone insert into the ear canal of one ear, and the placement of three surface electrodes onto the skin of the head and collarbone. The sites for these three electrodes will be cleaned with an alcohol pad and a gentle scrub gel to improve electrical contact. A conductive gel and light adhesive will hold them in place. After the experiment, the electrodes will be gently removed and the gel cleaned away with a damp paper towel.

  During the measurement, you will lie comfortably in a reclined easy chair and are encouraged to sleep. Sounds will be presented at a comfortable loudness and measurement time will be approximately 60 to 90 minutes.
Potential Risk and Discomfort

These methods are widely used in laboratories studying hearing. The only known risk with the measurements is the possibility that some individuals may temporarily experience redness where the surface electrodes were placed due to the skin cleaning procedure. In the very rare case of significant redness, we would discontinue the use of electrodes and invite the participant to exit the study. If preferred, participants can attend behavioral and electrophysiological tests separately in two measurement sessions in order to reduce the duration of individual sessions.

Potential Benefits

There are no direct benefits of this basic research to the individuals. There may be the possibility that the brief hearing assessment could identify a previously unknown hearing impairment. If this were to occur, we would encourage the participant to seek professional assessment from their family practitioner or audiologist. We would provide information about obtaining an assessment at the UW0 audiology clinic in Elborn College. Given the importance of temporal acuity in speech comprehension, it may contribute to our understanding of temporal processing of speech, which is of benefit to society in the long term.

Withdrawal

If you decide to withdraw from the study, you have the right to request (e.g., written, calling, etc.) withdrawal of data collected from you. If you wish to have your data removed, please let the researcher know and it will be removed from our records.

NOTE: once the study has been published we will not be able to withdraw your data.

Confidentiality

All information obtained in this study will be held in strict confidence and participant anonymity will be maintained. Full names will be collected as part of the consent form and for scheduling needs. Personal identifiers, specifically name, email and possibly telephone number, will be linked to a participant ID code in a password protected electronic master list. These personal identifiers will NOT be retained after data publication is complete. All data is collected using an anonymous participant ID. Your name will not appear in any publications or presentations of the findings of this study. Representatives of Western University’s Health Sciences Research Ethics Board may require access to your study-related records to monitor the conduct of the research.
Compensation
The entire study will be conducted over a 2 hour session and you will be compensated for your time at the rate of $5 per half-hour or part thereof, regardless of whether you withdraw midway through a half-hour period. If you are a SONA participant, the compensation will not be money and you will be compensated with 1 research credit per hour toward PSYC1000 for participating in this study. If you are enrolled in a course other than Psych 1000, your compensation will be based on your course outline.
If you have any questions about the time or compensation, please ask the investigators before you consider signing the consent.

Legal Rights of Participants
Your participation in this study is voluntary. You may decide not to be in this study. Even if you consent to participate, you have the right to not answer individual questions or to withdraw from the study at any time. If you choose not to participate, or to leave the study at any time, it will have no effect on your care, employment status or academic standing. We will give you new information that is learned during the study that might affect your decision to stay in the study. You do not waive any legal right by signing this consent form.

Contact Persons
If you have any questions or would like additional information about this study, please contact Dr. David Purcell, National Centre for Audiology, School of Communication Sciences and Disorders, University of Western Ontario, London, Ontario, N6G 1L1 (telephone: [redacted]).

If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Human Research Ethics (519) 661-3036, 1-844-720-9816, email: ethics@uwo.ca. The Research Ethics Board is not part of the study team. Everything that you discuss will be kept confidential.

This letter is yours to keep for future reference.
Consent Form

Project Title: Multifrequency Electrophysiological Estimate of Temporal Acuity

Document Title: Consent Form

Principal Investigator: David Purcell, Ph.D., Associate Professor, National Centre for Audiology, School of Communication Sciences and Disorders, Western University.

I have read the Letter of Information, have had the nature of the study explained to me and I agree to participate. All questions have been answered to my satisfaction.

Print Name of Participant          Signature          Date (DD-MM-YYYY)

My signature means that I have explained the study to the participant named above. I have answered all questions.

Print Name of Person Obtaining Consent          Signature          Date (DD-MM-YYYY)

Page 5 of 5       Version Date: 13/06/2018
Curriculum Vitae

Name: Negar Ahzan

Post-secondary Education and Degrees:
Shahid Beheshti University of Medical Sciences, Tehran, Iran, 2004-2009 B.Sc.

Honours and Awards:
Western Graduate Research Scholarship (WGRS), 2017-2019

Related Work Experience:
Graduate Teaching Assistant, The University of Western Ontario, 2017-2019
Graduate Research Assistant, The University of Western Ontario, 2017-2019