ESTIMATION OF BEHAVIORAL HEARING THRESHOLDS IN NORMAL HEARING LISTENERS USING AUDITORY STEADY STATE RESPONSES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

J. Kip Kelly, M.S.

* * * * *

The Ohio State University

2009

Dissertation Committee:

Lawrence L. Feth, Ph.D., Advisor
Christina M. Roup, Ph.D.
Ashok Krishnamurthy, Ph.D.

Approved by

____________________________________
Advisor
Graduate Program in Speech and Hearing Science
ABSTRACT

The ability to obtain frequency specific information regarding a patient’s hearing sensitivity in an objective manner allows the evaluation of patient populations who cannot be tested through traditional behavioral methods. One method for obtaining this information is the auditory steady state response (ASSR). ASSR permits the testing of multiple carrier frequencies simultaneously and both ears simultaneously, unlike the auditory brainstem response (ABR). ASSR replaces subjective examiner interpretation of the response with statistical analyses not subject to the variability of human observers.

Unlike ABR which has been in use for decades and utilizes relatively consistent stimuli and test protocols, the ASSR has only been in widespread clinical use for the past 6-8 years and consequently does not have the same level of standardization as ABR. ASSR can be elicited by a variety of stimulus types including but not limited to: (1) Sinusoidal amplitude modulated (SAM) tones, (2) Frequency modulated (FM) tones, (3) Mixed modulation (MM) tones, and (4) Toneburst (TB) trains. The ASSR response is found in the frequency domain at the frequency of modulation and is frequently differentiated from unrelated neural activity using an F-statistic to determine if the amplitude of the line spectra at the modulation frequency is statistically different from the surround physiologic noise.
The current study sought to evaluate several common stimuli used in ASSR testing to determine if a more recently introduced stimulus (TB) emerges as a more appropriate stimulus for generating the response. Response detection and collection parameters were standardized so that any differences seen could be attributed to the stimulus. Both behavioral and ASSR thresholds were measured using SAM, MM, and TB stimuli in ten young adults with normal hearing (≤ 15 dB HL from 250-8000 Hz). Comparisons were then made between stimulus types to determine which stimuli could best predict a behavioral response for a pure tone matching the carrier frequency.

The results of the current study indicate that the MM and TB stimuli provide lower ASSR thresholds than do SAM stimuli and that a regression model provides the most accurate estimates of behavioral threshold. The thresholds for an individually presented TB were consistently lower than for a TB at the same frequency that was presented in the multiple simultaneous paradigm (four simultaneous carrier frequencies presented to the ear). However the threshold predictions based on the two measurements were similar so little accuracy in prediction is lost by using multiple simultaneously presented tonebursts. The current study shows that while ASSR can provide reasonable estimates of hearing sensitivity when the mean data are examined for any given individual the accuracy of prediction can vary greatly.
Dedicated to the memory of my grandparents: Woodrow W. Kelly, Thelma M. Kelly, Kenneth L. Morris, M.D., Susie Morris, and my surrogate grandparents Hazel (Auntie B) and John (Big Uncle John) Percy—for always making sure that their grandchildren had access to the educational opportunities that they did not.

Dedicated also to my parents, Mike and Candi Kelly for their support and encouragement in all my endeavors.
ACKNOWLEDGEMENTS

I wish to thank my dissertation committee and mentors through my time at The Ohio State University: Christina Roup, Lawrence Feth and Wayne King. Dr. Roup provided invaluable assistance in the writing and organizing of the document. Thanks are also due to Drs. Feth and King for their guidance throughout the project and the use of their laboratories and equipment to collect the data for this project. Thank you also to Dr. Ashok Krishnamurthy for stepping in to complete the committee following Dr. King’s departure from the University. Special thanks are due to my subjects who endured several hours of testing and who came in at all hours of the day and night for their electrophysiological testing. Thanks are also due to my graduate student colleagues and good friends: Dean Hudson, Clint Keifer, Harry Patra, and Scott Seeman - thank you for the support and keeping me from losing my mind. Finally thank you to my family for all the support and encouragement throughout this process.
VITA

February 1, 1975…Born—Lafayette, Indiana

1997…………………B.A. Communication Sciences and Disorders, College of Wooster
1999…………………M.S. Communicative Disorders, University of Wisconsin-Madison
1999-2001….......Research Audiologist/Clinical Fellow, VAMC Mountain Home, TN
2006-Present………Research Audiologist, VAMC Mountain Home, TN

PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Speech and Hearing Science

Emphasis: Auditory Electrophysiology
# TABLE OF CONTENTS

Abstract .................................................................................................................. ii
Dedication ............................................................................................................... iv
Acknowledgments ........................................................................................................ v
Vita ........................................................................................................................ vi
List of Tables .......................................................................................................... x
List of Figures ......................................................................................................... xi

Chapters:

1. Introduction ........................................................................................................... 1
   1.1 Problems of Interest ......................................................................................... 5
   1.2 Purpose of the Current Study ........................................................................... 6

2. Review of the Literature ....................................................................................... 7
   2.1 The Electroencephalogram (EEG) ................................................................. 7
   2.2 Auditory Components of the EEG .................................................................. 10
      2.2.1 Neurophysiologic Responses to Auditory Stimuli .............................. 11
      2.2.2 Transient Responses .............................................................................. 16
      2.2.3 Steady State Potentials of the Auditory System .................................. 18
   2.3 The Auditory Steady State Response (ASSR) .............................................. 21
      2.3.1 Response Generators ............................................................................. 22
      2.3.2 Stimuli Definitions ............................................................................... 25
      2.3.3 Recording Techniques ........................................................................ 28
      2.3.4 Response Analysis ............................................................................... 29
   2.4 Advantages/Disadvantages of ASSR ............................................................ 31
      2.4.1 Interaction of Multiple Simultaneous Stimuli ....................................... 34
      2.4.2 Profound Hearing Loss and ASSR ....................................................... 34
      2.4.3 Clinical Applications of ASSR ............................................................. 35

3. Methods .............................................................................................................. 37
   3.1 Subjects .......................................................................................................... 37
   3.2 Stimuli ........................................................................................................... 37
      3.2.1 Pure Tones ............................................................................................. 38
      3.2.2 Sinusoidal Amplitude Modulated (SAM) Tones ................................ 39
      3.2.3 Mixed Modulation (MM) Tones ............................................................. 39
3.2.4 Toneburst (TB) Stimuli .................................................. 42
3.3 Procedures ................................................................. 47
   3.3.1 Psychoacoustic Threshold Procedures ....................... 47
   3.3.2 ASSR Threshold Procedures .................................... 50
3.4 Computer Modeling of Cochlear Excitation ............................... 52

4. Results
   4.1 Behavioral Threshold Data ........................................... 54
   4.2 ASSR Threshold Data ................................................... 56
   4.3 Threshold Estimates ................................................... 59
   4.4 Results of Computer Models of Cochlear and Auditory Nervous
       System Excitation ..................................................... 72

5. Discussion and Conclusions ............................................. 75
   5.1 Rationale for the Current Study ................................. 75
   5.2 Behavioral Thresholds .............................................. 76
   5.3 ASSR Thresholds ..................................................... 78
   5.4 Behavioral Threshold Estimates .................................. 82
   5.4 Modeling of the Auditory System ................................. 88
   5.6 Conclusions ........................................................... 89
   5.7 Implications for Future Research ................................. 91

Bibliography ................................................................. 93

APPENDIX A: Preliminary Qualification Questionnaire/Screening Forms .......... 102
APPENDIX B: Cochlear Excitation Model Outputs ....................................... 105
APPENDIX C: Individual Subject Data ............................................. 114
LIST OF TABLES

Table

4.1 Linear Regression Formulae.................................................................61
4.2 Quadratic Regression Formulae............................................................61
4.3 Mean Differences between ASSR and Pure Tone Behavioral Thresholds........62
4.4 Mean Prediction Errors........................................................................69
4.5 Mean Jack-knife Prediction Differences..................................................70
5.1 Cone-Wesson and Rance Regression Predictions......................................87
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Time Waveforms of SAM Tones</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>Power Spectra of SAM Tones</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Time Waveforms of MM Tones</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>Power Spectra of MM Tones</td>
<td>44</td>
</tr>
<tr>
<td>3.5</td>
<td>Time Waveforms of Tonebursts</td>
<td>45</td>
</tr>
<tr>
<td>3.6</td>
<td>Power Spectra of Tonebursts</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Mean ASSR and Behavioral Thresholds</td>
<td>55</td>
</tr>
<tr>
<td>4.2</td>
<td>Mean Multiple and Individually Presented Toneburst Thresholds</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Predicted and Measured Thresholds 500 Hz</td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>Predicted and Measured Thresholds 1 kHz</td>
<td>64</td>
</tr>
<tr>
<td>4.5</td>
<td>Predicted and Measured Thresholds 2 kHz</td>
<td>65</td>
</tr>
<tr>
<td>4.6</td>
<td>Predicted and Measured Thresholds 4 kHz</td>
<td>66</td>
</tr>
<tr>
<td>4.7</td>
<td>Cross Validation Threshold Comparisons</td>
<td>71</td>
</tr>
<tr>
<td>5.1</td>
<td>Representative Regression Lines</td>
<td>84</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

The auditory steady state response (ASSR) refers to a number of auditory evoked potentials obtained with modulated stimuli. The ASSR reflects activity in the auditory nervous system generated in response to the presentation of a modulated signal to the ear. The primary clinical purpose of ASSR is the estimation of hearing sensitivity in populations that cannot provide reliable behavioral responses, such as infants or difficult to test adults. ASSR thresholds provide estimates of hearing sensitivity across a range of frequencies. These estimates are then utilized in the identification, diagnosis, and remediation of an identified hearing loss. ASSR thresholds have been reported to be within approximately 10-35 dB of a behavioral threshold for a pure tone corresponding to the carrier frequency of the ASSR stimulus (e.g., Rance, Rickard, Cohen, DeVidi & Clark, 1995; Rance & Rickards, 2002; Herdman & Stapells, 2003; Lins & Picton, 1995). One of the biggest weaknesses of ASSR, however, is the considerable variability between ASSR thresholds and behavioral thresholds (e.g., John, Dimitrijevic, Van Roon, & Picton, 2001; Picton, John, Dimitrijevic, & Purcell, 2003; Cone-Wesson, Parker, Swiderski & Rickards, 2002; John, Purcell, Dimitrijevic & Picton, 2002). The lack of a standardized way to evoke, detect, and use ASSR thresholds in predicting hearing sensitivity accounts for a substantial portion of the variability seen in the research literature. In practice, the
large variability of the ASSR makes it difficult to determine what constitutes a normal response and what constitutes an elevated response indicative of hearing loss.

A probable source for some of the variability that has been observed across studies is the use of different stimuli. The most widely utilized stimuli to elicit an ASSR are sinusoidally amplitude modulated (SAM) tones and mixed modulation (MM) tones (a combination of amplitude and frequency modulation). More recently, trains of brief tonebursts (TB) have been used, but to date little work has been published using toneburst stimuli (Burkard & Szalda, 2004, 2005; Han, Mo, Lio, Lian & Huang, 2006; Mo & Stapells, 2007). Further, the majority of this work has been performed in animals rather than humans.

The number of studies which directly compare ASSR in normal-hearing or hearing-impaired listeners evoked with multiple types stimuli is limited (Dolphin, Chertoff & Burkard, 1994; Herdman & Stapells, 2003; John, Dimitrijevic, Van Roon & Picton, 2001; Picton, Skinner, Champagne, Kellet & Maiste, 1987). Such studies are crucial for characterizing what proportion of the observed variance in the ASSR response is due to differences in the stimuli used to evoke or measure an ASSR. Changes applied to stimulus parameters (e.g., signal bandwidth/envelope and modulation type) may lead to variations in the way the stimuli are processed in the auditory system resulting in differences in response characteristics of the ASSR. A wider bandwidth stimulus, for example, has the potential to excite a larger area in the cochlea resulting in greater activation of the auditory nervous system leading to lower ASSR thresholds (John, et al., 2001; Picton, et al., 2003). The broader acoustic spectrum is just one possible reason to expect the toneburst ASSR to exhibit larger average response amplitudes and lower
thresholds relative to the SAM ASSR. In addition to the bandwidth and spectral make-up of the signal, we must also consider the temporal characteristics of the stimuli. TBs have different temporal characteristics than both MM and SAM stimuli used in ASSR. For instance, the sharper envelope of the toneburst may be expected to enhance neural synchrony (Hecox, Squires, & Galambos, 1976), resulting in larger response amplitudes and lower thresholds relative to the continuous modulated tones.

In spite of our current imperfect knowledge of the ASSR, the technique has a number of putative strengths as an estimator of hearing sensitivity: (1) the response is more frequency specific than some established evoked potentials, e.g., the click evoked auditory brainstem response (ABR), (2) it is resistant to stimulus artifact, (3) the modulation scheme allows thresholds to be obtained for multiple carrier frequencies simultaneously, and (4) the response is insensitive to the attentive state of the patient for a wide range of modulation frequencies (Stapells et al., 2004).

In any electrophysiological recording, contamination of the response with an artifact from the stimulus is always a concern. Modulated stimuli, such as those used for ASSR, contain little or no acoustic energy at the frequency of modulation. Having little or no acoustic energy at the response (modulation) frequency is advantageous as the possibility of misidentifying a stimulus artifact as a response is greatly reduced or eliminated. The neural activity we refer to as ASSR arises from the auditory nervous system’s ability to phase lock to the modulation frequency of the stimulus (Picton, Skinner, Champagne, Kellet & Maiste, 1987). The response being located at the modulation frequencies makes it possible to simultaneously test multiple carrier frequencies by modulating each carrier at a different frequency (John, Lins, Boucher, &
Picton, 1998). Several studies have shown that thresholds for up to four carrier frequencies can be recorded simultaneously in one or both ears without significant interaction between the stimuli (e.g., Lins & Picton 1995; John et al., 2002), provided the carrier frequencies were at least one octave apart and stimulus levels were at or below 60 dB SPL (John, Lins, Boucher & Picton, 1998; Herdman & Stapells, 2001; Picton, Dimitrijevic & John, 2002).

Using multiple simultaneous stimuli results in substantial time savings in the clinic where obtaining four thresholds simultaneously has been demonstrated to be approximately 2.4 times faster than if the same four thresholds were measured individually (John et al., 2002b). The reason that we do not see recording times that are 4 times as fast (assuming 4 simultaneous stimuli) is because responses for some carrier frequencies (especially 500 Hz) take longer to record than others due to poorer neural responses and higher noise levels. Because of these differences in response acquisition time the multiple simultaneous stimuli paradigm will only be as fast as the slowest individual carrier frequency to generate a response (John et al., 2002b).

An additional strength of ASSR is that thresholds have been shown to be relatively resistant to degradation due to sleep and/or sedation (Cohen, Rickards & Clark, 1991; Aoyagi et al., 1993; John et al., 1998). The primary use of this test is estimation of hearing thresholds in infants and young children. The ability to record reliable responses during sleep or sedation a considerable strength, as patients in these groups generally must be asleep or sedated to allow testing to be completed.
1.1 Problems of Interest

While a great deal of research has been conducted using ASSR, there are still some problems of interest which need to be investigated: (1) the potential for interaction between stimulus components influencing responses when using multiple simultaneous toneburst stimuli, and (2) the lack of a standardized test protocol which has resulted in considerable variability across studies. Given that the ASSR is currently employed in clinical practice, it is imperative that these remaining problems be addressed to ensure appropriate clinical decisions are being made based on ASSR test results.

The wider acoustic spectrum of the toneburst compared to SAM tones results in greater potential for stimulus overlap and interaction due to sidebands of the stimulus components overlapping one another adding to the noise in the overlapping region. An increase in noise will require a higher amplitude response to make up the difference in signal to noise ratio (SNR) and potentially raise threshold. The resistance of multiple simultaneously presented toneburst stimuli to interaction effects has not been adequately demonstrated in the research literature. All the studies reporting minimal stimulus interactions were conducted using SAM or MM stimuli. There is not a clear consensus as to which stimulus provides the most reliable estimator of behavioral threshold.

Another potential source of variability is the lack of a standard response detection criterion. In many studies an F-test has been used to detect the response, however, these studies use different response detection criteria with the F-test resulting in some studies having more conservative response detection criteria than others. This lack of standardization in ASSR testing makes it difficult to accurately compare results across studies or clinics.
1.2 Purpose of the Current Study

As mentioned in the previous section there are numerous questions yet to be addressed regarding ASSR and its use in hearing threshold estimation. The current study will address the following questions:

(1) Are behavioral thresholds for ASSR stimuli comparable to behavioral thresholds for pure tones and it is reasonable to utilize the stimuli to estimate a pure tone threshold?

(2) Does the sharper envelope of the TB stimuli results in lower ASSR thresholds relative to SAM & MM stimuli?

(3) Does the use of a regression model provide a more accurate threshold estimation of the pure tone behavioral threshold than subtracting a mean correction factor?

(4) Does the time savings of presenting multiple simultaneous TBs offset any stimulus interaction effects relative to individual stimulus presentation as has been demonstrated previously with other stimuli?
CHAPTER 2

REVIEW OF THE LITERATURE

2.1 The Electroencephalogram (EEG)

Neural activity in the brain can be non-invasively measured as an electrical signal using electrodes placed at various locations on the scalp. This measurement of neural activity is known as the electroencephalogram or EEG. EEG measurement is useful in the diagnosis of neurological disorders that cause disruptions in normal brain activity. The EEG is recorded in one of two manners: (1) bipolar recordings which utilize pairs of active electrodes, or (2) referential recordings which link the active or recording electrode(s) to common reference or passive electrode(s) located in a neutral area. Both types of recordings also require a ground electrode which provides the amplifier with a reference voltage of electrical activity. Differential amplifiers then compare the differences in voltage between electrodes.

The EEG is generally subdivided into 5 bands based on the frequency of the observed waveforms. The slowest division is the Delta band which includes frequencies of up to 3 Hz. These slowest waveforms are associated with deep sleep in an adult human with normal neurological function. The Theta range occurs between 3-7 Hz and is generally associated with internal focus (e.g., meditation). Alpha rhythms occur between 8-12 Hz and are found in the normal adult brain during relaxed but alert states.
Beta rhythms occur between 13-30 Hz and are generally associated with active thought processes. Finally, Gamma rhythms (above 30 Hz) reflect information processing and through integration. Nunez (1995) notes the vast majority of EEG data are thought to originate from generators in the neocortex rather than subcortical structures, however, Nunez (1995) also goes on to point out the pertinent exception of averaged evoked potentials (e.g., ABR, ASSR) which arise from subcortical structures.

Recording the ongoing activity of the EEG is of little value without a way to coherently interpret the time recordings. A number of statistical measures have been used to analyze the information in EEG recordings, one of which is the Fourier transform. Fourier analysis of EEG signals provides investigators with a means of examining numerous simultaneous oscillations by separating the different rhythms (Quiroga, 1998). Despite its strengths, Quiroga (1998) points out some weaknesses in using Fourier transform in EEG analysis: (1) it performs best when the signal is stationary, but the EEG is a highly non-stationary signal, and (2) the latency of the transformed activity is lost. In the case of auditory evoked potentials (AEPs), the first weakness is not a major hinderance as AEPs are relatively stationary responses. The relative stationarity of evoked responses allows for the averaging of multiple responses to consecutive signals which enables them to be detected from within the non-stationary EEG at the surface of the scalp. The loss of direct response latency measures is a downside of utilizing Fourier transform in AEP recording, as latency can provide some clues about the site of generation for an AEP.

The further an electrode is from the generator site the more extraneous activity it will record making it more difficult to detect the desired activity. Additionally it
becomes harder to localize the exact generator of the response as the spatial resolution at
the scalp is decreased due to the increased distance from the source and the relatively
poor conductivity of the skull (Nunez, 1995). In addition to the increased noise in the
recording the amplitude of the desired potential progressively decreases as the distance
between the site of generation and the electrode increases. In the case of a stationary
response, such as an AEP, the response can be greatly improved through the use of
averaging additional response waveforms (Nunez, 1995).

When using scalp electrodes in measuring evoked potentials a response generated
by a single cell is far too small to measure. What is being measured on the scalp is a
volume conducted representation of the activity in the nervous system. The electrical
potentials underlying the neural response are transmitted through the tissue of the head to
the scalp via volume conduction. Volume conduction is the conduction of electrical
potentials through the tissues (e.g., brain, cerebrospinal fluid, skull, muscle, skin) of the
head when considering EEG (van den Broek, Reinders, Donderwinkel & Peters, 1998;
Nunez, 1995). Electrical activity in the brain results in the creation of a current dipole
which flows through the surrounding tissue. EEG is a measurement of the return currents
created by the dipole at the scalp. It is important to make clear that electrical activities
recorded on the scalp are not action potentials, but rather post synaptic potentials (PSPs);
either excitatory or inhibitory (EPSP or IPSP). The location of the electrodes relative to
the generators of the response, the alignment of the response generators and the
conductivity of the surrounding tissue will influence the amplitude and latency of the
recorded response.
2.2 Auditory Components of the EEG

AEPs are a class of neurophysiological responses, which are recorded in response to the presentation of an auditory stimulus (i.e., clicks, tone bursts, modulated tones, speech). AEPs are further divided into sub-groups based on their latency (in milliseconds) with reference to the onset of the stimulus signal. In humans, these responses are almost always recorded through the use of electrodes placed at various locations on the scalp. Scalp recordings are referred to as “far field” recordings due to their relative distance from the response generating structures. Synchronous depolarization in a population of neurons will result in a current dipole which may be detectible at a remote location (e.g., the scalp).

AEPs are only one component of the overall EEG and therefore must be separated from uncorrelated electrical activity in the brain (referred to as noise) through averaging and the use of high and low pass filters. Averaging is accomplished by recording multiple waveforms (often hundreds or thousands) in response to repeated stimuli and adding the resulting activity for each stimulus presentation response together. The portion of the EEG that is evoked by the stimulus will be relatively consistent across recordings resulting in an additive effect when the waves are combined. Uncorrelated activity is assumed to be random and will thus tend to cancel itself out, albeit not completely, when the individual waveforms are averaged together. Prior to starting any recordings the amount of noise present in the recording can reduced by using filters to eliminate frequency regions that do not contain any portion of the AEP that is being recorded. AEPs each occur within specific frequency ranges so by ignoring any activity located outside the known range of frequencies for a given potential there will be less
noise to eliminate through the use of response averaging. Reducing the noise levels will reduce the number of averages necessary to detect a response.

2.2.1 Neurophysiologic responses to auditory stimuli

*Pure tone*

When a pure tone is presented to the ear, increases in firing rates above spontaneous levels can be observed in the neurons of the auditory system sensitive to a tone of that frequency. The higher the intensity of a tone the greater the increase in firing rates over the spontaneous rate and the greater the number of neurons firing. The firing rate will spike with the onset of the tone and then decrease but remain above the spontaneous firing rate. Cochlear afferents, like all neurons, exhibit an all or none firing behavior. Therefore, they must encode differences in intensity or frequency by varying the rate at which they fire. Any given fiber will be most sensitive to a specific frequency (its characteristic frequency) and will increase its firing rate above the spontaneous level at a lower intensity level for that frequency compared to other frequencies. Fibers will respond to frequencies other than their characteristic frequency, but those signals will need to be of greater intensity to elicit a response. An intense broadband signal, such as a click, will result in a robust response because it stimulates a large population of neurons. Additionally, as it repeatedly turns on and off it generates numerous onset responses. This is in contrast to a pure tone which will stimulate a much narrower region if presented with equal power. When hearing threshold levels are approached, the population of neurons responding decreases and synchrony is reduced. This makes neural activity difficult to separate from background activity at the scalp even though the stimulus is still detectible from within the system (perceptually detectible by the listener).
The anatomy of the cochlea also influences the characteristics of the response as the stiffness of the basilar membrane varies along the length of the membrane due to changes in its thickness and width. The stiffness of the basilar membrane will influence how the traveling wave moves through the cochlea. As a wavefront passes the region most sensitive to the frequency of the stimulus the traveling wavefront will slow down and be damped (de La Rochefoucauld & Olsen, 2007). High and low frequency tones do not travel down the basilar membrane at the same speed, and low frequency tones must travel a greater distance to reach their characteristic frequency region resulting in a phase delay for the responses from lower frequency regions of the cochlea. These delays in phase for a lower frequency response put it out of phase with the already evoked responses from more apical regions of the cochlea making recording an AEP with a low frequency tone more difficult.

_Sinusoidal Amplitude Modulation_

When a SAM stimulus is presented to the ear, such as in ASSR measures, the fibers of the auditory nerve which are sensitive to the carrier frequency will fire in a manner that reflects the phase locking of the cochlea to the modulation envelope of the stimulus (Galambos & Davis, 1943; Rose, Brugge, Anderson & Hind, 1967). When the carrier frequency is low enough, the responding neurons will phase lock to both the carrier and the envelope of modulation frequency (Joris & Yin, 1992; Javel, 1980). In cases where the carrier frequency is too high for the neurons to phase lock to it, they will only lock onto the envelope of the modulation frequency. In the case of the high carrier frequency, the period histogram would appear to be that of an individual low frequency tone (Khanna & Teich, 1989; Javel, 1980). Spectral components which are not present in
the stimulus can be found in the period histogram of a modulated stimulus regardless of carrier frequency.

Joris and Yin (1992) as well as Yates (1987) showed that the magnitude or precision of phase locking to a SAM signal is dependent on the stimulus level and that it is not a linear function relative to the stimulus intensity. This dependence on the magnitude of the phase locked response is one reason that AEPs cannot be detected at the same level as the behavioral threshold. The magnitude of the phase locked response will become undetectable at the relatively remote electrode sites before the modulated signal becomes inaudible to a listener. Additionally, in a simple detection task the listener may be able to detect the stimulus at levels below where they can perceive the modulation in the signal. The electrophysiologic response does not share in this ability and can only be detected when the modulation is being encoded at a sufficient magnitude to be detected in the far field. The cochlear nucleus (CN), at least some cell types located within it, has been shown to be especially tuned to SAM stimuli. For example the chopper cells of the CN have been shown to phase lock to the modulation envelope of a SAM stimulus (Rhode & Greenberg, 1994). Chopper cells comprise a large portion of the cells found in the posteroventral cochlear nucleus (PVCN), although they are also present in other areas of the cochlear nucleus just not in as great numbers. Despite the fact that choppers do not phase lock particularly well to pure tones they are much better at phase locking to amplitude modulation (Rhode & Greenberg, 1994). This proclivity toward modulated signals makes them a very likely contributor to responses such as the ASSR. In the inferior colliculus (IC), neurons whose characteristic frequency is that of the SAM tone carrier frequency will also phase lock to the frequency of the modulation.
The cells of the IC appear to have band-pass characteristics when it comes to following the modulation envelope— they respond best within a certain range of frequencies and respond poorly above and below that range (Moller, 1976; Kim, Sirianni & Chang, 1990). Langer and Schreiner (1988) successfully recorded responses from the IC with modulation frequencies of 10-1000 Hz but found the greatest number of units responded to modulation frequencies between 10-300 Hz and especially between 30-100 Hz. This band pass characteristic helps explain why specific ranges of modulation rates are used depending on the level of the auditory nervous system being studied.

Frequency Modulation

Demodulation of the frequency modulation of a stimulus has been observed in responses recorded from afferent auditory nerve fibers in at least two studies (Britt & Starr, 1976; Sinex & Geisler, 1981). When a demodulator, in this case the auditory system, demodulates a FM signal the information obtained is a breakdown of frequency variation relative to the carrier at any given instant and produces an amplitude output that represents the instantaneous frequency variations of the signal (Electrical engineering training series, http://www.tpud.com/neets/book12/51c.htm). It was proposed in Popper and Fay (1992) that the above findings might indicate that the auditory system is first converting the FM signal to an AM signal through cochlear frequency filtering. This position is supported by several studies (Sinex & Geisler, 1981; Khanna & Teich, 1989; Saberi & Hafter, 1995; Quatieri, Hanna, & O’Leary 1997). Saberi and Hafter (1995) demonstrated that FM and AM share a common neural code at the level of the brainstem as evidenced by presenting an FM signal to one ear and an AM signal to the other with equal modulation and carrier frequencies. In the CN the response to an FM stimulus
varies by the type of cell responding to the stimulus. Some cell types show differences in
their responses based on frequency sweep direction (rising or falling) with some fibers
responding best to upward frequency sweeps and others responding best to downward
frequency sweeps (Watanabe, 1972). More recent work in Rhesus monkeys showed
neurons were highly selective in their responses to FM stimuli based on modulation rate
and direction (Rauschecker, 1997). Kay and Matthews (1972) showed that AM and FM
signal processing occurs through separate channels within the auditory nervous system.
Beyond the CN in the brainstem FM sensitive neurons again show band pass
characteristics and are most sensitive (to as little as a 1% frequency excursion) at
modulation frequencies of less than or equal to 100 Hz. The signals used in the Kay and
Matthews (1972) study were sinusoidally modulated as are the signals used in ASSR
recordings. As was seen in the CN, some fibers respond better to a particular direction of
frequency sweep. In the thalamus and cortex the neurons are still responding in a
periodic fashion and some early studies found cells in the medial geniculate body (MGB)
and auditory cortex which were exclusively responsive to FM signals and some that
responded to both FM and TBs but better to FM (e.g., Whitfield & Purser, 1964; Purser
& Whitfield, 1972; Watanabe, 1972; Nomoto, 1980).

_Toneburst (transient)_

In addition to the continuous modulated stimuli, more recent work in ASSR
recording has introduced trains of brief TBs. The greater spectral width of a short
duration TB, relative to a long duration tone, likely results in a larger population of
neurons being stimulated with the short duration tone. An increase in the number of
neurons firing in response to a stimulus translates into an increase in response amplitude
(Folsom, 1984). Additionally the sharp onset of short duration sounds results in greater neural synchrony than a smoother onset longer duration tone. The neural response to a short duration signal such as a click or toneburst can also be used to reasonably predict the response to pure tone stimuli or evaluate the integrity of the auditory nerve as in ABR testing (Don, Eggermont & Brackmann, 1979; Stapells & Picton, 1981; Stapells, Picton, Perez-Abalo, Read & Smith, 1985). This synchrony due to the abrupt signal comes at the cost of frequency specificity as the short duration has a negative impact on the frequency specificity of the signal. A click will primarily reflect high frequency hearing because by the time it reaches the apex the basal regions have already responded and the traveling wave (along the basilar membrane) becomes less sharp resulting in reduced synchronicity of firing for the fibers in that region. TBs, such as those used in ABR and more recently ASSR, represent a compromise between a continuous modulated signal and a click. The slightly longer signal improves frequency specificity relative to a click but is still brief enough to maintain neural synchrony.

2.2.2 Transient Responses

In a transient response (e.g., ABR), each response is evoked by a single stimulus and terminates after the presentation of that stimulus. Put another way, each response concludes prior to the subsequent stimulus (and response) and will not occur again until after the presentation of the next stimulus.

Auditory Brainstem Response (ABR)

The ABR, also known as the brainstem auditory evoked response (BAER) was first described in a 1970 paper by Jewett, Romano and Williston. In this paper Jewett et al. (1970) described a series of waveforms, stimulated by a click, which occurred
between approximately 2 and 7 milliseconds following the stimulus presentation. This paper also notes that the latency was remarkably consistent with repeated measures both within and between subjects (Jewett et al., 1970). Reversing the polarity of the stimulus did not result in a change in the polarity of the recorded waveforms which confirmed that the responses being recorded were not the cochlear microphonic.

In the normal auditory system the response waveforms are known to occur at certain latencies. Any significant departure from these latencies may be indicative of a pathological condition (e.g., hearing loss, retrocochlear pathology). The click used in ABR testing is very brief (typically 100 μs) which produces the neural synchrony necessary for a strong response. However, as the duration of a signal decreases the frequency spectrum will increase (Burkard, 1984). The click ABR is generally considered to mostly reflect the hearing sensitivity in the 2000-4000 Hz range (Gorga, Neely, Hoover, Dierking, Beauchaine, & Manning, 2004; Bauch & Olsen, 1985). Therefore, the threshold for a click ABR will tell the approximate hearing sensitivity of the tested ear at the frequency with the best hearing sensitivity within that frequency range, but will not give thresholds for any specific frequency and therefore has relatively poor frequency specificity.

The use of TBs rather than clicks makes it possible to obtain more frequency specific information with the ABR. A common duration for TBs used in ABR is 5 cycles (2 cycle rise-1 cycle plateau-2 cycle fall) meaning duration will vary by frequency of the TB. Clearly the TB is substantially longer in duration than the click and more frequency specific but still short enough to provide neural synchrony to generate a response. ABR testing of each frequency individually results in substantially longer test times compared
to the multiple simultaneous stimulus paradigm in ASSR which can provide similar diagnostic information in a shorter period of time. Testing four frequencies using individually presented TB stimuli would require at least eight recordings per ear, considerably increasing test time relative to a click stimulus, but providing considerably more information about hearing sensitivity. The amount of time an infant or child can be expected to remain quiet or still enough for testing is limited so speed is of particular importance. The ABR has been widely used in infant hearing screening programs, but in addition to its use as a screening tool, ABR can also be used to provide an estimate of hearing sensitivity across a range of frequencies if TB stimuli are used and has been shown to provide threshold predictions within approximately 10 dB (Sininger, Abdala, & Cone-Wesson, 1997; Stapells, Picton, Durieux-Smith, Edwards, & Moran, 1990).

2.2.3 Steady State Potentials of the Auditory System

Steady state responses are evoked when the changes in a stimulus occur at a rate that prevents the transient potential evoked by each individual stimulus occurrence from returning to the pre-stimulus baseline before the next response is elicited (Regan, 1966; Picton, John, Dimitrijevic & Purcell, 2003), this inability of the auditory nervous system to return to pre-stimulus baseline results in a response that is continuous and present for as long as the stimulus is presented. Steady state responses driven by a sound stimulus are represented by power spectra at specific frequencies and are not visible in the ongoing time waveform of the EEG. Viewing these responses requires the use of Fourier analysis and averaging techniques (Geisler, 1960). Although Galambos, Makeig and Talmachoff, (1981) are generally credited with being the impetus behind the investigation of using steady state responses in hearing evaluation, several other studies had investigated steady
state responses of the auditory systems in the 1960s and 1970s (e.g., Geisler, 1960; Schimmel, Rapin, & Cohen, 1974; Chatrian, Petersen, & Lazarte, 1960; Campbell, Atkinson, Francis, & Green, 1977; Hall, 1979b). Just as in the recording of transient responses, the signal to noise ratio of the response is enhanced by averaging multiple recordings.

The term “auditory steady state response” or ASSR has come to be associated with a specific response, but in broader terms it represents a class of responses that can be recorded from the auditory system. A brief discussion of other steady state potentials of the auditory system is provided to illustrate this point.

40 Hertz Response (40 Hz ERP)

Some of the first auditory steady state responses used modulation rates that were much lower than those used in what is known as the ASSR today. The 40 Hz response or 40 Hz event related potential (ERP) was first reported by Galambos et al. (1981). One of the stimuli used to elicit this response was a low frequency (500 Hz) tone pip. It was initially thought that these potentials could be used to obtain information about low frequency hearing and fill a gap left by the click ABR that primarily provides information about higher frequencies. The 40 Hz response as described by Galambos et al. (1981) is believed to consist of overlapping components of the auditory middle latency response (AMLR). Unfortunately, the responses at these lower modulation rates are largest in amplitude when recorded from higher levels in the brain and degrade with the onset of sleep or sedation. This should not be interpreted to mean that lower structures are not encoding the lower frequency modulation, they must be or how would this information reach higher centers in the brain? Herdman, Lins, van Roon, Stapells, Scherg, and Picton
(2002) recorded responses from brainstem and cortical generators from numerous sites on the scalp (46 electrodes) and found that the cortical sources produced larger responses than brainstem sources. A larger response will be easier to discern from the background noise and require less averaging time.

Envelope Following or Amplitude Modulation Following Response (EFR or AMFR)

A steady state evoked potential referred to as the envelope following response (EFR) has been described by numerous studies (Dolphin, Chertoff & Burkard, 1994; Dolphin & Mountain, 1992, 1993; Rickards & Clark, 1984; Kuwada, Batra, & Maher 1986). The name EFR reflects the phase locking of the response to the envelope of the stimulus (Dolphin, et al., 1994). Some of the aforementioned studies used amplitude modulated signals so the EFR is also referred to as the amplitude modulation following response (AMFR). In the aforementioned studies the response was elicited with a SAM tone and a two-tone complex. The two-tone complex consisted of two pure tones close in frequency played simultaneously. The response is located at the modulation frequency for a SAM tone and at the difference frequency (between the two tones) for the two-tone complex. Dolphin et al. (1994) found that the characteristics of responses obtained with SAM and two-tone stimuli were very similar and suggestive of the responses representing the same underlying physiology. The EFR has been shown to vary in latency based on the modulation frequency indicating that different generators are producing the response with different frequencies of modulation. Responses elicited with lower modulation or difference frequencies have longer latencies suggesting a more central site of origin (Dolphin et al., 1994). The effects of an interference tone (a tone that was not part of the stimuli) was also investigated by Dolphin et al. (1994) where they found that the further
removed in frequency the interference tone was the less impact it had on the response. This finding is analogous to today’s practice of presenting multiple simultaneous stimuli in ASSR testing with minimal effect on the responses provided the stimuli are spaced sufficiently apart as shown by John et al. (1998).

Dolphin and Mountain (1993) reported small (1-3 dB) decreases in response magnitude when multiple two-tone pairs were presented to one ear simultaneously. This is similar to the small effects that have been demonstrated with ASSR using a SAM tone (John et al., 1998). What this all boils down to is that the EFR, AMFR and the ASSR are the same response; ASSR has simply become the preferred term for referring to a steady state evoked potential of the auditory system elicited by a modulated stimulus.

2.3 The Auditory Steady State Response (ASSR)

The first descriptions of auditory steady state responses, as discussed previously, were made using lower modulation frequencies (20-40 Hz) than are typically used to measure ASSRs today (70-110 Hz). The primary problem associated with lower frequency steady state responses is that the responses are substantially attenuated during sleep or with sedation reducing their clinical utility in the populations where they would be most helpful. Generating the responses at higher modulation frequencies eliminates state of arousal as a confounding factor as the response is still measurable during sleep or sedation. The resolution of these problems is really a matter of physiology, rather than modulation frequency itself, as certain regions of the brain respond best at different frequencies and are impacted to different extents by state of consciousness (Stapells & Picton, 1981, Galambos et al., 1981). Therefore, what really makes the difference in
recording steady state responses is the site of generation more than the modulation rate itself. There is nothing inherently special about the 70-110 Hz modulation frequency range; it just happens that the brainstem, which is more resistant to sleep and anesthesia effects than are higher levels of the brain, responds best at that range of modulation rates.

2.3.1 Response Generators

The exact neural generators of the ASSR have not been clearly defined, but it has been shown that the general locations of the generators are at least somewhat dependant on the modulation frequency of the stimulus (e.g., Herdman, et al., 2002a; Arnold & Burkard, 2002; Szalda & Burkard, 2005, Cohen, Rickards & Clark, 1991, Herdman et al., 2002b). Generators most sensitive to lower modulation frequencies (~40 Hz) have been shown to be primarily cortical whereas the generators most sensitive to high modulation frequencies (~70-110 Hz) are primarily at the brainstem level (e.g., Herdman et al., 2002; Kuwada, Anderson, Batra, Fitzpatrick, Tessier and D’Angelo, 2002; Dobie & Wilson, 1994). Recent work, however, has shown that even in the high frequency ASSR, the cortex may also be making a contribution to the response (Arnold & Burkard, 2002; Szalda & Burkard, 2005).

In a 2002 study, Arnold and Burkard measured the EFR (i.e., the ASSR) with electrodes surgically implanted in both the inferior colliculus (IC) and the auditory cortex of a chinchilla using a two tone complex as their stimulus. They measured responses with difference tones of 20-320 Hz. The difference tone is the distance (in Hz) between the two tones and is analogous to the modulation frequency in a sinusoidal amplitude modulated tone. Difference tones of 80 Hz and 160 Hz produced the largest response in the inferior colliculus. For lower intensity stimuli (30-50 dB SPL) 80 Hz was the largest
response while at higher intensities (60-80 dB SPL) 169 Hz provided the largest response. Arnold and Burkard’s (2002) recordings showed contributions from both the inferior colliculus and the auditory cortex with the cortical component being substantially attenuated upon introduction of anesthesia. Further evidence of the inferior colliculus as the primary generator site for the high frequency ASSR is provided by Szalda and Burkard (2005) who found peaks in response amplitude with modulation frequencies of 109 Hz and 170 Hz using an electrode implanted in the inferior colliculus of a chinchilla. In this same study, Szalda and Burkard (2005) also recorded from an electrode implanted in the auditory cortex where they found a relatively flat amplitude function up to 70 Hz, above which amplitude decreased substantially. In anesthetized animals Szalda and Burkard (2005) noted a greater than 50% decrease in cortical response amplitude in the 70-210 Hz response range. Further evidence of different generators based on stimulus modulation frequency is provided by differences in latencies for responses evoked with different modulation frequencies. Kuwada et al. (1986) found that with modulation frequencies less than 55 Hz, the latency of the response was approximately 30 msec. When the modulation frequency was between 100-350 Hz, however, the latency of the response was 7-9 msec. In their 1992 study, Dolphin and Mountain reported latencies of approximately 12 msec when the modulation frequency was 10-50 Hz and 6.6 msec when the modulation frequency was 50-180 Hz. Latency estimates were calculated by examining the phase delay of the responses.

Kiren, Aoyagi, Furuse, and Koike (1994) demonstrated IC involvement with an ablation study in cats. AMFRs were measured with modulation frequencies between 20 to 120 Hz. Responses were recorded from the IC for all modulation frequencies with the
exception of 20 Hz. Following placement of a lesion in the IC, the AMFR was not recordable at any modulation frequency. Kiren et al. (1994) also recorded ABRs before and after placing lesions in the IC and there was no difference in pre/post lesion ABR. The Kiren et al. (1994) finding is of note since it has been proposed that the ASSR and ABR have the same generators (Stapells, Herdman, Small, Dimitrijevic, & Hatton, 2004). The Kiren et al. (1994) results should be given greater emphasis in weighing this matter as they provide physiological evidence to back up their findings, while in the Stapells et al. (2004) article the hypothesis that the ASSR consists of overlapping ABR wave Vs is not supported by any data. The Stapells et al. (2004) hypothesis is clearly not supported by the findings of Kiren et al. (1994). Based on anatomical studies, it would not appear that wave V of the ABR and ASSR represent the same neural activity.

The data from animal studies show that the IC tends to produce the most robust responses when a high modulation frequency is used, but one should be cautious in applying the findings in these and other animal studies to humans. First, as Szalda and Burkard (2005) point out in their study, the animal recordings are done in the near field whereas scalp recordings in a human are far field potentials. Differences in responses recorded in the near field from the IC and auditory cortex may not project exactly to responses recorded from the same structures in the far field due to the increased distance from the generators and how the generating neurons are aligned. The evidence currently available tends to point to multiple generators with the brainstem being the primary contributor at higher modulation frequencies and the cortex being the primary contributor when a lower modulation frequency is applied to the stimulus.
2.3.2 Stimuli Definitions

**Sinusoidal Amplitude Modulated (SAM) Signals**

Modulation of the amplitude of a carrier frequency is a common way to create a signal that will stimulate an ASSR. This amplitude modulation is generally sinusoidal in nature and can vary in the depth of modulation. When describing SAM stimuli, the depth of modulation \( m \) is generally referred to as a percentage between 0 and 100%. The SAM stimuli used in ASSR studies almost always utilize a modulation depth of 100%. For example, if the unmodulated amplitude of a signal was equal to 1 and this signal was amplitude modulated at 100%, the modulated signal’s amplitude would travel between 0 and 2. In addition to the depth of modulation, a SAM signal will also have a rate at which the modulation occurs per unit time (generally seconds). This value is completely independent of the depth of modulation, so a change in the modulation rate in a SAM signal will have no effect on the depth of modulation and vice versa. A pure tone that is modulated with sinusoidal amplitude modulation has a relatively simple acoustic spectrum consisting of three spectral peaks. These peaks consist of the carrier frequency and the carrier frequency +/- the modulation frequency. These two sidebands will be of equal amplitude and one-half the amplitude of the carrier frequency with 100% modulation.

**Frequency Modulated (FM) Signals**

Another type of stimulus which can and has been used to generate an ASSR is a pure tone that is modulated in frequency rather than amplitude. In FM, the signal varies in frequency to some degree around the carrier. Recall that in SAM the depth of modulation is generally referred to as a percentage, but in frequency modulation the
modulation index ($\beta$) is more complicated. In frequency modulation, $\beta$ represents the change in frequency relative to the modulation rate and not a simple change in frequency based on a percentage of the carrier frequency (Hartmann, 1998). So unlike in SAM where changing the rate of modulation has no effect on the depth of modulation, changing the frequency of modulation in an FM signal will have an effect on the value of the modulation index ($\beta$). For example if a 1000 Hz tone were modulated at a rate of 90 Hz ($f_m$) with a frequency excursion of 20% ($\Delta f$) then $\beta=\Delta f/f_m=200/90=2.22$. If however, we change the modulation frequency so that $f_m=70$ then $\beta=200/70=2.85$ illustrating how $\beta$ is affected by changing the frequency of modulation.

If the $\beta$ of an FM signal is small ($\leq 1.57$) then it will have the same power spectrum as an AM signal (Hartmann, 1998). In cases where the narrow band frequency modulation (NBFM) assumptions are not met, the spectrum of the signal consists of the carrier frequency +/- a large number of progressively smaller sidebands ($f_m, 2f_m, \ldots, n f_m$). It is important to note that in the ASSR literature, frequency modulation is not typically referred to by the modulation index or $\beta$ value. Rather the amount of frequency modulation is referred to by the frequency excursion expressed as a percentage. In essence it is being quantified by the numerator of the modulation index equation.

**Mixed Modulation (AM+FM) Signals**

Most ASSR studies and the majority of commercially available ASSR test equipment use a type of signal referred to as mixed modulation (MM), which consists of a pure tone that is modulated using both amplitude and frequency modulation. The two different modulation modes typically use the same modulation rates, but at least one study (Dimitrijevic, John, van Roon, & Picton, 2001) utilized different modulation
frequencies for AM and FM components. The most commonly used MM signals in ASSR studies utilize 100% AM and what they refer to as 10-25% FM (refering to the frequency excursion as a function of the carrier frequency) rather than specifying a $\beta$ value (e.g., Vander-Werff, Brown, Gienapp, & Schmidt-Clay, 2002; John 2001; Rance & Rickards, 2002).

**Tonebursts**

Another potential stimulus for generating an ASSR is a train of brief TBs. With a TB stimulus the response will be located at the repetition rate of the stimulus. A TB stimulus repetition rate is analogous to the modulation frequency in SAM, FM and MM stimuli. A TB consists of a brief duration sinusoid and can be defined in terms of its rise time, plateau, and fall time. The rise and fall times refer to the amount of time it takes the signal to go from zero to maximum amplitude and vice versa respectively. A plateau represents the amount of time the stimulus remains at the maximum amplitude. It is not a requirement for a plateau to be present; the TB can immediately start decreasing in amplitude once it reaches the maximum value. TBs can be of varied length but if it is too short then the signal becomes more click like and loses frequency specificity. However, if the signal is too long than it will not be as effective at generating the onset responses that a TB stimulus is typically used to elicit in AEP testing. In the current study a Blackman window was applied to the TB stimuli. A TB generated in this manner is essentially overmodulated SAM tone. Overmodulation occurs when the amplitude modulation exceeds 100%, resulting in additional sidebands compared to the single upper and lower sidebands seen in a SAM tone that is not overmodualted. This effective overmodulation occurs due to the Blackman window that is applied to each toneburst.
2.3.3 Recording Techniques

*Individual vs. multiple simultaneously presented stimuli*

ASSR can be recorded using individual or multiple simultaneously presented carrier frequencies from either one or both ears simultaneously. Extensive work has been done investigating the effects of presenting multiple simultaneous stimuli using SAM and MM stimuli (e.g., Dimitrijevic et al., 2002; John et al., 1998; John et al., 2002; Lins & Picton, 1995) but there is a relative paucity of published data using TB stimuli. Animal work by Dolphin (1996, 1997) demonstrated that presenting multiple simultaneous stimuli resulted in some small amplitude decreases but, these amplitude changes did not significantly impact the ability to record responses. It should be noted, however, that Dolphin (1997) used a wide range of modulation rates and that the greatest interaction effects were seen at frequencies outside the range typically used for ASSR in humans. John et al. (1998) saw similar effects in humans using multiple simultaneous stimuli with modulation frequencies around 40 Hz which showed more stimulus interaction than in the 80 Hz modulation frequency range. Multiple simultaneous stimulus presentation does result in some response attenuation (especially for lower carrier frequencies) at least in normal hearing subjects but, the position taken in the majority of the literature is that the time savings still outweighs the effects of stimulus interaction (e.g., John et al., 1998; Dimitrijevic, et al., 2002; John, Purcell, Dimitirjevic, & Picton, 2002). Stimulus interaction effects can also result in an enhancement of the response amplitudes rather than a decrease, in animal models (Dolphin & Mountain, 1993; Dolphin, Chertoff, & Burkard, 1994) and in humans (John et al., 1998). Most of the work investigating stimulus interaction effects of multiple stimuli has been done in normal hearing subjects.
and there is some evidence to show that in at least some hearing impaired subjects there is greater interaction between the stimuli (Dimitrijevic et al., 2002). This difference was attributed to two possible factors: (1) the higher stimulus levels needed to generate a response in the hearing impaired subjects, and (2) recruitment may be enhancing responses in hearing impaired ears (Dimitrijevic et al., 2002).

2.3.4 Response Analysis

As is true of all auditory evoked potentials, the ASSR must be detected in the background of unrelated neural activity (noise) and any other electrical noise in the environment (Picton, Dimitrijevic, John, & Van Roon, 2001). The periodicity of the ASSR results in a fairly simple frequency representation and can be efficiently analyzed with frequency based techniques such as Fourier analysis (Stapells, Makeig, & Galambos, 1987). The vast majority of published ASSR studies have utilized an F-test based on the discrete Fourier transform (DFT) of the recorded physiologic activity to determine if the response amplitude significantly differs from the surrounding noise.

An important caveat in these types of analyses is that a non-response result does not necessarily mean that a response was not present; it is possible that a response was present but too small to be detected with the chosen analysis parameters. Because of this it is important to consider what is meant by the term “threshold”. An ASSR threshold in its broadest terms is the lowest level at which the desired response can be recorded and differentiated from other non-related activity. Depending on how the criterion for a present response is defined, the resulting threshold can vary greatly. If the response detection criteria are changed, then threshold may shift as a response once considered too small to be detected now meets the response criteria. This flexible definition of threshold
is one of the major problems in interpreting an ASSR data given the lack of standard stimuli and analysis methods and the fact that an electrophysiologic threshold is not defined in the same manner as a behavioral threshold.

F-Statistic

The most commonly used method for detecting a response in ASSR is to do a power spectrum analysis with an F-statistic. Zurek (1992) pointed out that the power estimates at the modulation frequency and noise bins (adjacent frequencies) have a chi-square distribution. The chi-square distribution allows an F-statistic to be used in comparing the response and noise levels (Dobie & Wilson, 1996). Before the F-statistic can be applied, however, the recorded responses must first be transformed to the frequency domain. This is accomplished using fast Fourier transform (FFT) which yields real and imaginary components at each frequency bin.

The F-statistic as it is currently used in ASSR ignores phase information and only examines amplitude. If only noise is present in the response bin a normal distribution (mean zero, variance of noise power/2) will be observed. The power estimate at the modulation frequency has degrees of freedom equal to two (the real and imaginary components). The noise power estimate is the averaged power of the noise across all noise bins with degrees of freedom of 2m where (m) equals the number of noise bins resulting in $df = 2,2m$. The question that now arises is how many noise bins should be utilized? Increasing the degrees of freedom will provide greater statistical power; however, Dobie and Wilson (1996) found diminishing returns when values of m became very large. However, this was at least in part based on the computational demands placed
on the computer and is likely not as significant a problem given the improvements in computer technology since that paper was published.

Longer recording periods can improve SNR, but are not without cost. The most obvious cost being increased recording time for response acquisition. Additionally, since the response has a bandwidth, if the analysis bandwidth becomes narrower than the bandwidth of the response there can be spill-over of response energy into the adjacent noise bins. One must be careful, therefore, not to use a recording length that results in an analysis bandwidth narrower than the width of the response itself. Failure to do so will result in response energy spilling into one or more noise bins which will result in an inaccurate (elevated) noise estimate potentially resulting in an artificial increase in the threshold of the response. Lastly, noise (movement, tensing muscles, etc.) in physiologic recordings tends to be non-stationary and introduces large artifacts requiring entire recordings to be discarded. Using shorter recording periods and averaging the recordings help with this problem as throwing out the shorter intervals has a smaller impact on the overall recording time.

2.4 Advantages/Disadvantages of ASSR

Frequency Specificity

The vast majority of recent work in ASSR has been done using SAM or MM tones (100% AM and 10-25% FM). The acoustic specificity of a SAM tone is purported to be better than an ABR using TB stimuli in terms of acoustic frequency specificity (Stapells et al., 2004). However, a more acoustically specific stimulus does not necessarily translate into better cochlear place or neural specificity. Brainstem responses to both Blackman and linear windowed tonebursts have been shown to have similar
cochlear place and neural specificity despite the narrower acoustic spectrum of the Blackman gated tone compared to the linear gated tone (Oates & Stapells, 1997; Purdy & Abbas, 2002). Acoustic specificity of the stimulus is not unimportant, but cochlear place specificity are more important in making frequency specific estimates of hearing sensitivity (Herdman, Picton & Stapells, 2002; Picton, Dimitrijevic, & John, 2002). Cochlear place specificity refers to the portion of the basilar membrane that is being stimulated and contributing to a response (Oates & Stapells, 1997). The cochlear place specificity of the sinusoidally amplitude modulated tones (the narrowest of ASSR stimuli) has been shown to be similar to that of tonebursts (a comparatively wideband signal) used in ABR in both normal and hearing impaired ears (Herdman et al., 2002; Herdman & Stapells, 2003).

Recording Time

One of the purported advantages of ASSR over other AEPs is the reduced amount of time required to estimate an audiogram (John, Lins, Boucher & Picton, 1998; John & Picton, 2000b; Herdman & Stapells, 2001; Picton, Dimitrijevic & John, 2002). The primary populations tested using ASSR are infants and young children who cannot be counted upon to be quiet or still long enough to complete the test unless they are asleep or sedated. A shorter test time increases the chances of obtaining all the desired information during a single session with natural sleep and in the case of chemical sedation the amount of time the child must be sedated is reduced. This time savings comes from the ability to record multiple carrier frequencies from either one or both ears simultaneously in ASSR. Recording four carrier frequencies simultaneous has been shown to be 2-3 times faster than recording the response for each of those frequencies.
individually (Dimitrijevic et al., 2002; John, Dimitrijevic & Picton, 2002). Cone-Wesson et al (2002) compared ASSR and ABR recording times and reported needing up to 240 seconds per trial for their ABR recordings and no recordings taking longer than 104 seconds for ASSR. The ASSR system utilized in their study (Audera) presents only individual stimuli so testing four frequencies with ASSR could have taken up to 416 seconds versus up to 960 seconds for 4 frequencies of toneburst ABR. There are, however, some ambiguities in this study that make it difficult to accurately and fairly compare ASSR and ABR in terms of recording time. First, no mean times or standard deviations were reported by Cone-Wesson et al. (2002a); only the maximum time for each measure. Were these maximum times outliers or did most subjects need close to the maximum recording period? Without this information it is difficult to determine just how advantageous ASSR was in terms of recording time.

Secondly, the response decision for ASSR was made using the automatic detection algorithm in their evoked potential system (Audera) while the ABR response decision was made by a human observer and did not appear to have a defined stopping criteria (e.g., a fixed number of averages). An automated response detection algorithm may have been able to detect a response in the ABR testing faster than the human observer and made for a more equitable comparison. On the surface, the ASSR seems like a hands down winner in terms of test time, however, when the finer points of the comparison to ABR are examined it would appear that the time savings of ASSR is, perhaps, somewhat overstated.
2.4.1 Interaction of multiple simultaneous stimuli

Whereas it has been demonstrated that multiple simultaneous stimuli can be presented with minimal effects on the individual responses to the carrier frequencies with a SAM stimulus (John, et al., 1998; Herdman & Stapells, 2001, 2003; Herdman, Picton & Stapells, 2002), this same phenomenon has not been demonstrated with a TB stimulus. Given the substantially wider acoustic spectrum of the TB stimulus compared to a SAM stimulus, the potential exists for interaction between the stimulus components that could affect response amplitude.

2.4.2 Profound Hearing Loss and ASSR

Another purported advantage of the ASSR over ABR has been its ability to record responses from individuals with severe to profound hearing losses; a population where ABR cannot accurately evaluate hearing sensitivity due to output limits of current equipment. It has been thought that this additional response recording range would be helpful in identifying areas of residual hearing in profoundly hearing impaired individuals. More recent work, however, has shown that many of the responses elicited with high-level stimuli may in fact be artifactual or non-auditory in origin. Artifactual responses to high intensity stimuli have been recorded from air and bone conduction in several labs from individuals who were known to be deaf or profoundly hearing impaired and who denied hearing any of the stimuli that were presented (Gorga et al., 2004; Jeng, et al., 2004; Picton & John, 2004; Small & Stapells, 2004).

Picton and John (2004) and Small and Stapells (2004) have shown that at least in some cases what appeared to be a response was due to an artifact in the EEG from the high intensity stimuli. When the analog to digital conversion rate and EEG filters were
adjusted to remove aliasing energy many of the responses disappeared. In the Small and Stapells (2004) study, however, some responses remained at 500 and 1000 Hz in the deaf subjects who denied hearing the stimulus. Other studies using transient AEPs (e.g., Murofushi, Iwasaki, Takai, &Takegoshi, 2004; Welgampola & Colebatch, 2001; Sheykholeslami & Kaga, 2001) have demonstrated vestibular responses to high intensity (particularly low frequency) stimuli which may account for the remaining responses found after adjusting recording parameters in the 2004 Small and Stapells study.

2.4.3 Clinical applications of ASSR

The primary clinical application of the ASSR is in diagnostic testing of infants and other populations who are not able to provide accurate responses for behavioral threshold measures. Whereas ASSR is not the only measure that can be used to test these populations it has the purported advantage of being able to test multiple frequencies simultaneously and binaurally without significantly impacting the accuracy of threshold predictions. Even if frequencies were to be tested individually, it still allows for simultaneous testing of both ears and eliminates the need for subjective analysis of waveforms by an examiner. This allows threshold information to be collected in a shorter period of time than other methods used to objectively collect threshold data based on auditory electrophysiological measures. Determining approximate thresholds of hearing based on ASSR data allows an infant to be fit with amplification much sooner than they would be if the audiologist had to wait until they were old enough to provide reliable behavioral responses. By obtaining these estimates when the child is just a few months old, they can be fit conservatively with amplification and be provided with at
least some auditory stimulation until they are mature enough to provide behavioral responses at which time the amplification settings can be fine-tuned.

Analysis of ASSR data does not stop with the detection of the response itself. The ASSR threshold must now be adjusted to provide an estimate of the behavioral threshold. One issue in making this correction is the fact that in hearing impaired ears the ASSR can typically be recorded closer to behavioral threshold (probably due to recruitment) than in a normal hearing ear (Picton et al., 2001). Due to this finding, a regression model may prove to be a better model than simply subtracting a correction factor (i.e. an average difference between physiological and behavioral threshold) which might result in an underestimation of behavioral threshold in a hearing impaired ear (Picton et al., 2001).

While much has been learned about the ASSR and its application to clinical estimation of hearing sensitivity there are still numerous questions yet to be answered. It is important that we fully understand the effects that stimulus type, stimulus presentation method, response detection method and estimation procedures have on the ability of the ASSR to accurately predict the hearing sensitivity of an ear. The current study will seek to determine the impact of stimulus on the response by using identical collection and analysis protocols leaving stimulus as the only variable. Additionally it has not been demonstrated that a relatively new stimulus (TB) can be utilized with similar immunity to interaction between stimulus components as has been shown with a SAM stimulus. Finally, different methods have been employed to try and estimate a behavioral hearing threshold based on an ASSR threshold and these methods will be compared to determine the most accurate estimation method.
CHAPTER 3
METHODS

3.1 Subjects

Ten adult subjects (8 female, 2 male) completed the protocol for this study. Subjects were between 21 and 38 years of age (mean age = 27.5 years, SD = 5.2 years) and were recruited from the student population at The Ohio State University. Prior to being included in the current study potential subjects were required to meet the following criteria: (1) normal hearing sensitivity (thresholds no worse than 15 dB HL) from 250-8000 Hz; (2) tympanometry within normal limits for young adults (Roup, Wiley, Safady, & Stoppenbach, 1998); (3) no diagnosed otologic disease (e.g., Meniere’s, otosclerosis, etc.); and (4) no diagnosed neurological disorders (e.g., multiple sclerosis). To ensure middle ear function met the required criteria throughout the study, tympanometry was completed prior to each testing session. Preliminary exam and case history forms used in the current study can be found in Appendix A.

3.2 Stimuli

The stimuli used to measure behavioral and electrophysiological thresholds in this study consisted of: (1) pure tones (PT) for behavioral measures only, (2) sinusoidally amplitude modulated (SAM) tones, (3) mixed modulation (MM) tones, and (4) trains of brief tonebursts (TB). For all stimulus types, frequencies of 500 Hz, 1000 Hz, 2000 Hz
and 4000 Hz were used. During ASSR testing, carrier frequencies (CF) of 500, 1000, 2000, and 4000 Hz were presented for each stimulus type using modulation frequencies (MF) of 79.10 Hz, 86.91 Hz, 94.72 Hz, and 102.53 Hz respectively, reflecting the Fourier frequencies. Behavioral thresholds were tested using the above modulation rates rounded to the nearest integer values (79 Hz, 87 Hz, 95 Hz, and 103 Hz). All stimuli used during ASSR testing were generated by the Intelligent Hearing Systems (IHS) SmartEP system and passed directly to a set of insert earphones (Eartone 3A) connected to the IHS system. All stimuli used in behavioral testing were generated in MATLAB (Mathworks) and passed to a set of headphones (Sennheiser 850P).

3.2.1 Pure Tones

The pure tone stimuli utilized in this study were one second in duration and were only used in measuring behavioral thresholds. Pure tones can be represented mathematically by the following formula

\[ x(t) = A \sin(2\pi ft+N) \]  

(3.1)

where \(x(t)\) is the instantaneous amplitude of the signal, \(A\) is the maximum amplitude of the signal, \(f\) is the frequency of the signal, \(t\) is time and \(N\) is the starting phase of the signal. The amplitude of the pure tone and all other signals in the study were set to the maximum value allowed by the TDT system prior to being passed to the attenuators. Starting phase for all pure tone stimuli was 0 degrees.
3.2.2 Sinusoidal Amplitude Modulated (SAM) Tones

The SAM tones used in this study were one second in duration and amplitude modulated at 100%. The spectrum for these SAM tones consists of three acoustic peaks located at the CF and the CF±MF (e.g., if CF= 500 Hz and MF=79Hz then peaks of: 421, 500, 579 Hz comprise the signal). The SAM signal used in the current study can be described using the following formula

\[ x(t) = [1 + m \sin(2\pi F_m t)] \sin(2\pi F_c t) \]  

(3.2)

where \( x(t) \) is the instantaneous amplitude of the signal, \( m \) is the modulation index, \( F_m \) is the frequency of modulation, \( t \) is time, and \( F_c \) is the carrier frequency. Figures 3.1 and 3.2 depict several cycles of the SAM time waveform and the power spectra respectively. Each figure consists of four panels, one for each of the carrier frequencies. The time waveforms show signal amplitude as a function of time (ms) and the power spectra show the magnitude (dB) of the signal spectra as a function of frequency.

3.2.3 Mixed Modulation (MM) Tones (SAM+FM)

The MM tones used in this study were also one second in duration. The SAM and FM components of the MM signal used the same modulation frequency. The starting phase of the amplitude modulation was adjusted so that the SAM and FM components reached maximum amplitude and maximum frequency excursion at the same point in time (Dimitrijevic, et al., 2002).
Figure 3.1. Several cycles of the four SAM tones (100% modulation) used in the current study. The top left panel shows a 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
Figure 3.2. Power spectra of the four SAM tones (100% modulation) used in the current study. The top left panel shows the 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
The MM signals used in the current study can be described with the following formula

\[ x(t) = [1 + m \cdot \sin(2\pi F_{am}t + \pi \cdot (\varphi/180))] \cdot \sin(2\pi F_c t + \beta \cdot \sin(2\pi F_{fm}t)) \]  \hspace{1cm} (3.3) 

where \( x(t) \) is the instantaneous amplitude of the signal, \( m \) is the amplitude modulation index, \( F_{am} \) is the frequency of amplitude modulation, \( t \) is time, \( F_c \) is the carrier frequency, \( F_{fm} \) is the modulation frequency of the FM, \( \varphi \) is the phase lead of the AM signal relative to the FM signal and \( \beta \) is the modulation index of the FM. The MM signals in the current study all had frequency excursions of 20% of the carrier frequency resulting in the following \( \beta \) values (in parentheses) for the carrier frequencies listed: 500 Hz (1.27), 1000 Hz (2.30), 2000 Hz (4.21) and 4000 Hz (7.77). Figures 3.3 and 3.4 depict several cycles of the MM time waveform and the power spectra respectively. Each figure consists of four panels, one for each of the carrier frequencies. The time waveforms show signal amplitude as a function of time (ms) and the power spectra show the magnitude (dB) of the signal spectra as a function of frequency.

3.2.4 Toneburst (TB) Stimuli

The TB stimuli used in this study consisted of a 1 second train of repeated brief TBs. The TBs were each approximately 4 msec in duration (8 msec for the 500 Hz carrier frequency) and were individually Blackman windowed. The quiet interval between each TB (~5-8 msec) varied according to the repetition rate of the TB train due to the fixation of the stimulus window at 1 second. Figures 3.5 and 3.6 depict several cycles of the TB time waveform and the power spectra respectively. Each figure consists of four panels, one for each of the carrier frequencies. The time waveforms show signal amplitude as a
Figure 3.3. Several cycles of the four MM tones (100% amplitude modulation, 20% frequency modulation) used in the current study. Modulation frequency was the same for both AM and FM components. The top left panel shows a 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
Figure 3.4. Power spectra of the four MM tones (100% amplitude modulation and 20% frequency modulation) used in the current study. The same modulation rates were used for both the AM and FM components. The top left panel shows a 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
Figure 3.5. Several cycles of the four TBs used in the current study. The top left panel shows a 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
Figure 3.6. Power spectra of the four TB stimuli used in the current study. The top left panel shows the 500 Hz carrier modulated at 79 Hz. The top right panel shows the 1000 Hz carrier which was modulated at 86 Hz. In the lower left panel the 2000 Hz carrier modulated at 95 Hz is shown. The 4000 Hz carrier frequency modulated at 103 Hz is shown in the lower right panel.
function of time (ms) and the power spectra show the magnitude (dB) of the signal spectra as a function of frequency.

3.3 Procedures

All testing was completed in audiometric test booths (single walled) in one of three faculty labs within the Department of Speech and Hearing Science at The Ohio State University.

3.3.1 Psychoacoustic Threshold Procedures

Subject’s behavioral thresholds were evaluated using a two alternative forced choice (2AFC) paradigm. The task required the subjects to identify in which of two intervals they thought they heard the stimulus (in all trials one interval contained the stimulus and the other was silent). Each threshold search consisted of 50 stimulus presentations. Threshold for each search was the average of the levels at which reversals occurred, excluding the first three reversals. For each stimulus at least two threshold searches were completed and the mean of the thresholds for all searches was considered the participant’s final threshold. If the first two threshold search values differed by more than 2 dB a third threshold search was conducted and the mean of all three thresholds was considered the subject’s final threshold.

Experimental Set-Up

Subjects were seated in a single walled sound booth in front of a notebook computer. The psychoacoustic threshold test was controlled by a second computer located outside the booth using custom designed software written in MATLAB.
All test signals were digitally generated, in MATLAB, by the controlling computer and passed through: (1) a Tucker-Davis Technologies (TDT) System 2 digital-analog converter (DAC), (2) a smoothing filter (8 kHz low-pass), (3) a Tucker Davis System 2 programmable attenuator (PA4), and (4) a Crown amplifier (D-75) before being presented to the listener via a pair of Sennheiser (HP 850) headphones (Figure 3.1). The output from the TDT rack was calibrated at the headphones prior to any testing. Calibration was checked at the approximate midpoint and at the conclusion of data collection to ensure that the calibration remained stable throughout the study. Stimulus values obtained during the midpoint and final checks differed from the initial values by less than 1 dB.

**Feedback Provided to Subjects**

A notebook computer in the booth provided subjects with feedback regarding their responses and alerted them to when they needed to respond. This alerting and feedback was provided by illuminating one of a number of boxes on the notebook computer screen. Labeled boxes on the notebook screen were illuminated at appropriate times throughout the task to alert the listener to: (1) system readiness, (2) initiation of a threshold search, (3) an interval presentation, (4) enter a response, (5) acceptance of the response, and (6) which interval contained the stimulus (illuminated correct answer) following each trial.

**Threshold Search Procedure**

Testing started with signal presentation levels of 45-50 dB SPL for all signals as measured on a sound level meter prior to the initiation of testing. The signal was reduced in 8 dB steps until the first reversal. Following the first reversal a step size of 4 dB was utilized until two more reversals occurred (third overall reversal) after which the step size
was reduced to 2 dB where it remained for all remaining trials in that threshold search. A procedure known as a transformed 3 up, 1 down was used in the psychoacoustic measures. In this procedure three correct responses were required for the signal level to be reduced (attenuation increased) while a single incorrect response resulted in the signal level being increased (attenuation decreased), the 3 up 1 down procedure will track to the 79% point on the psychometric function (Levitt, 1971). Subject responses were passed to the controlling computer by the notebook computer for averaging. In each threshold search the subject’s threshold was determined by averaging the attenuation level at which reversals occurred (excluding the initial 3 reversals). Final threshold was the average of 2 threshold searches unless these two searches differed by more than 2 dB in which case a third threshold search was performed and all 3 threshold searches were averaged to determine final threshold.

Subject Instruction

During the psychoacoustic measures of threshold for pure tones and ASSR stimuli, subjects were told that they would hear two intervals (each paired with a flashing light); one of the intervals would be silent while the other would contain a stimulus. Participants were instructed to press a mouse button (left button for interval 1, right button for interval 2) corresponding to the interval in which they thought they heard the stimuli. If unsure of which interval contained the stimuli the subjects were instructed to guess. At least one practice threshold search was completed to confirm that the subject was comfortable with the task and was performing at an acceptable level.
3.3.2 ASSR Threshold Procedures

The second part of this study determined the subject’s ASSR thresholds for: (1) SAM tones, (2) MM tones and (3) TB stimuli. The same carrier frequencies and modulation frequencies used to measure the behavioral thresholds in the psychoacoustic portion of the experiment were used in the ASSR portion of the experiment. Thresholds were measured using multiple simultaneously presented stimuli for all stimulus types. In the case of TB stimuli, thresholds were also measured using individually presented stimuli. The multiple simultaneous stimuli presentations contained four carrier frequencies with the same type of modulation. ASSR was initially measured with a presentation level of 70 dB SPL, confirmed at the headphone with a sound level meter, for all stimulus presentation conditions. Presentation levels were initially reduced by 8 dB whenever a response was detected at a given level. Once a presentation level was reached where there were no responses the signal was increased by 4 dB until a response again was detected. The stimulus level was then lowered or raised in 2 dB steps until threshold was determined. In the case of the individually presented tonebursts following the initial recording at 70 dB SPL the presentation level was immediately reduced to the level determined to be threshold for that carrier frequency in the multiple simultaneous stimulus presentation condition. The search procedure from this point until threshold was determined was identical to that described for the multiple simultaneous stimulus conditions. Threshold for ASSR measures was defined as the lowest level at which a response was detected ($\alpha = 0.05$) provided there was also a response detected one presentation level (2 dB) higher.
Presence or absence of a response was determined by the equipment manufacturer's detection algorithm. According to this algorithm the IHS system only considered a response to be present when the following stipulations were met: (1) absolute response amplitude of >12.5 nV, (2) signal to noise ratio of ≥ 6.13 dB in the response bin and side bins, and (3) noise level of < 50 nV. Averaging was terminated after 180 accepted (quiet) sweeps for all recordings; if no response was present after 180 sweeps the response was deemed to be absent.

**Experimental Set-Up**

The IHS SmartEP system and the computer with which it was interfaced were located outside the sound booth (Figure 3.2). The pre-amplifier was inside the sound booth and connected via fiber-optic cables passed through a patch panel to the SmartEP system. ASSR stimuli were presented via EARTONE 3A insert earphones which directly interfaced with the SmartEP equipment outside the booth.

Prior to the collection of any data the ASSR system was acoustically calibrated per the equipment manufacturer’s instructions. Following calibration procedures the output from the earphone was confirmed to be within 1 dB of the level selected in the IHS software using the test mode rather than the calibration mode to ensure that adjustments made in the calibration utility had been applied to the stimulus file.

**Subject Set-Up**

Prior to placement of the electrodes each subject’s skin was cleansed with an alcohol pad and EEG prep gel (NuPrep). Once the subject’s skin was prepared disposable pre-gelled electrodes were attached to the scalp. A non-inverting (positive) electrode was placed at Fpz (approximately the hairline), an inverting (negative)
electrode was placed on the promontory of the mastoid behind the test ear and a ground electrode was placed on the promontory of the mastoid behind the non-test ear. Once the electrode leads were connected and appropriate impedance values (1-3 kΩ) were confirmed insert earphones (EARtone 3A) were placed in both ears. The insert earphone was placed in the non-test ear to provide additional sound attenuation. 

Subject Instruction

For the ASSR portion of the study subjects were informed that their skin would be cleansed at the electrode sites, three surface electrodes would be placed on their head and insert style earphones would be placed in both ears. Subjects were informed that they would hear numerous sounds of varying loudness in one ear, and that they did not need to respond to any of the sounds. Subjects were instructed to recline quietly with their eyes closed during testing and were encouraged to sleep.

3.4 Computer Modeling of Cochlear Excitation

Computer generated models were created to simulate the activation that the current study stimuli would create on the basilar membrane. Demonstrating a difference in cochlear excitation for various stimuli supports the hypothesis that differences should be expected in the response patterns elicited by the stimuli. The program used to create this model was Moore and Glasberg's ROEX (2005) program. This program was freely available as a software download from a website maintained by the authors at the time this study was completed.

---

1 This program was freely available from a website administered by the authors. At the time of this study the excit2005 program was available at http://hearing.psychol.cam.ac.uk/Demos/demos.html.
**ROEX Model**

The ROEX model predicts basilar membrane excitation generated by a sound taking into account the filtering effects of the outer/middle ear and the non-linearity of the cochlea. The excitation pattern is calculated based on the sound spectrum that is actually reaching the cochlea. The excitation pattern is plotted as a function of the outputs from the auditory filter center frequencies. Excitation is plotted in dB with the zero reference being the excitation generated by a 1000 Hz tone presented at 0 dB SPL. ROEX models were created for all stimuli at a relatively high (60 dB) and a relatively low (30 dB) presentation level. These levels were selected as they represent presentation levels where a normal hearing subject would be expected to have a response (60 dB) and at a level estimated to be near ASSR threshold for a normal hearing subject (30 dB).
4.1 Behavioral Threshold Data

Mean behavioral thresholds and standard deviations for the stimuli (SAM, TB, MM and PT) are summarized by frequency in Figure 4.1. Behavioral thresholds were similar, within 5 dB of each other, at corresponding carrier frequencies across stimuli with the exception of the threshold for the 500 Hz MM stimulus which was about 8 dB lower. In order to determine if significant differences existed between the behavioral ASSR stimuli thresholds and the corresponding PT threshold, behavioral threshold data for the ASSR stimuli (SAM, TB, MM) at each carrier frequency (500, 1000, 2000, and 4000 Hz) were compared to the corresponding PT using a repeated measures analysis of variance (ANOVA). Results of the repeated measures ANOVA revealed significant main effects for the 500 Hz, 1000 Hz, and 4000 Hz carrier frequencies (500 Hz: $F_{3,27} = 29.77$ p < .01; 1000 Hz: $F_{3,27} = 8.783$; p < .01; 4000 Hz: $F_{1.59,14.3} = 4.032$; p < 0.05). Mauchly’s test indicated that the assumption of sphericity was violated (chi-square = 23.54, p<0.05) at the 4000 Hz carrier frequency, degrees of freedom were adjusted with a Greenhouse-Geisser estimate (epsilon = 0.53). Post-hoc analyses (paired samples comparison with Bonferonni correction) revealed the only significant difference of interest (between an ASSR stimulus and the PT stimulus corresponding to the carrier frequency) in behavioral
Figure 4.1. Mean ASSR and behavioral thresholds for all subjects (n=10) for SAM, MM, TB and PT stimuli. Behavioral thresholds are the mean of 2 threshold searches. ASSR thresholds represent the lowest level at which a response could be recorded provided there was also a response at the next highest level (+2 dB).
thresholds for the 500 Hz carrier frequency to be between the MM stimulus and the PT stimulus. Specifically, the mean MM threshold was significantly lower than the mean PT threshold at 500 Hz. All other comparisons (SAM and TB vs. PT) were not significant, indicating similar thresholds across stimuli types at 500 Hz. Post-hoc testing at 1000 Hz revealed a significant difference between the thresholds for the TB and PT stimuli. Specifically, the mean PT threshold was significantly lower than the mean TB threshold. All other comparisons (SAM and MM vs. PT) were not significant, again indicating similar thresholds across stimuli types at 1000 Hz. No significant differences between ASSR stimuli thresholds and the PT threshold were found in post-hoc testing (paired samples comparison with Bonferonni correction) for the 4000 Hz carrier frequency.

4.2 ASSR Threshold Data

Mean ASSR thresholds and standard deviations for all stimuli (SAM, MM, TBM, TBS) are summarized by frequency in Figure 4.1. Differences in ASSR thresholds were analyzed using a linear mixed effects model with repeated measures to determine if there were significant differences in threshold as a function of stimulus type. The mixed effects model was selected over a general linear model because in the mixed effects model independent observations are not assumed as in the general linear model. The individual observations in the current study may not be entirely independent due to the subject selection process resulting in a shared group history making the mixed effects model more appropriate. Secondly, the mixed effects model allows one to select from a wide variety of covariance matrix assumptions to find the best fit for the data. In a mixed effects model the fixed factor is generally the primary variable of interest, which was
stimulus type in the current study. A factor identified as random in the mixed effects model consists of a categorical variable where only a random sample of all possible subjects have been tested (e.g., the 10 individuals selected as participants for the current study out of a population of ~100). The decision to treat subjects as a random factor was based on the ASSR thresholds being highly variable between subjects while behavioral thresholds were much more consistent across subjects. Had ASSR threshold patterns been more consistent across subjects then it would not have been necessary to treat the subjects as a random factor. The analyses of ASSR thresholds for multiple simultaneous stimuli conditions revealed a significant main effect for stimulus at the 2000 Hz and 4000 Hz carrier frequencies but not at the 500 Hz or 1000 Hz carrier frequencies for the multiple simultaneous stimuli conditions. Post-hoc analyses (paired samples comparisons) revealed the following significant differences between ASSR stimuli by carrier frequency: 2000 Hz (SAM/TBM); 4000 Hz (SAM/TBM). The TBM stimulus, as expected, resulted in lower thresholds relative to the SAM stimulus at both 2000 Hz and 4000 Hz.

The current study also sought to determine the impact using multiple simultaneous stimuli has on a TB ASSR threshold. In order to determine this effect the thresholds generated by TB stimuli in both the multiple simultaneous and individual presentation methods were compared. When the TBS and TBM thresholds were compared, using a linear mixed effects model, significant differences between the thresholds for the two stimulus presentation methods were found at all carrier frequencies (500 Hz: $F_{1,9} = 5.56, p < 0.05$; 1000 Hz: $F_{1,9} = 10.92, p < 0.05$; 2000 Hz: $F_{1,9} = 7.89, p < 0.05$ and 4000 Hz: $F_{1,9} = 15.14, p < 0.05$). In all cases an individually presented TB
Figure 4.2. Mean ASSR thresholds and standard deviations for all subjects (n=10) using the multiple simultaneously presented and individually presented TB stimuli.
yielded a lower threshold than the simultaneously presented TBs, the differences between thresholds ranged from approximately 7 to 11 dB as shown in Figure 4.2. The estimates of covariance parameters provided by the mixed effects model indicated the random effect of subjects was not significant in the multiple simultaneous stimuli conditions. The insignificant findings indicate that the random sample of subjects did not impact the relationship between the fixed factor of stimulus and threshold meaning that any observed differences were due to stimulus not subject effects. For the individually presented TB stimuli versus the multiple simultaneous TB stimuli the estimate of covariance parameters was only significant for the 1000 Hz carrier frequency (Wald Z = 1.97; p < 0.05) indicating that the subject effect was significant and that the random sample of subjects, not just the stimulus, had an influence on the dependant variable (threshold). The implication of a significant finding here is that the results may have been skewed by the subject sample and may not accurately reflect the overall population from which the subjects were selected. The considerable variability in ASSR thresholds and relative stability of behavioral thresholds across subjects, by stimulus type, can be seen in the individual data plots (Appendix C).

4.3 Threshold Estimates

ASSR thresholds must be corrected or adjusted to provide behavioral threshold predictions. Previous studies (e.g. Cone-Wesson et al., 2002c; Rance et al., 1995; Rance et al., 1995; Aoyagi et al., 1999; John et al., 2001; Dimitrijevic et al., 2002) have used either a linear regression model or subtracted a mean difference correction factor (mean difference between ASSR thresholds and measured behavioral thresholds) to make
threshold predictions. The current study utilized both of those measures and adds a third type of regression analysis (quadratic). The quadratic regression model was used due to data in some conditions following a more quadratic rather than linear pattern. The linear and quadratic regression formulae used to make threshold predictions are shown in Tables 4.1 and 4.2. Mean difference values are shown in Table 4.3 and represent the average difference between ASSR thresholds and behavioral thresholds for a pure tone whose frequency corresponds to the carrier frequency of the modulated tone.

The range of values for measured and predicted thresholds utilizing the mean difference correction values, linear regression formulae and quadratic regression formulae in Tables 4.1 through 4.3 are displayed in Figures 4.3 through 4.6. Each figure shows the predicted and measured threshold ranges at one carrier frequency for each stimulus type. The mean and median predictions were relatively uniform across mean difference prediction and regression based prediction but the variability of the estimates differed considerably. The use of a regression model rather than subtracting a fixed value resulted in a less variable prediction; however, predictions made by the linear and quadratic regression models had similar variability in most conditions (See Figures 4.3 through 4.6).

The differences between the predicted behavioral thresholds and the measured behavioral thresholds were analyzed with a series of repeated measures ANOVAs. These analyses revealed significant main effects for 1000 Hz (MM stimulus: F2,18= 8.325, p < 0.03; TBS stimulus: F2,18= 12.184, p < 0.01), 2000 Hz (TBS stimulus: F1,1,18= 8.239, p <0.05) and 4000 Hz (MM stimulus: F1,1,18= 5.995, p < 0.035, TBM stimulus: F1,1,18=...
Table 4.1. Linear regression formulae used for predicting behavioral thresholds for a pure tone using ASSR thresholds obtained in the current study. The ASSR threshold replaces (x) and the resulting (y) is the predicted behavioral threshold. Formulae are displayed by stimulus type and frequency.

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>500 Hz</th>
<th>1000Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>y = 0.171x + 9.35</td>
<td>y = 0.251x + 1.63</td>
<td>y = 0.576x -10.34</td>
<td>y = 0.754x -33.18</td>
</tr>
<tr>
<td>MM</td>
<td>y = 0.303x + 1.94</td>
<td>y = 0.344x + 0.01</td>
<td>y = 0.461x -2.99</td>
<td>y = 0.664x -28.27</td>
</tr>
<tr>
<td>TBM</td>
<td>y = -0.1x + 25.87</td>
<td>y = 0.21x - 4.28</td>
<td>y = 0.601x - 6.54</td>
<td>y = 0.41x -9.26</td>
</tr>
<tr>
<td>TBS</td>
<td>y = 0.154x + 11.28</td>
<td>y = 0.446x -2.88</td>
<td>y = 0.511x + 2.11</td>
<td>y = 0.436x -5.71</td>
</tr>
</tbody>
</table>

Table 4.2. Quadratic regression formulae used for predicting pure tone behavioral thresholds in the current study. The ASSR threshold replaces (x) and the resulting (y) is the predicted behavioral threshold. Formulae are displayed by stimulus type and frequency.

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>500 Hz</th>
<th>1000Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>y = 0.005x² + (-0.345x) + 21.48</td>
<td>y = 0.004x² + (-0.139x) + 10.53</td>
<td>y = -0.002x² + (0.787x) - 15.42</td>
<td>y = 0.039x² + (-3.921x) + 100.82</td>
</tr>
<tr>
<td>MM</td>
<td>y = -0.021x² + (2.705x) -66.35</td>
<td>y = -0.076x² + (6.507x) -120.48</td>
<td>y = -0.01x² + (1.469x) -26.16</td>
<td>y = 0.022x² + (-1.813x) + 39.24</td>
</tr>
<tr>
<td>TBM</td>
<td>y = -0.024x² + (2.753x) -58.23</td>
<td>y = -0.023x² + (2.5x) -49.86</td>
<td>y = -0.003 x² + (0.889x) - 12.65</td>
<td>y = 0.011x² = (-0.675x) + 15.72</td>
</tr>
<tr>
<td>TBS</td>
<td>y = 0.035x² + (-3.727x) + 114.76</td>
<td>y = -0.01x² + (1.179x) -15.52</td>
<td>y = -0.03 x² + (2.611x) -29.91</td>
<td>y = -0.004 x² + (0.827x) – 13.17</td>
</tr>
<tr>
<td>Frequency (Hz)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>1000</td>
<td>2000</td>
<td>4000</td>
</tr>
<tr>
<td>Difference in Thresholds (ASSR vs. PT behavioral) dB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAM – PT</td>
<td>40.8</td>
<td>35.4</td>
<td>31.3</td>
<td>49.6</td>
</tr>
<tr>
<td>MM – PT</td>
<td>38.8</td>
<td>26.6</td>
<td>27.7</td>
<td>47.3</td>
</tr>
<tr>
<td>TB(M) – PT</td>
<td>40.4</td>
<td>29.2</td>
<td>22.9</td>
<td>39.4</td>
</tr>
<tr>
<td>TB(S) – PT</td>
<td>34.6</td>
<td>23.8</td>
<td>17.1</td>
<td>28.0</td>
</tr>
</tbody>
</table>

Table 4.3. Mean difference correction values used to predict pure tone behavioral thresholds for all stimulus types: SAM, MM, TB(M) and TB(S) in the current study.
Figure 4.3. Range of predicted and measured thresholds for all subjects (n=10) using a 500 Hz stimuli (MD= mean difference, LR = linear regression, QD = quadratic regression, MEAS = measured behavioral threshold). Each stimulus type had it’s own regression formula or mean correction value. Error bars represent the 5th and 95th percentiles; the box edges represent the 25th and 75th percentiles. The thin cross bar is the median threshold level and the thick cross bar represents the mean.
Figure 4.4. Range of predicted and measured thresholds for all subjects (n=10) using a 1000 Hz stimuli (MD = mean difference, LR = linear regression, QR = quadratic regression, MEAS = measured behavioral threshold). Each stimulus type had its own regression formula or mean correction value. Error bars represent the 5th and 95th percentiles; the box edges represent the 25th and 75th percentiles. The thin cross bar is the median threshold level and the thick cross bar represents the mean.
Figure 4.5. Range of predicted and measured thresholds for all subjects (n=10) using a 2000 Hz stimuli (MD = mean difference, LR = linear regression, QD = quadratic regression, MEAS = measured behavioral threshold). Each stimulus type had its own regression formula or mean correction value. Error bars represent the 5th and 95th percentiles; the box edges represent the 25th and 75th percentiles. The thin cross bar is the median threshold level and the thick cross bar represents the mean.
Figure 4.6. Range of predicted and measured thresholds for all subjects (n=10) using a 4000 Hz stimuli (MD = mean difference, LR = linear regression, QD = quadratic regression, MEAS = measured behavioral threshold). Each stimulus type had its own regression formula or mean correction value. Error bars represent the 5th and 95th percentiles; the box edges represent the 25th and 75th percentiles. The thin cross bar is the median threshold level and the thick cross bar represents the mean.
8.731. p < 0.013) carrier frequencies. Mauchley’s test indicated the assumption of
sphericity was violated at 2000 Hz for the TBS stimulus (chi-square = 12.9, p < 0.05),
4000 Hz for the MM stimulus (chi-square = 16.8, p < 0.05) and 4000 Hz TBM stimulus
(chi-square = 12.9, p < 0.05). Degrees of freedom were adjusted using Greenhouse-
Geisser estimates where indicated (2000 Hz TBS: epsilon = 0.53; 4000 Hz MM: epsilon
= 0.55; 4000 Hz TBS: epsilon = 0.55). Post-hoc (paired comparisons with Bonferroni correction) indicated significant differences between the mean difference correction method and both forms of regression at 1000 Hz for the MM and TBS stimuli. For both types of stimuli (MM and TBS) the regression models provided closer threshold predictions than did the mean difference correction method. Further, the linear regression provided the closer estimate of the two regression predictions. A significant difference was also found between predicted thresholds generated with the mean difference correction and both regression models (p < 0.05 with Bonferroni adjustment) at 2000 Hz for the TBS stimulus. In this case the mean based prediction was the most accurate followed by the quadratic and linear regression based predictions respectively. Post-hoc analysis (paired comparisons with Bonferroni correction) also found a significant difference between the linear and quadratic regression predictions for a 4000 Hz TBM (linear regression provided the most accurate prediction followed by the mean difference and finally quadratic regression). Post-hoc analysis (paired comparisons with Bonferroni correction) for the MM stimulus at 4000 Hz showed significant differences between the mean difference and the two regression based predictions with the linear regression providing the most accurate prediction followed by the mean difference and quadratic regression respectively. The lack of significant differences for prediction error in most
stimulus conditions (11 of 16 conditions showed) indicates that overall the three
correction methods tended to exhibit similar prediction errors. The mean prediction
errors and the standard deviations of the prediction error are shown in Table 4.4. These
values were obtained by subtracting the predicted behavioral threshold for each subject
from their measured behavioral threshold, for example if the predicted threshold was 16.5
dB and the measured behavioral threshold was 14.0 dB then the prediction error is 2.5 dB.
The prediction error values obtained for each subject were then averaged for each
condition and the standard deviation for each condition was calculated.

In order to test the robustness of the predictive models, a series of jack-knife
analyses were completed to determine the similarity of predictions for subjects when their
data was included and excluded from the model used in making the prediction. A jack-
knife analysis is a form of bootstrapping which re-samples subsets of the data to
determine if there are any biases caused by outliers within the dataset. To complete this
jack-knife analysis one subject was removed from the data pool and the regression
models were recalculated then used to predict the thresholds for that subject. This
procedure was repeated for each individual subject resulting in ten new regression
formulae for each condition. The smaller the difference between the original and jack-
knife predictions, the more robust the model indicating that outliers in the dataset did not
substantially skew the model. Differences between the threshold predictions based on the
original regression formula and the jack-knife regression formula ranged between
approximately 1 to 11 dB, with most differences being between 2 and 4 dB. Table 4.5
shows the mean differences between the jack-knife procedures threshold estimates and
the threshold predictions based on the entire dataset for each stimulus by frequency. The
<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>6.7(9.4)</td>
<td>8.6(6.9)</td>
<td>6.2(3.6)</td>
<td>5.2 (2.9)</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>4.0 (4.2)</td>
<td>5.3 (3.2)</td>
<td>5.6 (2.9)</td>
<td>4.5 (3.0)</td>
</tr>
<tr>
<td>Quadratic Regression</td>
<td>4.6 (4.0)</td>
<td>5.9 (3.4)</td>
<td>6.9 (5.9)</td>
<td>9.6 (4.0)</td>
</tr>
<tr>
<td>MM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>6.1 (4.5)</td>
<td>6.2 (5.0)</td>
<td>6.9 (4.9)</td>
<td>5.3 (3.3)</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>4.3 (3.6)</td>
<td>5.4 (3.4)</td>
<td>4.5 (4.1)</td>
<td>4.1 (3.1)</td>
</tr>
<tr>
<td>Quadratic Regression</td>
<td>4.7 (4.0)</td>
<td>5.9 (3.5)</td>
<td>8.0 (3.5)</td>
<td>7.6 (4.3)</td>
</tr>
<tr>
<td>TBM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>10.5 (5.3)</td>
<td>8.1 (6.4)</td>
<td>5.2 (2.9)</td>
<td>9.2 (6.7)</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>4.5 (4.0)</td>
<td>5.5 (2.2)</td>
<td>2.9 (2.8)</td>
<td>4.6 (3.1)</td>
</tr>
<tr>
<td>Quadratic Regression</td>
<td>4.7 (4.0)</td>
<td>7.1 (4.1)</td>
<td>6.7 (4.0)</td>
<td>10.1 (4.7)</td>
</tr>
<tr>
<td>TBS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>10.7 (4.8)</td>
<td>7.5 (3.4)</td>
<td>6.3 (5.2)</td>
<td>10.2 (10.1)</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>4.6 (3.5)</td>
<td>7.2 (5.4)</td>
<td>9.8 (6.9)</td>
<td>8.0 (7.1)</td>
</tr>
<tr>
<td>Quadratic Regression</td>
<td>4.7 (4.0)</td>
<td>7.3 (4.3)</td>
<td>6.8 (3.9)</td>
<td>10.2 (5.0)</td>
</tr>
</tbody>
</table>

Table 4.4. Mean prediction errors (in dB) between predicted and measured thresholds by stimulus type and prediction method. Values in parentheses represent the standard deviation of the prediction error.
<table>
<thead>
<tr>
<th>Stimulus</th>
<th>500 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>2.9</td>
<td>2.8</td>
<td>2.4</td>
<td>1.5</td>
</tr>
<tr>
<td>MM</td>
<td>7.7</td>
<td>1.9</td>
<td>3.1</td>
<td>3.9</td>
</tr>
<tr>
<td>TBM</td>
<td>2.7</td>
<td>5.5</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>TBS</td>
<td>2.3</td>
<td>4.4</td>
<td>2.0</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Table 4.5. Mean differences (in dB) between the jack-knife linear regression predictions for each subject (n=10) and the original linear regression prediction models for the same subject.
Fig 4.7. Cross validation comparisons for a representative subject. The top left panel shows predictions for SAM stimuli, the top right MM stimuli, bottom left TBM stimuli, and bottom right TBS. Predictions were made excluding the subject indicated below the predicted threshold on the x-axis.
values in Table 4.5 were obtained by finding the mean difference between subject’s jack-knife and full dataset based predicted thresholds.

A cross validation comparison using the regression models constructed in the jack-knife procedure was completed for all subjects. The predictions for a representative subject can be seen in Figure 4.7. For any given condition the thresholds from each of the prediction models were within a range of 2 to 4 dB different and the largest difference in thresholds across any one condition being about 8 dB. As in the jack-knife procedure, results of the threshold prediction were similar across models for all subjects indicating that the threshold prediction model is robust and was not substantially influenced by outlying subjects.

4.4 Results of Computer Modeling of Cochlear Excitation

**ROEX Model**

Models of cochlear (basilar membrane) excitation were constructed to demonstrate that the stimuli used in the current study resulted in different activation of the cochlea and could therefore be expected to generate different response patterns as the neural activity represented by the ASSR is reflective of the excitation generated on the basilar membrane by a stimulus. Based on the ROEX models the most robust pattern of basilar membrane excitation was generated by the TB and MM tones and the least robust excitation pattern was generated by the SAM tone. The differences in excitation pattern were in the width of the excited area as the amount of excitation (in dB) at the carrier frequency were equal across stimuli. However, there was not much difference in the excitation patterns for the TB and MM stimuli in the ROEX model. This supports the
notion that using a MM or TB stimulus will allow you to record the ASSR to lower thresholds than a SAM stimulus but does not provide support for a TB stimulus being able to generate a lower threshold than an MM stimulus. Greater excitation should lead to a stronger neural response that would presumably be recordable to a lower stimulus level than a response which started out with a less robust excitation pattern in the cochlea. At 500 and 1000 Hz the TB resulted in the most excitation while at 2000 and 4000 Hz the MM tone resulted in the most excitation in the cochlea. Since the MM tones and TBs are exciting larger areas of the basilar membrane than the SAM stimuli it was anticipated that wider acoustic spectrum TB and MM tone would result in lower response thresholds due to the larger areas of excitation along the basilar membrane. Based on the outputs of the ROEX models one would expect to see substantial differences between the SAM stimuli response and the MM and TB responses. The differences in excitation generated by the MM and TB stimuli are much more subtle so these two stimuli would likely not generate substantially different responses based on the model.

TB stimuli were presented using both a multiple simultaneous stimuli presentation paradigm and an individual stimulus presentation paradigm in an effort to determine the influence presenting multiple simultaneous TBs has on the response threshold relative to a threshold for a TB stimulus presented individually. There was considerable overlap between areas of excitation for component stimuli at the high presentation level using the multiple component stimuli but at the lower level the excitation area overlap was reduced considerably. Overlap of excitation areas only occurred between 500 and 1000 Hz TB stimuli and the 1000 and 2000 Hz TB stimuli.
The measured data followed the predictions generated from the ROEX model in that the thresholds generated with multiple simultaneous MM and TB stimuli were lower than those elicited with a multiple simultaneous SAM tones. However, the MM threshold at 2000 Hz and 4000 Hz was closer to the SAM threshold than to the TB threshold (TB threshold lower) contrary to what was expected based on the model. The outputs of the ROEX models are provided in Appendix B.
CHAPTER 5

DISCUSSION AND CONCLUSIONS

5.1 Rationale for the Current Study

ASSR is a relative newcomer to AEP based prediction of hearing sensitivity and as such there are still a number of details about the test that have not been fully established. A number of different stimuli and analysis methods have been used but none has emerged as a standard way of conducting the test. This variability in methodology makes comparison across studies problematic as there may be several parameters that differed making it hard to determine which one was responsible for any observed differences. The current study standardized the test protocol, with the exception of stimulus type, so that any differences in thresholds would be attributable to the stimulus and not the recording or analysis method. Several methods to arrive at a predicted threshold are also examined but these analyses were carried out separately from the stimulus effects analyses. A more recent stimulus (TB) has also been incorporated into ASSR testing and may prove a superior signal if in fact, as it has been suggested but Stapells et al. (2004), the ASSR is really overlapping wave V responses of the ABR. In their 2004 paper Stapells and colleagues do not provide any evidence to support their assertion and in fact there is some evidence in a lesion study (in an animal model) that shows ABR and ASSR have separate generators (Kiren et al., 1994).
5.2 Behavioral Thresholds

In addition to the electrophysiological thresholds behavioral thresholds were obtained for all ASSR stimuli utilized in the current study (SAM, MM, TB) along with behavioral thresholds for pure tones corresponding to the carrier frequencies of the ASSR stimuli. While previous ASSR studies have utilized step sizes of 5 or 10 dB (e.g., Lins & Picton, 1995; Dimitrijevic, et al., 1991; Rance, et al., 1995; Rance, et al., 1999; John, et al., 2001), a 2 dB step size was chosen for the current study so that differences between behavioral and ASSR thresholds for corresponding stimuli could be measured with greater precision. These comparisons were made to demonstrate that in behavioral testing the ASSR stimuli have similar audibility to a pure tone which corresponds to the carrier frequency of a modulated ASSR signal.

The differences between the behavioral thresholds for ASSR stimuli with carrier frequencies of 2000 Hz and 4000 Hz were not significantly different from a pure tone of 2000 Hz or 4000 Hz respectively. For the 1000 Hz carrier frequency the difference in mean behavioral threshold between the TB stimulus and PT stimulus was significant but still relatively small (<5 dB) and would either not have been detected or been within 1 step using a standard clinical step size (5 dB). So, while different statistically the thresholds would not be considered clinically different since they would have been equal or within one step of each other. The only other carrier frequency showing a significant difference in behavioral threshold was 500 Hz and only for the MM threshold relative to the PT threshold. In the case of the 500 Hz MM threshold the large difference observed between it and the other 500 Hz thresholds was potentially due to an artifact which cued listeners to the stimulus interval and likely does not reflect a true difference. Based on
the similarity of and lack of clinically significant difference in the behavioral thresholds for pure tones and ASSR stimuli with corresponding carrier frequencies, predicting pure tone thresholds based on ASSR thresholds for TB, MM and SAM tones is an appropriate comparison despite the use of different signals. Additional support for the use of modulated signals in making puretone threshold estimates is provided by studies showing similar cochlear place specificity for some ASSR stimuli (SAM tones) and pure tone stimuli (Herdman, Picton, & Stapells, 2002; Herdman & Stapells, 2003).

Behavioral thresholds were collected with an adaptive threshold procedure (Levitt, 1971). Specifically, a 3 up 1 down, 2AFC procedure was utilized, which targets the 79% point on the psychometric function. The 2AFC procedure was selected to try and reduce the potential bias that can be introduced by the listener in a simple up/down procedure. In the simple up/down procedure the listener may determine that a sequential pattern is being used and could anticipate a stimulus and respond without actually detecting the stimulus (Levitt, 1971). A 3 up 1 down procedure was selected in an effort to obtain highly precise behavioral thresholds. The threshold search procedure in the ASSR measures is equivalent to the simple 1 up 1 down procedure (Wetherill & Levitt, 1965) which targets the 50% rather than 79% point on the psychometric function. Under a best case scenario it is unlikely that the behavioral and ASSR thresholds tracking to different points on the psychometric function had a substantial effect on the threshold predictions in the current study based on the work of Watson, Frank and Hood (1972). In their study Watson et al. (1972) measured psychometric functions in a group of young normal hearing listeners (similar to the subject pool in the current study) and found a mean difference of about 3 to 4 dB in stimulus level was needed to move from chance levels of
identification to near perfect signal identification in the frequency range of stimuli used in the current study (500-4000 Hz). However in the Watson et al. (1972) study there was also considerable variability across subjects and those small differences (3-4 dB) were reached after thousands of trials, considerably more training that was given to subjects in the current study. Most of the subjects in the current study had at least some experience in other psychoacoustic experiments but had, at most, a half dozen or so training trials with the current study protocol. These factors make it difficult to determine just how much any given subject’s thresholds might have changed if a 1 up 1 down procedure had been used and is admittedly a weakness of the study.

5.3 ASSR Thresholds

A simple 1 up 1 down threshold search procedure with a single interval was selected for the ASSR portion of the study for two reasons. First, the ASSR is an involuntary physiologic response so there is no danger of a listener recognizing a presentation pattern and adjusting their responses accordingly as Levitt (1971) described for a behavioral measure. Secondly, the amount of time required to obtain an ASSR threshold is considerably longer than a behavioral threshold (30-60 minutes vs. 4-5 minutes), therefore, a 3 up 1 down procedure would not have been practical from a time standpoint. Completing 50 trials (as was done in the behavioral testing) would have taken 2 ½ hours or more for just one stimulus condition with ASSR testing. The use of a 3 up 1 down procedure would have been further complicated by the multiple simultaneous stimuli—if one component drops out does stimulus level increase or keep going down because other responses are still present.
There is currently not a standard stimulus for eliciting, detecting or analyzing the ASSR, so different manufacturers have selected different stimuli and analysis parameters for their ASSR equipment. At the time of the current study little published data was available with threshold measures comparing the most commonly utilized stimuli under otherwise identical conditions. The lack of a standard stimulus or analysis method makes comparing results difficult because changing these parameters may result in a shift in threshold without a concomitant change in hearing sensitivity.

ASSR thresholds were measured using multiple simultaneous SAM, MM and TB stimuli. While all of these stimuli have been or are in use in measuring ASSR no published papers were found where all stimuli types were measured in the same group of subjects using identical analysis and collection parameters. Testing a single group of subjects with all stimuli using the same response detection protocol allows for the fairest comparison of the response characteristics of the different stimuli. Prior work has demonstrated that a MM stimulus results in a lower threshold than a stimulus modulated with only SAM (e.g., John, et al., 1998; Dimitrijevic, et al., 2002), however, this has not been demonstrated for TB generated responses. One of the goals of the current study was to determine if the sharp onset and brief duration of the TB stimulus results in a lower threshold compared to the MM stimulus.

The ASSR threshold analyses only showed a significant threshold difference for stimuli with carrier frequencies of 2000 Hz and 4000 Hz. Post-hoc analysis of the ASSR threshold data (paired samples testing with Bonferonni correction) showed the only significant difference in threshold at 2000 Hz and 4000 Hz was between the SAM and TBM stimuli. The thresholds for different ASSR stimuli at 500 Hz while not
significantly different from one another by stimulus type tended to be higher (poorer) relative to other carrier frequencies. This is not especially surprising as noise levels tend to be higher and neural synchrony poorer at lower frequencies. Prior studies have also reported poor responses at low frequencies (e.g., John et al., 1998; Rance & Rickards, 2002). The lack of significant findings between the stimuli in many conditions is contrary to other prior studies where significant differences have been noted between SAM and MM stimuli (John et al., 1998; Dimitrijevic, et al., 2002; Picton, et al., 2003). These findings suggest that at least for the equipment used in the current study the type of stimulus used (SAM, MM, TBM) does not have a significant impact on threshold for the lower frequencies, perhaps due to the rather poor (high) thresholds that were recorded.

**Multiple vs. Single Frequency Toneburst Presentation**

In all cases responses generated by a single frequency TB presentation exhibited a lower threshold than did the multiple simultaneous frequency TBs. In the multiple simultaneous stimulus each component TB is calibrated separately so that the SPL of each component TB is the same as when it is presented as an individual TB. The differences in thresholds obtained using multiple simultaneous vs. individually presented TBs were both statistically and clinically different. These differences ranged between approximately 7 dB and 11 dB.

For individually presented TBs the predicted thresholds were within 1 to 10 dB of the corresponding measured behavioral threshold for a pure tone (mean differences varied by frequency but were within 3 to 6 dB of the measured thresholds). When ASSR was measured with multiple simultaneous TBs the predicted thresholds were within 1 to
25 dB of the measured pure tone behavioral threshold (mean differences varied by frequency but were within 7 to 10 dB of the measured behavioral thresholds).

Clearly individual presentation of the carrier frequencies provides a closer and less variable estimate of behavioral threshold but also has been shown to take 2-3 times as long as multiple simultaneous presentation (Herdman & Stapells, 2001, 2003; John et al., 2002). Multiple simultaneous frequency TB presentation will not show a four-fold decrease in time due to some carrier frequencies (generally 500 Hz) tending to need additional averaging to reach significance. While the mean difference in prediction error between simultaneous and individually presented was only about 4 dB then range of prediction errors differed by considerably more (approximately 15 dB). In light of this finding at least at the moment the time savings of using multiple simultaneous TBs is outweighed by the loss of accuracy of prediction.

**Stimulus Interaction Effects**

The current study also sought to determine if multiple simultaneous TBs have similar resistance to stimulus interaction effects, as has been shown with SAM stimuli (Herdman & Stapells, 2001; John, et al., 1998). Given the wider acoustic spectrum of the toneburst components (numerous sidebands above and below the carrier frequency compared to the single sideband on either side of the carrier frequency in the SAM stimulus) it was anticipated that stimulus interaction effects would be greater with toneburst stimuli than with the SAM stimuli. As demonstrated in the computer models, the areas of excitation within the cochlea for each component overlap considerably more in the toneburst stimuli than in the SAM stimuli. This greater overlap increases the chances of one stimulus component affecting the threshold for another. The thresholds
were lower at all carrier frequencies in the individually presented TB threshold paradigm compared the multiple simultaneous TB presentation. The differences in threshold between the multiple simultaneous and individual TB presentation conditions were between approximately 7 and 11 dB. The threshold differences at all frequencies tested were found to be significantly different using a linear mixed effects model (500 Hz: $F_{1,9}=5.55\, p<0.05$; 1000 Hz: $F_{1,9}=10.917, \, p<0.01$; 2000 Hz: $F_{1,9}=7.811, \, p<0.05$; 4000 Hz: $F_{1,9}=15.142, \, p<0.01$). Prior work has suggested that stimulus interaction does not occur when the stimulus level is below 60 dB SPL (John et al., 1998). The results of the current study, however, indicate that stimulus interaction effects are occurring at much lower presentation levels with TB stimuli and that TB stimuli are less resistant to interaction effects than SAM stimuli and should be presented individually rather than simultaneously.

5.4 Behavioral Threshold Estimates

The accuracy of the behavioral threshold estimates was determined by how close they were to measured behavioral thresholds. For the pure tone behavioral threshold/multiple simultaneous stimuli ASSR comparisons the MM and TBM signals provided the closest estimates at all frequencies except 500 Hz where the SAM stimulus resulted in the closest prediction.

Threshold Prediction Methods

Three prediction methods were compared to determine which method produced the closest and most consistent estimates of behavioral hearing thresholds for pure tones.
As mentioned previously the first of these methods involves finding a mean difference between ASSR thresholds and measured behavioral thresholds and then subtracting this mean difference from the ASSR thresholds (see Table 4.3 for mean difference values for the current study). The product of this subtraction represents the estimated threshold. This method resulted in the largest prediction error for all stimulus types at 500 Hz and 1000 Hz and for the 4000 Hz TBS stimulus as shown in Table 4.4.

The second method for threshold estimation examined in this study was to calculate and use linear regression equations (see Table 4.1 for individual formulae). Linear regression is the most commonly utilized method to predict behavioral thresholds based on ASSR thresholds. The ASSR threshold is inserted into the regression equation and solving this equation provides the estimated threshold. A separate formula is required for each carrier frequency and stimulus type and the estimates for each carrier frequency are made separately. With the exception of the 2000 Hz TBS condition the linear regression model yielded the lowest prediction error of all methods utilized in the current study (see Table 4.4).

In several conditions the data collected in the current study were observed to have a quadratic rather than linear relationship between the behavioral and ASSR thresholds and using a quadratic rather than linear model would provide a more accurate prediction. For this reason quadratic regression models were also used to try and improve the predictive ability of the model (See table 4.2 for the individual quadratic formulae). Just as in the linear regression the ASSR threshold is plugged into the regression equation to estimate the behavioral threshold and a separate formula is needed for each stimulus and carrier frequency. The linear and quadratic regression models resulted in very similar
Figure 5.1. The top panel represents a condition where ASSR thresholds (X axis) and behavioral thresholds (y axis) had a somewhat quadratic relationship. The top panel represents the 1000 Hz MM ASSR thresholds (x-axis) compared to a 1000 Hz pure tone threshold (y-axis). The bottom panel represents the more typical linear relationship between the ASSR threshold (x-axis) and pure tone behavioral threshold (y-axis). The bottom panel plots the 2000 Hz SAM threshold (x-axis) against the 2000 Hz pure tone threshold (y-axis). In both panels the open circles are observed data points, the solid line is a linear regression line and the dashed line is a quadratic regression line.
mean predictions and prediction ranges (see Table 4.4). Based on these results the accuracy of behavioral threshold predictions made based on linear regression are reasonably accurate and it can be used in this application without needing to move to a more complex model. The slight quadratic shape observed in some of the data is likely due to the small sample size in the current study as previous similar studies with much larger sample sizes have not reported a quadratic shape to their data (e.g., Rance et al., 1999; Rance & Rickards, 2002; Cone-Wesson, et al. 2002b). The range of predictions for all prediction methods and the measured thresholds can be seen graphically in Figures 4.3 through 4.6.

When using a regression formula to estimate behavioral thresholds based on ASSR thresholds it is important to know how the formula was derived. The data being examined must be collected in the same manner as the data used in calculating the regression model. Failure to do so can dramatically impact the accuracy of the prediction. To illustrate this point Table 5.1 shows estimates of behavioral threshold that used regression equations from prior studies (Cone-Wesson et al., 2002b; Rance & Rickards, 2002) with ASSR data collected in the current study. The Cone-Wesson et al. (2002b) and Rance and Rickards (2002) studies used different response detection techniques, slightly different stimuli and different step sizes compared to the current study. Additionally both the Rance and Rickards (2002) and Cone-Wesson et al. (2002b) studies utilized single frequency presentations rather than multiple simultaneous frequency presentations. Threshold estimates were made using the formulae from other studies and data collected in the current study to illustrate the importance of knowing where the estimation method applied to an ASSR threshold comes from and whether it is
appropriate to use with your measures. In most cases the estimated thresholds derived with the formulae from other studies were substantially less accurate in their estimate than the estimates using current study formulae. This illustrates the need to collect data using the same parameters as the model selected to predict the behavioral threshold.

The jack-knife and cross correlation analyses were completed to determine if the prediction for an individual subject remained stable when their threshold was predicted while they were removed from the data pool and when it was predicted using various subgroups of subjects respectively. Across subjects the predictions made from a model that did not include that particular subject’s thresholds (jack-knife) were within 5 dB of the original prediction 81% of the time with only one case where the difference in the original and jack-knife prediction was greater than 10 dB. The consistency of these analyses indicates that the model is relatively robust (at least for a normal hearing ear) and could be used to predict thresholds of persons not included in the model with similar accuracy to those included in the model. Cross validation analysis showed relatively stable predictions as well across subjects (Figure 4.7) using each of the regression models created in the jack-knife procedures. The differences between predictions within each stimulus were between approximately 1 and 8 dB with most differences in the 2 to 4 dB range. This analysis shows that the model was not skewed by outlying subjects and could be expected to provide similar predictions in a new group of normal hearing subjects.
### Table 5.1

Panel A depicts the mean predicted behavioral thresholds for pure tones based on the Cone-Wesson et al. (2002b) regression formulae using ASSR thresholds from the current study. Measured behavioral thresholds for pure tones from the current study are shown in parentheses. Panel B shows the predicted behavioral thresholds for pure tones based on the Rance and Rickards 2002 regression formulae again using ASSR thresholds from the current study. Measured behavioral thresholds for pure tones from the current study are shown in parentheses.

#### Panel A

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>500 Hz</th>
<th>1000Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>48.6(19.6)</td>
<td>46.4(14.0)</td>
<td>51.6(18.1)</td>
<td>57.1(11.6)</td>
</tr>
<tr>
<td>MM</td>
<td>46.4(19.6)</td>
<td>35.7(14.0)</td>
<td>48.3(18.1)</td>
<td>56.0(11.6)</td>
</tr>
<tr>
<td>TBM</td>
<td>47.8(19.6)</td>
<td>39.0(14.0)</td>
<td>43.7(18.1)</td>
<td>43.7(11.6)</td>
</tr>
<tr>
<td>TBS</td>
<td>40.3(19.6)</td>
<td>32.3(14.0)</td>
<td>34.6(18.1)</td>
<td>28.7(11.6)</td>
</tr>
</tbody>
</table>

#### Panel B

<table>
<thead>
<tr>
<th>Stimulus Type</th>
<th>500 Hz</th>
<th>1000Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td>34.4(19.6)</td>
<td>22.1(14.0)</td>
<td>27.6(18.1)</td>
<td>40.8(11.6)</td>
</tr>
<tr>
<td>MM</td>
<td>32.2(19.6)</td>
<td>10.5(14.0)</td>
<td>23.0(18.1)</td>
<td>39.8(11.6)</td>
</tr>
<tr>
<td>TBM</td>
<td>33.6(19.6)</td>
<td>14.0(14.0)</td>
<td>16.9(18.1)</td>
<td>27.7(11.6)</td>
</tr>
<tr>
<td>TBS</td>
<td>26.1(19.6)</td>
<td>6.7(14.0)</td>
<td>4.5(18.1)</td>
<td>13.1(11.6)</td>
</tr>
</tbody>
</table>
5.5 Modeling of Cochlear Excitation

Computer generated models of cochlear excitation were generated utilizing the various stimuli in the current study to demonstrate that the different stimuli result in different activation of the cochlea and would then presumably result in differing activity in the auditory nervous system which would then be reflected in the response pattern for an ASSR. Showing differences in cochlear excitation in the model supports the hypothesis that different thresholds based on stimulus type should be expected when all other controllable recording parameters are the same. A more robust excitation pattern in the cochlea can be expected to result in stronger response and thus a lower ASSR threshold.

The excitation patterns for the MM and TB stimuli were very similar and neither clearly stood out as generating greater excitation in the models. However, the SAM stimuli excitation patterns were clearly smaller and less diffuse than those of the MM and TB stimuli. Given the smaller and narrower response pattern the SAM stimulus could reasonably be predicted to generate higher thresholds than the MM and TB based on these models. The similarity of model output for the MM and TB stimuli is not surprising given their spectral similarity. The ROEX models based on MM and TB stimuli were very similar and would lead one to expect a lower threshold for MM/TB stimuli relative to a SAM stimuli but do not support the MM and TB stimuli having different response patterns. Differences observed in the data would therefore be something transparent to the model.
The measured data were, in general, consistent with the output of the ROEX model given that the MM and TB stimuli were consistently better (lower) than those generated with a SAM stimulus but not especially different from one another.

5.6 Conclusions

Consistent with previous work (e.g., John et al., 1998; Dimitrijevic, et al., 2002; Picton et al., 2003) the MM stimuli resulted in lower thresholds than a SAM stimulus; although in the current study the differences did not reach significance. The TBM stimuli also produced lower thresholds than the SAM stimuli, again not significantly so except at 2000 Hz and 4000 Hz. The TBM stimuli did not elicit a significantly lower threshold compared to the MM stimulus in the multiple simultaneous presentations contrary to what was expected but in agreement with the ROEX model results. Given these results the sharper onset envelope of the TB stimulus does not appear to be a major factor in eliciting the ASSR. This finding is more in line with the physiological based findings of Kiren et al. (1994) showing separate generators for ABR and ASSR rather than the ABR and ASSR having the same generators as was suggested (without any evidence) by Stappells et al. (2004).

Behavioral thresholds for pure tone and ASSR stimuli were demonstrated to have similar audibility in almost all cases making it a reasonable assumption to use thresholds from a modulated stimulus to predict the threshold for a pure tone. Threshold predictions with all three correction methods were similar but the regression analyses (linear and quadratic) generally provided closer predictions than did the mean difference subtraction method. The quadratic regression did not improve the prediction accuracy substantially
over the linear regression model so there is little justification in switching from a linear to quadratic model. The most important factor in choosing a method for making behavioral threshold estimates with ASSR data is to ensure that the data are collected using the same response detection parameters and the same stimuli as were used to create the regression equation applied to the data.

The individually presented TB stimuli resulted in lower thresholds for all frequencies but these differences were relatively small for the most part (7 to 11 dB) but the presence of apparent stimuli interaction at low stimulus levels calls into question the appropriateness of using multiple simultaneous TB stimuli. The current study only included normal hearing subjects so further research is necessary using hearing impaired subjects to determine if these findings hold true in hearing impaired ears. Further study is needed to determine if the use of multiple simultaneous TBs is appropriate.

The current study shows that in the mean data ASSR based predictions of behavioral threshold are reasonably accurate. However, when the individual data is examined while some subjects’ behavioral thresholds were predicted with a great deal of accuracy other subjects' threshold predictions were off by a considerable margin. In clinical practice this presents a serious problem with using the ASSR as the individual is the concern not the mean data. Inaccurate predictions could lead to the wrong clinical decision possibly resulting in harm being done to the patient or needed intervention not being provided. Furthermore the fact that persons known to have normal hearing sensitivity had high thresholds is a concern as little dynamic range is then available to differentiate a hearing impaired ear from a normal hearing ear. While ASSR has potential as a diagnostic tool in audiology it has been rushed into service and is not ready
to be a stand alone test until it has been demonstrated to be a more consistent predictor of behavioral threshold at the level of the individual.

5.7 Implications for Future Research

Generalizing information or conclusions based on the current study is limited due to the small sample size and exclusion of any hearing impaired subjects. The regression equations used to predict behavioral threshold in the current study would only be valid with data collected using the same protocols and stimuli. This is an important finding as the system used in the current study, which is commercially available, has built in regression equations from a published study which used a different stimuli and response detection algorithm so the predicted thresholds based on these equations are likely not accurate and could result in inappropriate clinical intervention.

This study certainly would have been strengthened had the behavioral and ASSR measures targeted the same point on the psychometric function. While under the best case scenario the difference between the behavioral thresholds as measured and what behavioral thresholds might have been had they also targeted the 50% point on the psychometric function only differs by a few (3 to 4 dB) as shown by Watson et al (1972) even within their data the variability was considerable so the difference could potentially be much greater. It should also be noted that the subjects in the Watson et al. (1972) study were highly practiced (thousands of threshold searches) whereas the subjects in the current study, while mostly experienced psychoacoustic listeners, had only a handful of training trials specific to the current study prior to study data being collected making it
more likely that the differences between the 50% and 79% points on the psychometric was larger than under the ideal scenario.

The psychometric function for an ASSR threshold is also an unknown which further complicates the comparison of the two measures. The function is clearly shifted in sensitivity as ASSR thresholds are higher than behavioral thresholds but the shape of the curve has not been investigated to date. Targeting the same point on the psychometric function with the two measures becomes moot if the shape of the ASSR psychometric function differs from the behavioral psychometric function. Determining the shape of the psychometric function for an ASSR would be an avenue of future research and could lead to improvements in predictive ability of the test or at least a better understanding of the relationship between behavioral and electrophysiological thresholds.
BIBLIOGRAPHY


Burkard, R. & Szalda, K. (2004). Two-tone versus toneburst stimuli to elicit the ASSR to single and paired frequency components: inferior colliculus (IC) and auditory cortex (AC) responses, pre and post carboplatin. Poster presented at Mid-Winter meeting of the Association for Research in Otolaryngology, Baltimore, MD.


APPENDIX A

Preliminary Qualification Questionnaire/Screening Forms
Subject Case History Qualification Questionnaire:

1. Do you currently have or recently had (past 2 weeks) an ear infection?

2. Do you currently or have you recently had (within the past week): a cold, the flu or other upper respiratory illness?

3. Do you have a history of chronic middle ear disease?

4. Have you ever had surgery on your either of your ears? (Other than PE tubes as a child)

5. Have you been diagnosed with: Meniere’s disease, sudden onset hearing loss or any other inner ear disease?

6. Do you have a history of dizziness or imbalance due to inner ear disease?

7. Do you have any neurological conditions/diseases?
Subject Data Sheet

Subject ID: _______________                  Ear Tested: ________
Age/Gender: ____________

Otoscopy: (Y)es or (N)o
WNL: ______

Tympanometry: (Y)es or (N)o

Session 1:      WNL: _______
Session 2:      WNL: _______
Session 3:      WNL: _______

Hearing Screening
(P)ass or (F)ail at 15 dB HL for all frequencies

<table>
<thead>
<tr>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
<th>8000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

Cochlear Excitation Model Outputs
This appendix contains the outputs from a computer program used to model the excitation that would occur on the basilar membrane when the various stimuli used in the current study were presented to the ear. Basilar membrane excitation modeling was completed with the ROEX program developed by Glasberg and Moore (1995). Outputs from the ROEX model are presented in two manners. First the ASSR stimuli used in the current study are plotted by frequency so that all stimuli types of the same carrier frequency are plotted on the same figure to display the differences in spectral spread for each stimulus. Two stimulus levels are presented to demonstrate the changes that occur in spectral spread as stimulus intensity decreases. The two levels selected are meant to represent levels where a normal hearing ear would have a supra-threshold response and at a level near expected threshold for a normal ear. The second way the ROEX outputs are displayed is with the multiple simultaneous stimuli presentation where four carrier frequencies for each stimulus type are plotted on the same figure.
Figure B.1. Excitation pattern models for all 500 Hz stimuli with a stimulus presentation level of 60 dB SPL (top) and 30 dB SPL (bottom)
Figure B.2. Excitation pattern models for all 1000 Hz stimuli with a stimulus presentation level of 60 dB SPL (top) and 30 dB SPL (bottom)
Figure B.3. Excitation pattern models for all 2000 Hz stimuli with a stimulus presentation level of 60 dB SPL (top) and 30 dB SPL (bottom)
Figure B.4. Excitation pattern models for all 4000 Hz stimuli with a stimulus presentation level of 60 dB SPL (top) and 30 dB SPL (bottom)
Figure B.5. Excitation patterns for multiple simultaneously presented tonebursts at levels of 60 dB SPL (top) and 30 dB SPL (bottom).
Figure B.6. Excitation patterns for multiple simultaneously presented mixed modulation tones at levels of 60 dB SPL (top) and 30 dB SPL (bottom).
Figure B.7. Excitation patterns for multiple simultaneously presented SAM tones at levels of 60 dB SPL (top) and 30 dB SPL (bottom).
APPENDIX C

Individual Subject Data
Figure C.1. Subject 1 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.2. Subject 2 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.3. Subject 3 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.4. Subject 4 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.5. Subject 5 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.6. Subject 6 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.7. Subject 7 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.8. Subject 8 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.9. Subject 9 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.
Figure C.10. Subject 10 ASSR thresholds and behavioral thresholds for pure tones and modulated (ASSR) stimuli.