Auditory-Nerve Fiber Responses to Amplitude Modulated Tones and Multi-Tonal Stimuli

by

Holden Cheng
B.A., Physics
University of California, Berkeley, 1999

Submitted to the Harvard-MIT Division of Health Sciences and Technology in Partial Fulfillment of the Requirements for the Degree of Master of Science in Health Sciences and Technology

at the
Massachusetts Institute of Technology

September 2005

© 2005 Massachusetts Institute of Technology
All rights reserved

Signature of Author..............................................................................................................

Harvard-MIT Division of Health Sciences and Technology
August 15, 2005

Certified by ..............................................................................................................................

John J. Guinan, Jr.
Associate Professor of Otology and Laryngology, Harvard Medical School
Affiliated Faculty in Health Science and Technology, Harvard-MIT
Thesis Supervisor

Accepted by ............................................................................................................................

Martha L. Gray
Edward Hood Taplin Professor of Medical and Electrical Engineering
Co-Director, Harvard-MIT Division of Health Sciences and Technology
Auditory-Nerve Fiber Responses to Amplitude Modulated Tones and Multi-Tonal Stimuli

by

Holden Cheng

Submitted to the Harvard-MIT Division of Health Sciences and Technology on August 15, 2005 in partial fulfillment of the requirements for the Degree of Master of Science in Health Sciences and Technology

ABSTRACT

In normal-hearing ears, sound waves are amplified within the cochlea and a small fraction of the sound energy travels backward out into the ear canal, producing sounds known as “otoacoustic emissions” (OAE) that can be measured with a sensitive microphone. One class of OAE, called “stimulus-frequency-otoacoustic-emissions” (SFOAEs), has been hypothesized to be produced by a process known as “coherent reflection filtering” (CRF). The CRF theory provides a prediction between the SFOAE group delay and the group delays of tone responses on the basilar membrane within the cochlea. Using single and multiple-tone stimuli, we collected data from the firing patterns of single auditory-nerve-fibers (ANFs) from which basilar-membrane tone-response group delays can be calculated for both high and low best-frequency (BF) positions along the basilar membrane. These calculated basilar-membrane group delays were compared to published SFOAE group delays. Our results suggest that group delays calculated from the tip, the lower-frequency tail, or the above-BF region of ANF tuning curves do not match the CRF theory prediction.

In obtaining the data to test the CRF theory, we used two methods for obtaining ANF group delays at frequencies above BF: a previously published method and a simpler new method based on the same principle. Surprisingly, the two methods produced different results. Control measurements suggest that the previously published method does not do what it was expected to do.

Thesis Supervisor: John J. Guinan, Jr.

Title: Associate Professor of Otology and Laryngology, Harvard Medical School Affiliated Faculty in Health Science and Technology, MIT
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>1</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>I. Background and Significance</td>
<td>4</td>
</tr>
<tr>
<td>A. Introduction</td>
<td>4</td>
</tr>
<tr>
<td>B. Group Delays and the Theory of Coherent Reflection Filtering</td>
<td>5</td>
</tr>
<tr>
<td>C. Phase Plateaus with AM Tones</td>
<td>7</td>
</tr>
<tr>
<td>II. Research Methods</td>
<td>9</td>
</tr>
<tr>
<td>A. Animal Preparation</td>
<td>9</td>
</tr>
<tr>
<td>B. Stimulus and Data Collection</td>
<td>10</td>
</tr>
<tr>
<td>C. Multi-tone Phase Analysis</td>
<td>11</td>
</tr>
<tr>
<td>D. Two-tone Method</td>
<td>13</td>
</tr>
<tr>
<td>E. AM-tone Method</td>
<td>14</td>
</tr>
<tr>
<td>III. Results</td>
<td>16</td>
</tr>
<tr>
<td>A. Group Delays</td>
<td>16</td>
</tr>
<tr>
<td>B. Group Delays across CF</td>
<td>18</td>
</tr>
<tr>
<td>C. Multi-tone vs Two-tone</td>
<td>25</td>
</tr>
<tr>
<td>D. Phase Plateaus with AM tones</td>
<td>30</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>32</td>
</tr>
<tr>
<td>A. ANF vs SFOAE Group Delays</td>
<td>32</td>
</tr>
<tr>
<td>B. Differences in Group-Delay Methods</td>
<td>33</td>
</tr>
<tr>
<td>References</td>
<td>35</td>
</tr>
</tbody>
</table>
I. Background and Significance

A. Introduction

The sound processing in the mammalian auditory system involves many steps in the periphery (Figure 1) before being interpreted in the brain. In the presence of a sound source, the propagating sound waves are collected by the outer ear and travel into the ear canal, where they vibrate the eardrum. The sound pressure at the eardrum is conducted by the middle ear ossicles to a fluid-filled hearing organ, known as the cochlea. Motion of the last ossicle, called the stapes, produces a pressure wave (called the “fast wave”) that spreads throughout the fluid of the cochlea at the speed of sound. The fast wave also initiates a slower wave (called the “traveling wave”) due to the pressure difference across the basilar membrane (BM). The traveling wave propagates from the cochlear base to the cochlear apex and has a tuned response such that high-frequency sound produces the largest BM motion in the cochlear base and low-frequency sound produces the largest BM motion in the cochlear apex. Once the traveling wave reaches its maximum amplitude at the characteristic place, it abruptly collapses afterward at a position known as the cutoff region. In the classical view, BM motion from the traveling wave produces in-phase motion of the other structures of the organ of Corti and leads to bending of inner-hair-cell stereocilia which leads to excitation of auditory-nerve fibers (ANFs).
B. Group Delays and the Theory of Coherent Reflection Filtering

In this study, we gathered group delays across CFs to provide additional insight on the “theory of coherent reflection filtering” in the cochlea. Proposed by Shera and Zweig (Shera and Zweig, 1993; Zweig and Shera, 1995), the theory of coherent reflection filtering (CRF) states that at low-to-moderate sound levels, evoked otoacoustic emissions (OAE) in the ear are generated by “coherent scattering” of the traveling wave off small and random perturbations in the cochlea (Shera and Zweig, 1993; Zweig and Shera, 1995; also see: Shera and Guinan, 2003). One of the main predictions of this theory is that the delay of stimulus frequency otoacoustic emission (SFOAE) should be approximately equal to twice the BM group delay at the tuning
curve tip with CF equaled to the stimulus frequency (Shera and Zweig, 1993; Zweig and Shera, 1995; Shera and Guinan, 2003). The hypothesis is that SFOAE delays need to account for the round-trip traveling time (forward and backward), while the BM group delay takes only the forward traveling time. Mathematically, it can be expressed as: $\text{Delays}_{\text{SFOAE}} = 2 \times \text{Delay}_{\text{BM}}$.

The validity of the CRF theory has been tested recently by Siegel et al (2005) when they compared group delays of BM measurements, ANF measurements, and SFOAEs across different species. They concluded that in the chinchilla and in other species, the group delay of SFOAEs is does not equal to twice the BM group delay over most areas of the cochlea (Siegel et al, 2005). To provide further insight into this problem, we obtained ANF data in cats to calculate BM group delays across CF, and to compare them to SFOAE group delays. We hypothesized that the reflection source responsible for the group delays in SFOAEs isn’t located in the tip regions of the tuning curve, as presumed in the coherent reflection filtering theory. We addressed this hypothesis by measuring group delays in the tuning curve tip, low-frequency tail and high-frequency upper-edge regions, and comparing them to the published cat SFOAE group delays from Shera and Guinan (2003).

The group delays at all three regions of the tuning curve were calculated by using van der Heijden and Joris (2003) multi-tone method to extract phases and group delays from ANF responses (as in Cheng, 2005). This method is potentially useful for high-frequency stimuli because it circumvents the ANF phase-locking limitation (to stimuli over 4 kHz) by presenting a multiple-tone complex that produces low-frequency beats, which form the envelope of the stimuli. The beat frequency carries the difference in phase between two interacting tone responses. The envelope can be entrained by the ANFs and measured, and the mathematical reconstruction of the beat phases yields the relative phases and group delays of the original tones.
(see methods for more details). The ability to extract group delays by calculating the phase difference divided by the frequency difference of nearby phase points allows us to find the group delays across ranges of fibers’ CFs.

As an alternate method, the group delays at the high-frequency edges were also derived from a new method using two tones. Based on the same principle as the multi-tone method, the two-tone method uses two simultaneous tones to create a beat frequency which encodes the phase differences between the interacting tones. Comparison between the two methods should provide further insight on the validity of our methods.

C. Phase Plateaus with AM Tones

This study also extends a previous topic dealt with in Cheng, 2005, which searched for neural correlates of phenomena found in BM motion studies. The question was: “Are there phase plateaus in ANF responses at high sound levels and above CF?” BM measurements in the cochlear base have been shown to exhibit phase plateaus at frequencies well above CF (Ruggero, 1997; Robles and Ruggero, 2001; Ruggero et al, 2000). The phase plateau is thought to occur when the traveling wave changes into an exponentially decaying evanescent wave at a point of total internal reflection (Watts, 2000). Determining the existence of a corresponding phase plateau in ANF will show if the evanescent wave is capable of driving inner-hair-cell stereocilia and ANFs.

In Cheng (2005), we found what might be evidence of phase plateaus in ANF responses to AM tones at high sound levels and at frequencies far above CF, but these measurements lacked adequate controls. We used AM tones as the stimulus of choice because the multi-tone method of van der Heijden and Joris (2003, 2005) has problems when its stimuli at high sound
levels produce distortion products, which obscure the results. In this report, we carried out the same procedure but with better controls. We used two different earphones with different distortion properties to produce supposedly identical AM-tone stimuli on the same ANF unit.

The results show that the previous measurements of a potential phase plateau in ANF firing was due to distortions produced by the earphone.
II. Research Methods

17 cats weighing between 4 to 7 pounds have been used for these experiments. The experiments were done in the Eaton-Peabody Laboratory (EPL) of Auditory Physiology. All experiments were in compliance with protocols approved by the Committee on Animal Care at the Massachusetts Eye and Ear Infirmary.

A. Animal Preparation

The experimental methods follow several published experiments involving ANF recording (such as Stankovic and Guinan, 1999, and Kiang et al, 1965). Anesthesia on cats was done by intra-peritoneal injection of Dial in urethane. The initial dose was 75mg per kilogram of weight, and supplemental boosters at 10% of the initial dose were used if there was any sign of a toe-pincher reflex. A tracheotomy was performed and a trachea tube was inserted for optional connection with a Harvard Apparatus animal respirator. The animal was positioned lying down with the head held erect and placed in a soundproof room. The bulla cavities on both sides were then exposed, revealing the middle ear cavities and the round windows. Silver electrodes were placed on or near each round window to measure cochlear compound action potentials. The animal’s temperature was monitored by a rectal thermometer and maintained near 38°C with a heating pad.

The posterior area of the parietal bone of the skull was exposed, followed by aspiration of the cerebellum to reveal the cochlear nucleus. Cotton balls and a small metal retractor were used to push the cochlear nucleus medially, exposing the auditory nerve. 3 M KCl filled glass
micropipettes mounted on a remote-control micro-manipulator were used to search for and record from auditory nerve fiber units.

B. Stimulus and Data Collection

The acoustic assembly consisted of a 1-inch condenser earphone and a 1/4-inch condenser microphone in a calibrated probe tube. In some experiments a second earphone, the DT48 dynamic earphone, was connected by a 3-inch tube. The tip of the assembly was placed inside the external meatus a couple of millimeters away from the eardrum. Stimulus generation and data collection were controlled by a Windows PC computer running National Instruments LabView 6.1 or 7 and MathWorks Matlab 6.1 or 7. The animal’s status of vitality (i.e., animals’ heart rate, breath rate, CO2 level, and EEG) was monitored by a Macintosh G3 personal computer with LabView 6.0. The vitality computer continuously collected data within a 15-second interval and displayed the resulting averages versus time for the previous 2 hours. An alarm sounded when the vitality data fell outside the animal’s physiological range.

Glass micropipettes at impedances from 10–30 MΩ were used to record action potentials in single auditory nerve fibers. We searched for ANF units by remotely varying the depth of microelectrode penetration while presenting broadband noise at about 75–85 dB SPL. When an ANF unit was found, the trigger level and gain were adjusted to ensure the best triggering (i.e., the least amount of extra and/or missed spikes). Only data from ANF units with perfect or near perfect triggering were used for analysis. A tuning curve was measured using 12 to 30 frequency steps per octave, followed by a measurement of the spontaneous rate (SR) within a 15-second interval. The characteristic frequency (CF) of the unit was determined by finding the lowest threshold in the tuning curve.
C. Multi-tone Phase Analysis

The limitation in most ANF studies is that ANFs cannot follow the fine time structure of high-frequency stimuli. ANFs’ response synchrony below 1 kHz is good, but starts to decline above 1 kHz and is gone by 4–5 kHz. This phase-locking limitation hinders the ability to extract accurate phase information from ANF units at high frequencies. To overcome this ANF temporal limitation, van der Heijden and Joris (2003, 2005) devised a method that uses a complex comprised of multiple tones that produce beats. The beats occur from interactions among each pair of tones, acting on the nonlinearities inherent in the cochlea. The beats were low in frequency and thus able to be encoded by the nerve, even if the individual tones exceed the phase-locking limit (above 4 kHz). For example, two tones at frequencies $f_1$ and $f_2$ produce an envelope with beat frequency, $f_2 - f_1$. A mathematical equation for the real part of the two tones is,

$$z(t) = \text{Re}[A_1 \exp\{i(w_1t + \phi_1)\} + A_2 \exp\{i(w_2t + \phi_2)\}],$$  \hspace{1cm} (1)

where $z(t)$ is the sum of the complex tone’s waveform, $A$’s are the amplitude, $w$’s are the angular frequencies, and $\phi$’s are the phases. The general waveform can be applied for $N$ number of tones, as shown below,

$$z(t) = \text{Re} \left\{ \sum_{i=1}^{N} A_i \exp\{i(w_it + \phi_i)\} \right\}.$$  \hspace{1cm} (2)

A simple mathematical equation for the interactions between the tones is the square envelope of $z(t)$,
\[ |z|^2 = \sum_{i=1}^{N} A_i^2 + \sum_{m=1}^{N} \sum_{k=m+1}^{N} 2A_m A_k \cos \{(w_k - w_m)t + \phi_k - \phi_m\}. \] (3)

The last term indicates that each possible pair of tones interacts to form a beat frequency of \(w_k - w_m\). Furthermore, the magnitude and phase of the “primary” tone complex (e.g., \(w_k\) and \(\phi_k\)) are related to the envelope magnitude and phase. The result can be summarized as follows,

\[ A_{km} = 2A_k A_m, \] (4)

\[ \phi_{km} = \phi_k - \phi_m \left( \text{mod } 2\pi \right). \] (5)

The beats’ magnitude \((A_{km})\) and phase \((\phi_{km})\) are at the left-hand sides of equations 4 and 5. These are the values extracted from the responses of ANF units to multiple tones. Equations 4 and 5 are then used to reconstruct the relative phases and magnitude of the responses to the primary components (at the right-hand side of the equations). Because there are more beats than there are primaries (see the example below), equations 4 and 5 will yield an overdetermined set of equations. These equations can be solved numerically by the least-square method to determine the best estimates of the original values for the primaries. Van der Heijden and Joris (2003) showed that this method, when applied to low-CF fibers, produces results consistent with direct measurements of response phase from low-frequency multi-tone stimuli.

In our experiments, a continuous tone complex of 4–6 primary frequencies lasting 1 sec was used as the stimulus. The primary tones were separated in frequency to allow unique beats with adjacent frequencies no more than 1 kHz apart so they are below the ANF phase-locking limit. A Fast Fourier Transform (FFT) was applied to the spike histograms to reveal the magnitude and phase of the beats. To obtain a good measurement, a run was repeated to accumulate enough averages so that the beat frequencies’ vector strength was above the Rayleigh
criteria at \( p < 0.001 \) (Mardia, 1972). Also, if there were cubic distortion \((2f_1-f_2)\) components above a criterion, the run was dismissed. Equations 4 and 5 were used to compute the phases and magnitude of the responses to the primaries, respectively. Group delays can be determined by calculating the phase slope between each adjacent phase point. We removed all reconstructed group delays that were negative because the controls done in the Cheng 2005 showed that the negative reconstructed group delays do not match the true group delays.

The relative phases and group delays calculated from van der Heijden and Joris (2003) multi-tone phase-extraction method will be referred to as the “reconstructed phases” (RP) and the “reconstructed group delays” (RGD). The ANF responses also entrain to stimulus frequencies below 4 kHz, and the phases and group delays from these entrainments will be called the “true primary phases” (TPP) and “true primary group delays” (TPGD).

D. Two-tone Method

In certain instances, we used multi-tone series consisting of only two tones for the upper-edge regions of the tuning curve. The two-tone runs had one fixed tone presented around the tuning curve tip, while another tone was presented along the upper edges of the tuning curve. The frequency difference between the near CF tone and the “edge” tone was chosen to be less than 4 kHz. In a typical run, there were a series of edge tones, and each was presented with the same fixed tone. The phase was determined by the beat frequency produced by the edge tone with the fixed tone, and the group delay was determined by the slope between neighboring phase points of the beat responses. This had two advantages over the van der Heijden and Joris (2003) multi-tone method: 1) Higher sound levels could be produced by the acoustic transducer, and 2) any \(2f_1-f_2\) distortion products that is produced will be outside of the fiber’s tuning curve.
E. AM-tone Method

The amplitude modulated (AM) tones were short (200 ms) tone bursts presented with varying carrier frequencies and sound levels, but with the modulation frequency constant at 100 Hz and the modulation depth fixed at 0.3 as in Gummer and Johnstone (1984). Spike rate, phase and synchrony of ANF responses to the AM tones were measured from the spike timing events as in Stankovic and Guinan (2000). As a test of consistency, we also presented identical AM runs using two different acoustical sources. One earphone was the default 1-inch reverse-driven condenser earphone from Brüel and Kjær (B&K), and the other was a dynamic earphone, the Beyer DT48. Each earphone has different acoustical and distortion characteristics. The square-law condenser earphone’s (B&K) output had been supposedly linearized at high sound levels by driving it with an inverse-square function, while the DT48 is capable of much larger output and was generally more linear at high sound levels. The acoustical output and frequency characteristics of each earphone were different (see figure 2), but for frequencies less than 30 kHz and levels less than about 100 dB SPL, they were similar.
Figure 2: In ear acoustical calibration of 1-inch condenser earphone (B&K, ...) and dynamic earphone (DT48, ...) for cat#17.
III. Results

A. Group Delays

We used three criteria to ensure that the data from the van der Heijden and Joris (2003) multi-tone method were adequate. The first criterion was perfect or near perfect triggering of the auditory nerve fibers in response to the stimulus. The second was high signal-to-noise ratio of the beat frequencies by evaluating the synchrony of each beat frequency and comparing to a criterion that was twice the Rayleigh number. If two or more beat frequencies from the same common primary frequency were above the criteria, then that primary frequency was removed. If more than 2 primaries were removed, then the entire run was rejected. The third criterion was a low number of distortion products, such that if 2 or more distortion products were above twice the Rayleigh criterion, the entire run was rejected.

Figure 3 shows composite graphs of group delays from the van der Heijden and Joris (2003) method separated into CF groups. We divided the results according to fiber CFs and plotted group delays as a function of stimulus frequency in reference to CF. In all the panels except for CF-region below 1 kHz, the group delays were calculated by the reconstruction of the primaries using the van der Heijdgen and Joris (2003) multi-tone method. The composite group delays below 1 kHz were taken from the true primary group delays as discussed in the methods. All group delays have a 1-msec nerve-conduction-delay correction to show calculated basilar membrane (BM) group delays.

The figure shows that the group delays were usually longest at CF and fell to shorter group delays at frequencies away from CF. Group delays were also lower at lower CFs (note that the axes on all 6 bottom graphs, 1 kHz and above, are on the same scale).
Figure 3: Composite plots of group delays grouped into 7 categories by CFs. Each point represents the group delay calculated from a pair of adjacent phase measurements; an individual nerve fiber can yield more than one point. For panels with CFs above 1 kHz, we used the multi-tone method to calculate the reconstructed group delays. For the panel with CFs below 1 kHz, we also used the multi-tone data, but extracted the group delays from the true primary phases, not from the beats which were low in frequency and noisy for these CFs. Note that the 6 bottom panels have the same axes scale. All group delays have the 1-msec-nerve-conduction-delay correction.
B. Group Delays across CF

We compared ANF group delays taken at different regions of the tuning curve to stimulus frequency otoacoustic emission (SFOAE) group delays. Figure 4 shows group delays taken at the ANF’s tuning curve tips as a function of CF. To define data that are at tip of the tuning curve, we chose stimulus frequencies that are within 0.03 octave of the unit’s CF. Group delay minus 1-msec nerve conduction delay as a function of CF (green +) is compared with group delays of SFOAEs (thick solid line, Shera and Guinan, 2003) as well as the group delays from ANFs (thin solid line, van der Heijden and Joris, 2003 and 2005).

![Figure 4: Tuning-curve-tip group delays (+) within 0.03 octave of CF compared to SFOAE group delays (Shera and Guinan, 2003) and BM-calculated group delays (van der Heijden and Joris, 2003 and 2005) as calculated and plotted by Siegel et al (2005). Each point represents the group delay calculated from a pair of adjacent phase measurements, plotted at the CF of the unit. All group delays have the 1-msec-nerve-conduction-delay correction.](image-url)
Since our study essentially used the same method as van der Heijden and Joris, the two sets of data measuring group delay data at the tuning curve tips are similar (figure 4). The results shown in Figure 4 support the notion that the delay of SFOAE does not equal twice the delay in the ANF evaluated at the tuning curve tip. Furthermore, the ANF group delay at CFs less than 1.4 kHz is actually longer than the SFOAE group delays. However, above 2 kHz, the SFOAE delays and the ANF group delays overlap considerably.

We graphed group delays in the low-frequency-tail of the tuning curve by taking individual units and evaluating the group delays at stimulus frequencies more than 0.25 octave below the unit’s CF. The delays at the low-frequency tail were generally lower than at the tip (see figure 3). Figure 5 compares low-frequency-tail group delays (red □) to first-peak-click response latencies (○ and △) from Lin and Guinan (2000) and the SFOAE group delays from figure 4. Three low-CF units had group delays that were calculated from their true primary phases (◇). Click latencies were measured as the time from the onset of the click stimuli to the onset of the first peak in the ANF units’ peri-stimulus time histogram response (Lin and Guinan, 2000), minus 1 msec for nerve-conduction-delay. The lower-frequency-tail group delays and first peak click latencies are plotted as a function of units’ CF. First peak latency and tail-region group delays match well for CFs over about 3 kHz. At lower CFs below 2 kHz, the tail-region delays are considerably higher. In comparison to group delays of SFOAEs, tail-region group delays are shorter for CFs above 1 kHz.
Figure 5: A comparison of low-frequency-tail group delays, first peak click latencies (from Lin and Guinan, 2000), SFOAE group delays from (from Shera and Guinan, 2003), and BM-calculated group delays (from van der Heijden and Joris, 2003). Lower-frequency-tail group delays are divided into two groups – ones derived from the reconstructed phases, and the other from the true primary phases. For the lower-frequency-tail group delays, each point represents the group delay calculated from a pair of adjacent phase measurements, plotted at the CF of the unit. True primary phases are only applicable for CF below 4 kHz (see methods). All group delays have the 1-msec-nerve-conduction-delay correction.
We calculated the group delays in the above-CF-region of the tuning curve with two methods. For the first method, we used the composite plots (figure 3) and evaluated the average group delays for stimulus frequency over 0.18 octaves above CF. These group delays are not a function of individual units, so the composite group delays were plotted as a function of the average CFs. The second method is to use two-tone stimuli along the above-CF upper edges of the tuning curve. We used only phases with statistically significant synchrony (as defined by Stankovic and Guinan, 2000) and evaluated all group delays for stimulus frequencies 0.18 octaves or more above CF. The group delays derived from this method are plotted as a function of unit CF. Both group delays will be referred to as the high-frequency-tail group delay, but they differ in their method of acquisition – “composite” or “two-tone.”

Figure 6 shows the high-frequency-tail group delays evaluated by using the composite plots (blue ⭐) or using the two-tone method (blue ●). In comparison to each other, both group delays are noticeably different for CFs less than 8 kHz, even though they both supposedly measure group delays in the high-frequency-tail region. The high-frequency-tail group delays calculated from the composite plots are shorter than the SFOAE group delays, while the group delays calculated from the two-tone method are the same or sometimes longer than the SFOAE group delays.

Figure 7 is a compilation of group delays that are generally shorter than the SFOAE group delays. Included are the first-peak click latency data (○ and △) from Lin and Guinan (2000), low-frequency-tail group delays (□), and high-frequency-tail group delays derived from the composite plots (⭐).

In figure 8, we took the group delays from figure 7 and multiplied them by a factor of 2 to compare them with the SFOAE group delays. Some points below 1 kHz are off the scale and
thus not seen. The turquoise-color symbols are the doubled counterparts of the respective group delays in each graph. None of these delays appear to fit the hypothesized relationship (i.e., $\text{Delay}_{\text{SFOAE}} = 2 \times \text{Delay}_{\text{BM}}$).

Figure 6: High-frequency-tail group delays calculated from two methods: 1) taking group delays from the composite plots from figure 3 (★), and using the 2-tone methods (●). The composite group delays are averages of group delays 0.18 octaves above CF in the CF regions defined in Fig. 3, each plotted at the average CFs for units that contributed to the average group delay. For the 2-tone group delays, each point represents the group delay calculated from a pair of adjacent phase measurements, plotted at the CF of the unit. All group delays have the 1-msec-nerve-conduction-delay correction.
Figure 7: Compilation of group delays from figures 4 to 6 that are below SFOAE group delays. All group delays have the 1-msec-nerve-conduction-delay correction.

Figure 8 (next page): Comparison of “doubled” group delays to the SFOAE group delays. Group delays were doubled (turquoise) for click latencies, low-frequency-tail group delays, and high-frequency-tail group delay (from composite plots), and plotted with the SFOAE and calculated-BM group-delay curves from figure 4. All group delays have the 1-msec-nerve-conduction-delay correction.
C. Multi-tone vs Two-tone

The plot in figure 6 illustrates a discrepancy between the two methods for measuring group delays in the high-frequency-tail region – one using the van der Heijden and Joris multi-tone method and the other using the two-tone method. To explore this discrepancy, we compared the group delays calculated from the two methods on individual single ANF units.

Figure 9 and 10 illustrate this comparison for two different ANF units. First, panel D in both figures illustrates the presentation of the two-tone stimuli inside the ANF tuning curve. The tone near CF (△) is fixed and presented on every run, and is presented simultaneously with one of the “edge” tones (○) that is picked at random. The two tones produce a firing rate (panel B) and a single difference-frequency beat between them, which is entrained by the ANF. Panel A shows the synchrony index and phases of the beat frequency as a function of the edge tones frequencies. Phases that are statically significant in synchrony (as defined by Stankovic and Guinan, 2000) are displayed with circles. The group delays were calculated from the slope of the phase points and plotted in panels C. Figure 9 shows that all 6 phase points are statistically significant, while only 3 phase points are significant in figure 10, thus giving only 2 group delays (red ●). Panel C also shows the group delays that were calculated from the multi-tone method near the frequencies of the two-tone stimuli (□).

In both figures, group delays calculated from the multi-tone method were shorter than the group delays calculated from the two-tone method. In figure 9, the group-delay difference between the two methods is about 1.6 milliseconds at 3.4 kHz. In figure 10, the group-delay difference near 4.05 kHz varies from 0.5 to about 1 millisecond.
Figure 9: Group delay comparison using the two-tone and the multi-tone methods. Panel D is the representation of the two-tone stimuli. On each trial, the fixed near-CF tone (△) is paired with one edge tone (○). Panel B shows the firing rate of both tones together and of the near-CF tone alone as a function of edge tones frequencies. Panel A shows the phases and synchrony indices of the beat frequency between the two tones. Statistically significant phase and synchrony points are denoted with o’s and x’s, respectively. Panel C shows the calculated group delays from the two-tone method (2T) in addition to the group delays from a multi-tone run (MT) on the same unit and near the same stimulus frequencies. Group delays from both methods have the 1-msec nerve-conduction-delay correction.
Figure 10: Group delay comparison between the two-tone and the multi-tone methods. Panels are the same as in figure 9, except that panel B also shows spontaneous rate (blue horizontal line, not shown in figure 9 because spontaneous rate is near zero), and panel C labels the two statistically significant two-tone group delays as red-filled circles (●).
As a control, we compared phases of the ANF responses between multi-tone and single-tone stimuli at the high-frequency-tail region in individual units. Units with CFs below 4 kHz were chosen so that there would be phase locking in the responses to the single tones. We first presented a multi-tone run on a unit, followed by the individual tones of the same multi-tone complex (or frequencies within 1% of the multi-tone frequencies). Figure 11 illustrates the phases derived from the two methods. The phase curves have been shifted such that the first phase points overlap (the phases derived by the multi-tone method have an arbitrary reference). The phases from the single tone method (澳大) were not the same as those from the multi-tone method (□), especially at higher levels and frequencies. In the right column, the phases are similar in the first 3 points, but not in the last two. In the left column, none of the phase points after the first one overlap. The steeper phase curves from the single tone method yield longer group delays than the delays from the multi-tone method. Least-square-error estimates of the slopes indicate group-delay differences between the two methods of at least 1 msec (Figure 11).
Figure 11: Comparison of phases between the multi-tone and single-tone method. Tuning curves of two fibers (left and right columns) are shown with the multi-tone runs (bottom). The phases of the multi-tone run are compared with the phases of single tones that were separately presented at frequencies and levels near the individual components of the multi-tone complexes. Least-square-error fit of the phase points yields slopes that are listed in the top panels.
D. Phase Plateaus with AM tones

In a previous study, we found what might be evidence of a phase plateau in ANF responses to AM-tones at high levels and at frequencies far above CF (Cheng, 2005). As a check for consistency, similar measurements of auditory-nerve-fiber (ANF) responses were made using identical AM-tone stimuli presented by two different earphones. One earphone was the default 1-inch Brüel and Kjær (B&K) reverse-driven condenser earphone, and the other earphone was the Beyer DT48 dynamic earphone. The synchrony indices of the ANF responses to the two different earphones are shown in figure 12 in two different ANF units and cats. Orange (thin) lines are ANF responses to the 1-inch condenser earphone, while blue (thick) lines are ANF responses to the dynamic earphone. The x’s in the synchrony plots indicates a phase synchrony of significance as described by Stankovic and Guinan (2000).

The synchrony plots show that the ANF responses to the two different earphones presenting the same AM stimulus are not the same. In particular, there is a lack significant synchrony at high levels in the ANF responses to the lower-distortion dynamic earphone. This control study indicates that the evidence of a phase plateau found previously is likely due to distortion from the condenser earphone (this conclusion will not be discussed further).
Figure 12: Comparison between auditory-nerve-fiber responses from two earphones presenting the same AM tones at high sound levels (in two different units and cats). Panels show synchrony indices as a function of AM-tone sound levels for the 1-inch condenser earphone (---) and the dynamic earphone (—). Synchronies that are statistically significant (as defined by Stankovic and Guinan, 2000) are marked with x’s.
IV. Discussion

A. ANF vs SFOAE Group Delays

The group delays derived from the slopes of the phase-versus-frequency functions of auditory-nerve-fiber (ANF) responses depict the physiological delay in auditory stimulus processing. These delays have a frequency-dependent component that is commonly assumed to reflect the travel times of a traveling wave from the base of the cochlea to the apex (Hillery and Narins, 1984, 1987). Figures 3 to 8 are consistent in showing that the travel time is noticeably longer at lower frequencies. Presumably this is because the cochlea is arranged such that the low-frequency CFs is located further away from the high-frequency CFs at the base.

The theory of coherent reflection filtering predicts that the group delays in stimulus-frequency otoacoustic emissions (SFOAEs) with stimulus frequency equaled to the basilar membrane (BM) characteristic frequency (CF) should be approximately equal to twice the group delay measured in the BM. The reason is because SFOAE group delays takes into account the travel time of the traveling wave in both the forward and backward directions. We attempted to find possible mechanisms that would satisfy this prediction by comparing group delays in different regions of the ANF tuning curve to published SFOAE group delays. In figure 4, we showed that the calculated BM group delay measured from ANF responses near the tuning-curve tips do not support this prediction. In fact, figures 5 to 8 (most notably figure 8) show that group delays measured in other regions of the tuning curve also do not fit the prediction.

Assuming that the results in figure 3 to 8 are valid, the findings suggest that perhaps certain aspects of the theory of coherent reflection filtering are flawed. In figure 4, the group
delays of the calculated BM group delay match the group delays of SFOAE for frequencies above 2 kHz. If the results are valid, then this suggests that one forward traveling wave and a near-instantaneous backward traveling wave would give the correct SFOAE group delay for frequencies of 2 kHz and higher. Other plots (figures 5-8) suggest that none of the group delays derived from other parts of the tuning curve match the SFOAE group delays. A possibility is that more than one source is responsible for the SFOAE delay, such that the forward traveling wave and the backward traveling wave originate from different parts of the tuning curve. Another possibility is that there is a wave along the cochlea at CFs of 2 kHz and lower that accounts for the SFOAE group delays at these frequencies, but that auditory-nerve responses of this wave were not captured by our methods.

B. Differences in Group-Delay Methods

The validity of the van der Heidjen and Joris (2003) multi-tone method to calculate group delays from ANF responses was put into question when we compared it to two other methods. At frequencies above CF, the multi-tone method produced noticeable shorter group delays in comparison to the delays from the two-tone and single-tone methods (figures 9, 10, and 11). The differences in group delays were usually 1 msec or more.

The findings reveal two main points. 1) As shown in figure 11, the ANF responses to the individual tones in a multi-tone complex are not the same as when these tones are presented as single tones. 2) The difference between the multi-tone and two-tone methods is most surprising because both rely on the same principle (i.e., creating low frequency beats that encode phase differences) to calculate group delays. One possible explanation is that in the multi-tone method,
the rather large number of (5 or more) primary frequencies produce more mutual suppressions than just using one or two tones, and that this can affect the group delays.

Future experiments and analysis are needed to verify and more accurately quantify the discrepancy between results obtained with the multi-tone method and the other methods, and also to understand why they are different. Possible studies may include comparing the two-tone method to the single-tone method, and exploring group delay differences in other parts of the tuning curve. Also, the study of the coherent reflection filtering theory can benefit from more group-delay data for CFs below 1 kHz. Since ANF units below 1 kHz have adequate phase-locking, the use of the possibly flawed multi-tone method is not required.

Acknowledgement

I would like to thank my advisor, John J. Guinan, Jr., in making this thesis possible. I would also like to thank my fellow colleagues in EPL and in the Speech and Hearing Science Program for their critic and help, especially Waty Lilaonitkul, Bradford Backus, Leonardo Cedolin, and Tony Miller. Also, thank you Connie Miller for doing a great job in animal preparation. And lastly, I thank my family and Charmsak “Touch” Thongsornkleeb for your support.
References


